May 3-5, 2019
Dhaka, Bangladesh

2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT 2019)



Proceedings of

2019 1st International Conference on Advances in Science, Engineering and Robotics Technology 2019 (ICASERT 2019)

May 3-5, 2019
Dhaka, Bangladesh

Table of Contents

Design of CMOS On-Chip Transformer Coupled Matching Network for Millimeter-Wave Amplifiers with Optimal Chip Area 1
Md Sazzad Hossain, Minoru Fujishima, Takeshi Yoshida, Shuhei Amakawa and Md. Mamunur Rashid

Enhancement of Global Bandwidth of Internet Users Employing Internet Exchange Point Leveraging VPN Technology 7
Al- Nahian Bin Emran, Aman Ullah Juman and Gazi Zahirul Islam

Comparison of Performance of Cognitive Radio Network under Single and Multi-User Scenario 12
Md Abul Kalam Azad, Md. Imdadul Islam, Muhammad R. A. Khandaker and Jugal K. Das

Impact of SUs under FC in Co-operative Cognitive Radio Network in PUEA Environment 17
Fahmida Hossain, Anzuman Yeasmin Snigdha and Imdadul Islam

Sleep-based DBA Algorithm for Energy Efficiency by End Points Collaboration in EPON 22
Kamrun Nahar Runa, Monir Hossen and Shuvashis Saha

Server Placement for Serving VIP Clients in Case of Link Failure 28
Sadia Sharmin, Shafayat Ahmed, Md. Saidur Rahman and Sourav Kumar Mondal

DDoS Mitigation and Intrusion Prevention in Content Delivery Networks using Distributed Virtual Honeypots. 34
Md. Mahmudur Rahman, Shanto Roy and Mohammad Abu Yousuf

Merging the communication protocols 6LoWPAN-CoAP and RPL-CoAP (Simulation and performance analysis using Cooja simulator) 40
Faria Hossain, Arif Mahmud, Faija Juhin and Tasnim Ara

A Low Cost Dual-Band Patch Antenna Array for Future 5G Communication Applications 46
Md Eshtiuk Hossain, Prosen Kirtonia and Mst Fateha Samad

Differentially Fed Compact Triple-band Antenna for Implantable Medical Devices 50
Prosen Kirtonia, Mst. Fateha Samad, Md Kamal Hosain and Md Eshtiuk Hossain

Design & Analysis of a Micro-strip Patch Antenna for RFID, WiMAX and X-band Applications 54
Mst. Dilshad Jahan, Chinmoy Das and Nayan Sarker

Traffic Load Based Efficient Energy Management Technique for 5G Small Cell Networks . 60
Syed Zahidur Rashid, Atikur Rahaman, Abdul Gafur, Yaheya Solaiman and Erfanul Hoque Bahadur

GDPR Compliance: Implementation Use Cases for User Data Privacy in News Media Industry 66
Ahmedur Rahman Shovon, Shanto Roy, Arnab Kumar Shil and Tanjila Atik

Applying Model-Free Reinforcement Learning Algorithm in Network Slicing for 5G 72
Mostafa Zaman Chowdhury, Md. Tanvir Hossan and Yeong Min Jang

Throughput Analysis of Compressive Sensing in Cooperative Opportunistic Radio Network 76
Shohely Tasnim Anindo, Teethiya Datta and Sk. Shariful Alam

Design of a High Voltage Closed Loop Charge Pumping System with a self-regulation mechanism 82
Alif Md. Asif Khan, Tahmida Islam, Md Wasif Absar, Md Rubel Sarkar and Md Shariful Islam

Effects of Wave Function Penetration on Gate Capacitance and Threshold Voltage of p-Channel Double Gate Junction Less Transistor 88
Md. Mohsinur Rahman Adnan and Mohammad Shahidur Rahman

Design of CMOS On-Chip Millimeter-Wave Transformer Coupled Balun and Power Divider-Combiner with Optimal Amplitude and Phase Imbalance 93
Md Sazzad Hossain, Minoru Fujishima, Takeshi Yoshida, Shuhei Amakawa and Md. Mamunur Rashid

Designing an Electro-Mechanical Ventilator Based on Double CAM Integration Mechanism 99
Md. Rakibul Islam, Mohiuddin Ahmad, Md. Shahin Hossain, Muhammad Muinul Islam and Sk. Farid Uddin Ahmed

Performance evaluation of an optical drop filter based on simple homogenous photonics crystal resonator 104
Md. Aminur Rahman and A.K.M Ehtesanul Islam

A High Speed and Low Power 32nm FinFET Dynamic Latch Comparator Design 108
Mir Muntasir Hossain and Satyendra N. Biswas

MOSFET-like Carbon nanotube Field Effect Transistor based Full adder design 112
Bappy Chandra Devnath and Satyendra N. Biswas

A Novel Approach to detect the Compound Level of Fertilizer and Pesticide in the Soil. . 116
Abu Shufian, Riadul Islam and Md. Mominur Rahman

Performance Analysis of EWOD Lens with Minimum Response Time and Optimized Damping Oscillation 122
Sumaiya Rahman and Mst. Fateha Samad

Impact of Increased Quantum Confinement on Gate Capacitance and Threshold Voltage of P-Channel Junctionless Transistor 127
Md. Khairul Alam Titu, Sajid Hasan and Md. Mohsinur Rahman Adnan

Design and Analysis of Dual-Core D-Shape Photonic Crystal Fiber Biosensor 132
Mohammad Al Mahfuz, Md. Nazmus Sakib, Mohammad Hasibul Hasan Hasib, Akash Bosu and Md. Roniuzzaman

An Efficient Soft Error Tolerant Approach to Enhance Reliability of TCAM 136
Rumana Afrin, Muhammad Sheikh Sadi and Jan Jrjens

An investigation on the surface states evolution of polycrystalline silicon nanowires for gas sensing application 142
Mohammad Mojammel Al Hakim, Afrida Arshad, Sharmin Shoma and Urmila Saha

Design and Analysis of high Efficient Latched Comparator 148
Dona Basak and Satyendra Nath Biswas

Performance comparison between Graphene Nano-Ribbon FET & conventional CMOS based on Arithmetic Logic Unit (ALU) 153
Md. Omar Faruque, Tanvir Hasan and Safayat-Al Imam

Numerical Assessment of the Backward Facing Step for NACA 0015 Airfoil using Computational Fluid Dynamics 156
Asif Kabir, Md. Saadbin Chowdhury, Md. Jahirul Islam and Mehran Islam

A Spatiotemporal Analysis and Forecasting of Electricity Generation-Mix in Bangladesh . 162
Sujoy Barua, Anik Nath, Fahim Shahriyar and Nur Mohammad

Compensation of Frequency Fluctuation in a Single Area Power System Using LQR-VC . 167
Fariya Tabassum and Dr Md Sohel Rana

Low Power Consuming Solar Assisted Vapor Compression Refrigerator To Preserve Emergency Medicines And Vaccines In Rural Areas 172
Akbar Hossain, Shovan Saha, Md Arifur Rahman and Mahbub Alam Himel

Robust Control of Terminal Voltage of an Isolated Electric Power Generating Unit 178
Fariya Tabassum and Dr Md. Sohel Rana

Design & Implementation of a Microcontroller Based Automatic Power Factor Rectification System for Different Loads 183
Md. Mayen Uddin, Abdullah Al Mahmud and Naeemul Islam

IMPROVED LOAD FREQUENCY CONTROL OF TWO-AREA POWER SYSTEM ... 189
Abdullah Al Mamun, Mohammad Mostafizur Rahman, Anowar Hosen, Sanjita Akter and Md. Mijanur Rahman

A Comparative Study of Various Methods of Phasor Measurement Unit Algorithms 194
Zahidur Rahim Talukder, Kazi Nishat Tasnim and Md. Shamim Reza

Water Cooled Chiller Based HVAC System Used in a Linear Generator for Oceanic Wave Energy Conversion 200
Selim Molla and Omar Farrok

Generation of Electrical Power by using High Graded Permanent Magnet Linear Generator for Wave Energy Conversion 206
Md. Samiul Bashir and Omar Farrok

Improved Bangla Language Modeling with Convolution 211
Shuvendu Roy

Denoising Sequence-to-Sequence Modeling for Removing Spelling Mistakes 215
Shuvendu Roy

Extracting RDF Triples from Raw Text 220
Yeasmin Ara Akter and Md Ataur Rahman

Bangla Spell Checking and Correction Using Edit Distance 224
Muhammad Ifte Khairul Islam, Rahnuma Islam Meem, Faisal Bin Abul Kasem, Aniruddha Rakshit and Md. Tarek Habib

Filtering Political and Sports News of Social Media from Textual Information 228
Lutfun Nahar and Zinnia Sultana Sultana

Sentiment Analysis & Emotion Extraction: a review of research paradigm 234
Lutfun Nahar

Topic Classification from Text Using Decision Tree, K-NN and Multinomial Nave Bayes . 242
Md Ataur Rahman and Yeasmin Ara Akter

Survey on Text-Based Sentiment Analysis of Bengali Language 246
Nayan Banik, Md. Hasan Hafizur Rahman, Shima Chakraborty, Hanif Seddiqui and Muhammad Anwarul Azim

Doly: Bengali Chatbot for Bengali Education 252
Md. Kowsher, Farhana Sharmin Tithi, M Ashraful Alam, Mohammad Nurul Huda, Mir Md. Moheuddin and Md. Golam Rosul

Threat and Abusive Language Detection on Social Media in Bengali Language 258
Puja Chakraborty and Md. Hanif Seddiqui

Adheetee: A Comprehensive Bangla Virtual Assistant 264
Syed Mohidul Islam, Most Fowziya Akther Houya, Syed Mynul Islam, Salekul Islam and Nahid Hossain

A Bangla Spell Checking Technique to Facilitate Error Correction in Text Entry Environment 270
Tanni Mittra, Sadia Nowrin, Linta Islam and Deepak Chandra Roy

Question Bank Similarity Searching System (QB3S) Using NLP and Information Retrieval Technique 276
Md Raihan Mia and Abu Sayed Md. Latiful Hoque

Textual Question Answering for Semantic Parsing in Natural Language Processing 283
Jaydeb Sarker, Mustain Billah and Md. Al Mamun

Information Retrieval System in Bangla Document Ranking using Latent Semantic Indexing 288
Md. Nesarul Hoque, Rabiul Islam and Md. Sajidul Karim

Face2Anime Generating Anime from Real Human Image with Adversarial Training 293
Shuvendu Roy

Forecast Upcoming Traffic States by Exploiting Nearest Junctions and Big Data 298
Md. Akter Hossain and Mohammed Nazim Uddin

Pest-Inspector: An Insect Detection System by Learning Data Representations 304
Md. Akter Hossain, A. H. M Tareq and Mohammed Nazim Uddin

Kinect 3D Reconstruction Fusion Mapping with C# Algorithm By RGB-D Sensor 310
Shoumik Rahman Mehedy, Ashikur Jhalak and Rashedur M. Rahman

Detection of Lung Cancer from CT Scan Images using GLCM and SVM 316
Mehdi Hassan Jony, Fatema Tuj Johora, Parvin Khatun and Humayan Kabir Rana

Single Image Haze Removal with Non Linear Complex Diffusion Process 322
Sharmin Sultana, Nazifa Akter, Rahat Hossain Faisal and Md. Mostafijur Rahman

Underwater Image Enhancement Based on Fusion Technique via Color Correction and Illumination Adjustment 328
Shahjadi Silma Subah, Md. Ashfaqul Islam and Md. Mohyminul Islam

Automatic Examinee Validation System using Eigenfaces 333
Nuruzzaman Faruqui, Mohammad Abu Yousuf and Md. Fazlul Karim Patwary

Prediction of Potato Disease from Leaves using Deep Convolution Neural Network towards a Digital Agricultural System 340
Md Al-Amin, Tasfia Anika Bushra and Md Nazmul Hoq

Human Identification Based on Color Stimuli 345
Md.Mahmudul Hasan, Nafiul Hasan, Md. Mohaiminul Islam and Mahojabin Suhi

Design of a Convolutional Neural Network Based Smart Waste Disposal System 349
Md. Samiul Haque Sunny, Debopriya Roy Dipta, Shifat Hossain, Hossain Mansur Resalat Faruque and Eklas Hossain

Brain Tumor Classification Using Convolutional Neural Network 354
Sunanda Das, O.F.M. Riaz Rahman Aranya and Nishat Nayla Labiba

Improved Segmentation Approach for Plant Disease Detection 359
Md. Arifur Rahman, Md. Mukitul Islam, Gazi Mohammad Shahir Mahdee and Md. Wasi Ul Kabir

An Automated Approach to Detect Breast Cancer Tissue Using Ultrasound Images 364
Dr. T M Shahriar Sazzad, Misbah Ul Hoque, A.K.M. Ashek Farabi, Ismail Hosen and Mursheda Akter Somi

An Automated Dermatological Images Segmentation Based on a New Hybrid Intelligent ACO-GA Algorithm and Diseases Identification Using TSVM Classifier 368
Md. Humayan Ahmed, Romana Rahman Ema and Tajul Islam

Design and Development of a Hand Gesture Based Virtual Mouse 374
Kabid Hassan Shibly, Samrat Kumar Dey, Md. Aminul Islam and Shahriar Iftekhar Showrav

Non-parametric development of the ranked tree structure for the accurate facial landmarking localization. 379
Supratip Ghose and Habibur Rahman

A Vision Based Three-layer Access Management System with IoT Integration 384
Nafize Ishtiaque Hossain, Ali Reza Galib, Raihan Bin Mofidul and Dhiman Kumar Sarker

Computer Vision and Smoke Sensor Based Fire Detection System 389
Nahid Chowdhury, Dewan Ruhul Mushfiq and Ehtesham Chowdhury

Implementation of Principal Component Analysis on Mask and Non-mask Face Recognition 394
Md. Sabbir Ejaz, Md. Rabiul Islam, Md. Sifatullah and Ananya Sarker

Brain Tumor Detection Using Convolutional Neural Network 399
Tonmoy Hossain, Fairuz Shadmani Shishir, Mohsena Ashraf, Md Abdullah Al Nasim and Faisal Muhammad Shah

BFAGC: A Bias-Free Adaptive Gamma Correction Method for Image Enhancement 405
Nazifa Akter, Sharmin Sultana, Rahat Hossain Faisal and Md. Mostafijur Rahman

Automatic Attendance System: A Face Recognition Approach 411
Manukumar Rudresh, C. Lalnunchama, Arun Kumar, Vijay Bhaskar Semwal and Vinay Anand Tikkiwal

A Dynamic Face Recognition using HAAR & CNN: A Transfer Leaming Approach 415
Amaan lqbal, Kunal Kumar, Arun Kumar, Vijay Bhaskar Semwal, and Vinay Anand Tikkiwal

Heart Failure Risk Prediction and Medicine Recommendation using Exploratory Data Analysis 419
Sumaya Habib, Maisha Binte Moin, Sujana Aziz, Kalyan Banik and Hossain Arif

Chapman-Kolmogorov Relation Based Median String Algorithm for DNA Consensus Classification 425
Mohammad Shibli Kaysar and Mohammad Ibrahim Khan

Performance Analysis and Benchmarking of Clustering Algorithms with gene datasets ... 431
Meskat Jahan

Linear Regression Models for Total Body Water Measurement Based on Bioelectrical Impedance Analysis 436
Mst. Farzana Khatun, Dr Md Sohel Rana and Dr Ajay Krishno Sarkar

Improved Piecewise Constant Approximation Method for Compressing Data Streams 442
Atik Mahbub, Farhana Haque, Habibul Bashar and Mohammad Rezwanul Huq

Mathematical Models for Fat Free Mass Measurement Based on Bioelectrical Impedance Analysis 448
Tahasin Ahmed Fahim and Dr Md Shamim Anower

Statistical Approach for Classifying Sentiment Reviews by Reducing Dimension using Truncated Singular Value Decomposition 453
Asmaul Husna, Habibur Rahman and Emrana Kabir Hashi

Building Predictive Models for Infant Mortality Rate in Bangladesh 458
Mst Lucky Sultana, Mst. Irin Sultana and Md Mahmudul Hasan

Design of Twenty-One Level Modular Multilevel Converter Using a Three-level Submodule Block 464
Ashraf Siddiquee, Shirajum Munira and Md. Shamim Reza

Comparative Study of Enhancing Stability of a Wind Farm connected to the Grid using PID Controller based STATCOM and Capacitor Bank 469
Ajmayen Fayek, Shah Md. Salimullah, Md. Shakhawath Hossain, Rakib Hossain, Md. Shadman Hasan Shakib, Ariful Islam Anik and Moynul Hossen Khan

Load Scheduling of a Refrigerated Warehouse with Homogeneous Compressors Under Dynamic Pricing 473
Shah Md. Nehal Hasnaeen, Shohanur Rahman and Md. Forkan Uddin

Compensation for Voltage Disruption in a Smart Grid during Cyber-Attack 479
Md. Abdur Rahman and Dr Md. Sohel Rana

Fuzzy Logic based Optimization for Electric Vehicle Charging Station in Bangladesh 484
Ashish Kumar Karmaker and Md Raju Ahmed

Modeling and Design of a Fuzzy Logic Based PID Controller for DC Motor Speed Control in Different Loading Condition for Enhanced Performance 489
Mirza Muntasir Nishat, Fahim Faisal, Mahdi Rahman and Md Ashraful Hoque

A Systematic Approach to Design Optimal Controller of Single Phase VSC for Distributed and Renewable Energy Application 495
Md Khurshedul Islam, Farhina Haque and Naeemul Islam

Implementation of Finite Control Set Model Predictive Control on 7-level Flying Capacitor Multilevel Inverter Using Space Vector 501
Nasir Jubair Islam, Md.Selim Sarker, N.M.Merazul Islam and Md.Shamim Reza

Design and Comparative Analysis of Dual-stage Servo controller in Hard Disk Drive 507
Alamgir Hossain and Md Arifur Rahman

Intelligent Speed Control of DC Motor using ANFIS D Controller 512
Tanjim Tarannum

A New Topology of Input Switched Single Phase AC-DC Buck Boost Converter for High Performance 517
Khadiza Akter, Sairatun Nesa Soheli and Md. Nafiul Islam

A Novel 36-Pulse Star Rectifier with Passive Injection Circuits at DC Link 521
Md. Abdul Malek, Md. Abu Bakar Siddik and Md. Shoaib Akther

Experimental Results of a locally developed BLDC Motor Controller for electric tricycle . 527
Fahim Bhuiyan, Nazmus Sakib, Mohammad Rejwan Uddin and Khosru M Salim

HOMER Optimized Off-grid Hybrid Energy System: A Case Study on Rohingya Relocation Center in Bangladesh 531
Ferdous S. Azad, Istiak Ahmed, Syed Raihan Hossain and Rashedul Amin Tuhin

Improved Differential Strobe Terminal Based Open-loop Control of Smart Power Factor Metering System 536
Shahriar Rahman Fahim, Sakib Shahrear Avro, Subrata K. Sarker, Md. Rokunuzzaman and Sajal K. Das

Converter based Frequency Adjustment and Protection of Grid-Tied Wind Farm 542
Kanij Ahmad, Nur Mohammad and Muhammad Quamruzzaman

Design of a cost effective Smart Autonomous Power and Switch Controlled Led Street Light 547
Monir Hossain, Md. Tawhid Islam Opu, Mohaiminul Islam Ratul and Navila Rahman Nadi

Analysis of Air Foils and Design of Blades for a Low-Speed 250W Vertical Axis Wind Turbine Suitable for Coastal Areas of Bangladesh 553
Mahtab Murshed, Md. Yeasin Arafat and M. Abdur Razzak

Genetic Expression Analysis To Detect Type Of Leukemia Using Machine Learning 559
Umid Dey and Md Sajjatul Islam

Sensor-Less Solar Irradiance Estimation and Maximum Power Point Tracking Using Artificial Neural Network 565
Arif Rubayet and M. M. Atiqur Rahman

Electricity Theft Detection to Reduce Non-Technical Loss using Support Vector Machine in Smart Grid 570
Rafia Nishat Toma, Md Nazmul Hasan, Abdullah Al-Nahid and Bo Li

A Machine Learning Approach to Protect Electronic Devices from Damage Using the Concept of Outlier 576
Sunanda Das, Muhammad Sheikh Sadi, Md. Ahsanul Haque and Md. Milon Islam

Ulcer Detection in Image Converted from Video Footage of Wireless Capsule Endoscopy . 582
Abdullah Al Mamun and Md. Sohag Hossain

Feature-Based Mobile Phone Rating Using Sentiment Analysis and Machine Learning Approaches 586
Abdullahil Kafi, M. Shaikh Ashikul Alam, Sayeed Bin Hossain, Siam Bin Awal and Hossain Arif

Unsupervised Content-Based Image Retrieval Technique Using Global and Local Features 592
Md. Farhan Sadique, Bishajit Kumar Biswas and S.M. Rafizul Haque

Bangla Handwritten Character Recognition: an overview of the state of the art classification algorithm with new dataset 598
Md Nazmul Hoq, Nadira Anjum Nipa, Mohammad Mohaiminul Islam and Sadat Shahriar

Personality Detection from Texts using Convolutional Neural Network 604
Md. Abdur Rahman, Asif Al Faisal, Tayeba Khanam, Mahfida Amjad and Md Saeed Siddik

An Empirical Study of Cervical Cancer Diagnosis using Ensemble Methods 610
Enamul Karim and Nafis Neehal

Bangla Document Categorisation using Multilayer Dense Neural Network with TF-IDF. . 615
Manisha Chakraborty and Mohammad Nurul Huda

Neural Network based Software Defect Prediction using Genetic Algorithm and Particle Swarm Optimization 619
Safial Islam Ayon

Artificial Neural Network Inducement for Enhancement of Cloud Computing Security 623
Md. Raihan Uddin, Md. Taslim Arefin and Kh. Mohaimenul Kabir

Non Functional Requirements Classification with Feature Extraction and Machine Learning: An Empirical Study 629
Md. Ariful Haque, Md. Abdur Rahman and Md Saeed Siddik

Design and Development of Question Answering System in Bangla Language from Multiple Documents 634
Samina Tasnia Islam and Dr. Mohammad Nurul Huda

Using Boosting Approaches to Detect Spam Reviews 638
Sifat Ahmed and Faisal Muhammad Shah

Performance Analysis of Different Decision Tree Based Methods for Identifying Drop Out Students 644
Nafisa Tasnim, Mahit Kumar Paul and A.H.M. Sarowar Sattar

Students Academic Performance Evaluation Method Using Fuzzy Logic System 650
Abdul Aziz, Md. Asaf-Uddowla Golap and M. M. A. Hashem

Using Social Networks to Detect Malicious Bangla Text Content 656
Tanvirul Islam, Nadim Ahmed and Subhenur Latif

Classification of Bengali Questions Towards a Factoid Question Answering System 660
Syeda Tamanna Alam Monisha, Sourav Sarker and Md Mahadi Hasan Nahid

Evidential SVM for binary classification 665
Md. Eusha Kadir, Pritom Saha Akash, Amin Ahsan Ali, Mohammad Shoyaib and Zerina Begum

Sentiment Analysis of Bengali Texts on Online Restaurant Reviews Using Multinomial Naive Bayes 671
Omar Sharif, Mohammed Moshiul Hoque and Eftekhar Hossain

Clinical Human Gait Classification: Extreme Learning Machine Approach 677
Prithvi Patil, K.Shusheel Kumar, Neha Gaud and Vijay Bhaskar Semwal

A Performance Improved Kretschmann Configuration Based Surface Plasmon Resonance (SPR) Sensor 683
Tauhidul Haque and Hasan Khaled Rouf

Leakage Detection in Water Pipeline Using Microcontroller 687
Muhammad Arman Uddin, Mohammed Mohibul Hossain, Samin Yeasar Seaum, Akil Ahmed and Hasibul Hasan Sabuj

Smart Gas Furnace 691
Md. Mominur Rahman, Abu Shufian, Farzana Akter, Md. Tarikul Islam Juel and Riadul Islam

DOORMOR: A Fully Functional Prototype of a Search and Rescue Robot 696
Fahim Faisal and Farhana Sharmin

VADet: An Arduino based automated vehicle accident detection and messaging system .. 702
Nazir Ahmmed, Nusrat Jahan Jenny, Most Fowziya Akther Houya, Anika Ibnat Binte Alam and Md. Adnan Arefeen

ParkEasy: An Embedded Model to Mitigate Car Parking Problems Using IoT Technology 707
Samrat Kumar Dey, R. M. Rubayet Shamim, Md. Atiqul Islam and Md. Mahbubur Rahman

DESIGN AND IMPLEMENTATION OF BOILER AUTOMATION SYSTEM USING PLC 714
Md.Iquebal Hossain Patwary, Shafayat Hossain, Rubel Ahmed, S.M. Tahsin Labib, Sohan Ul Alam and Md.Ashikul Islam

Face Recognition System Based on Raspberry Pi Platform 720
Md Asaduz Zaman Mamun, Nafis Mustakim, Noushad Hossain, Mohammad Mustafizur Rahman, Nadimul Islam and Zayed Hossain Sayem

Automatically Controlled Energy Conservation System for Corporate Office Based on Microcontroller. 724
Koushik Ahmed, Joydip Paul, Md. Mominur Rahman, Abu Shufian, Md Shahnewaz Tanvir and Md. Mohaimanul Islam Sagor

Affordable Smart ECG Monitoring Using Arduino & Bluetooth Module 730
Mohammed Mahmudur Rahman, Muhammad Armanul Hoque, Md Azizul Hoque Rimon and Md. Redwan Sammir

Six tiers multipurpose Security Locker System Based on Arduino. 734
A. Z. M. Tahmidul Kabir, Nirmol Debnath, Utshaw Rafin Akther, Tawsif Ibne Alam and Fukrul Hasan

Design and Implementation of a Vault Security System 739
Md. Mahadi Hasan Moon, Debashish Kumar Ghosh and Nafiz Ahmed Chisty

Design and Implementation of Remote Controlling and Monitoring System for Automatic PLC Based Packaging Industry 745
Raihan Bin Mofidul, Md. Shahadath Hossain Sabbir, Amit Kumer Podder and Mohammad Shaifur Rahman

Development of an Automated Low-cost Book Scanner and Translator 750
Sabila Nawshin, Samir Kumar Das, Najia Hossain, Joyita Das Mela, Abu Shafin Mohammad Mahdee Jameel and Salekul Islam

A Strategic Routing Analysis for Agro Sensor Communications in Mobile Ad hoc Networks 755
Sampad Banik, Md Munjure Mowla and Iftekhar Ahmad

Return Loss and Gain Improvement for 5G Wireless Communication Based on Single Band Microstrip Square Patch Antenna 761
Mohammad Faisal

Performance Evaluation of SDN Controllers in Wireless Network 766
Saiful Islam, Md. Ariful Islam Khan, Sanjida Tasnim Shorno, Sumonto Sarker and Md. Abubakar Siddik

Low-Cost Underwater Wireless Communication System Development Using Piezo-Ceramic Transducer 771
Md Ragib Shaharear, H.A.Shakhawat Hossen Prayash, Md Farhanul Islam and Tuhin Dey

Mobility Aware Cache Management in 5G Future Generation Wireless Communication System 775
Mayeesha Fairooz Ahmad and Mohammad Arif Hossain

An Intelligent Assistive Tool for Alzheimer's Patient 781
Kazi Shahrukh Omar, Afia Anjum, Tahrima Oannahary, Md. Rezaul Karim Rizvi, Diana Shahrin, Tasmiah Tamzid Anannya, Sanjida Nasreen Tumpa, Md Mahboob Karim, Md Fazle Rabbi and Muhammad Nazrul Islam

A Prototype of Automated Vault Locker Solution for Industrial Application 787
Shafayet Hossain, Mazid Ishtique Ahmed and Md. Niaz Mostakim

The Third Tab (3T): Assistant for Visually Impaired People 793
Sohal Islam, Kazi Arham Kabir and Md. Jakaria Rahimi

A Case Study on a Proposed Adaptive and Energy Efficient Street Lighting System for Chittagong City 798
Saif Hannan, Goutom Barua Milton, Md. Humayun Kabir and Mohammed Jashim Uddin

Detection of Safe Angle for Water Vehicles due to Vulnerable Swing 803
Md. Shahidul Islam and Sarfaraz Newaz

A Comparative Study of Walking Assistance Tools Developed for the Visually Impaired People 809
Mahfuza Khanom, Muhammad Sheikh Sadi and Md. Milon Islam

Performance Evaluation of a Wearable 2.45 GHz Planar Printed Meandering Monopole Textile Antenna on Flexible Substrates 814
Md. Mamun Ur Rashid, Ashiqur Rahman, Liton Chandra Paul and Ajay Krishno Sarkar

Intelligent Path-finder for the blind. 820
A. Z. M. Tahmidul Kabir, Nirmol Debnath, Mohitosh Pramanik and Tawsif Ibne Alam

A multi-controlled semi-autonomous wheelchair for old and physically challenged people . 825
Md.Eftekhar Alam, Mohammed Abdul Kader, Umme Hany, Rumana Arjuman, Aisha Siddika and Md.Zakaria Islam

Towards Developing an Intelligent Wheelchair for People with Congenital Disabilities and Mobility Impairment 831
Tasneea Hossain, Md Sabbir-Ul-Alam Sabbir, Asma Mariam, Toki Tahmid Inan, Muhammad Nazrul Islam, Khairul Mahbub and Muhammad Tawsif Sazid

Wearable Visual-MIMO for Healthcare Applications 837
Tae-Ho Kwon, Ki-Soo An, Partha Pratim Banik, and Ki-Doo Kim

RoboFI: Autonomous Path Follower Robot for Human Body Detection and Geolocalization for Search-Rescue Missions using Computer Vision and IoT 841
Ahmed Imteaj, M A Isfar Jubair Chowdhury, Mohammad Farshid and Abdur R Shahid

Development of Microcontroller based Light Following Emergency Evacuation Chair 847
Delowar Hossain, Hasan Alam, Ajmain Ar Rafi, Kamruzzaman Rakib and Umma Hany

A Prototype of Heavyweight Carrying Robot for Garments Industry in Bangladesh 852
Md. Mizanur Rahman, Sanjoy Kumar Sana, Md. Rokib Hasan, Sadia Kamal, Junaed Jahin and Thajid Ibna Rouf Uday

Q-Learning Based Particle Swarm Optimization Algorithm for Optimal Path Planning of Swarm of Mobile Robots 858
Syed Irfan Ali Meerza, Moinul Islam and Md. Mohiuddin Uzzal

Offloading SLAM for Indoor Mobile Robots with Edge-Fog-Cloud Computing 863
Victor Kathan Sarker, Jorge Pea Queralta, Tuan Nguyen Gia, Tomi Westerlund and Hannu Tenhunen

Simulation of Ferrofluid Based Soft Actuator Using Magnetic Field Optimization Approach 869
Tasnova Tanzil Khan, A. M. Harsha S. Abeykoon, Tanujjal Bora and Golam Fahad Mullick

Design of a Prototype of a Ferrofluid Based Soft Actuator 875
Tasnova Tanzil Khan, Tanujjal Bora, A. M. Harsha S. Abeykoon and Golam Fahad Mullick

Stable walking of an underactuated bipedal robot with unbalanced masses using one actuator 880
Mohammad Farhan Ferdous and Jannat Binta Alam

Artificial Intelligence Humanoid BONGO Robot in Bangladesh 886
A S M Ahsanul Sarkar Akib, Mohammad Farhan Ferdous, Milon Biswas and Hossain Mohammad Khondokar

Design & Implementation of Voice Command Based Bipedal Surveillance Robot 892
Himel Das, Himel Chakraborty and Md. Sarwar Uddin Chowdhury

Development of Myoelectric Robotic Hand using 3D Printer 897
Sivapong Nilwong, Delowar Hossain, Eneo Petoku and Genci Capi

Revised Scheme for Antimagic Graph Labeling for Splitting Graph and Shadow Graph Associated with Cycle 901
Zareen Tasneem, Farissa Tafannum and Md Mahbubur Rahman

Icons For the Mass: An Approach Towards Text Free Smart Interface 906
Eshtiak Ahmed, Md. Mahade Hasan, Mirza Omar Faruk, Muhammad Farhad Hossain, Md. Arifur Rahman and Ashraful Islam

Efficient Compression Scheme for Large Natural Text Using Zipf Distribution 910
Md. Ashiq Mahmood and K.M. Azharul Hasan

Design and Optimization of AlGaAs/InP Multi-junction Solar Cell 916
Md. Masud Rana, S. Md. Firoz Khan and Md. Shafin Shafayat

Cows Urine is a prominent source of electrolyte for Microbial Fuel Cell 921
Md. Abdul Halim, Rituparna Kundu, Md Ibrahim, Md. Rasel Mollah, Hasan Md. Bokhari and Islam Md. Aminul

An Intelligent Street Light System Based on Piezoelectricity Generator with Noticeable Zebra Crossing Point and Bus Stand. 925
A. Z. M. Tahmidul Kabir, Nirmol Debnath, Mirza Tanvir Sadik, Sakhwat Hossain Lisun and Fukrul Hasan

Electrical and Morphological Characterization of Zinc-Doped a- Fe2O3 Thin Films at Different Annealing Temperature 931
Arman Ibn Shahjahann

Fabrication of Digital Sol-Gel DIP Coating System with Smart Beaker Carrier 935
Sarafat Hussain Abhi, Md. Shoaib Akther and Dr. Md. Faruk Hossain

Enhancement of Sensitivity for Surface Plasmon Resonance Biosensor with Higher Detection Accuracy and Quality Factor 939
Mohammad Hasibul Hasan Hasib, Jannati Nabiha Nur, Kamrun Nahar Shushama, Imteaz Rahaman, Md. Masud Rana and Mohammad Al Mahfuz

Design and Implementation of Solar Power Wireless Battery Charger 944
Abu Shufian, Md. Mominur Rahman, Koushik Ahmed, Riadul Islam, Mahmodul Hasan and Toufiqul Islam

Offshore Wind Energy Estimation in the Bay of Bengal using Satellite Wind Measurement 949
Navila Rahman Nadi, Ferhat Bingl and Merete Badger

Electrical Energy Harvesting from the Foot Stress On Foot Overbridge Using Piezoelectric Tile 957
Tarik Bin Abdul Akib, Hasanuzzaman Mehedi and Md. Nazmuschayadat

Aspects of Numeral Value of Cu2ZnSnS4 (CZTS) Absorber Layer Fabricated in Sol-gel Process by Dip Coating Technique 962
Sabiha Dilruba, Tabassum Perveen Ananna, Afrina Sharmin and Zahid Hasan Mahmood

A Statistical Estimation of Solar Power for Energy Mix in Bangladesh 966
Sujoy Barua, Zainal Abedin, Anik Nath and Chitra Biswas

A Comparative study on Waste-to-energy (WTE) technologies for municipal solid waste (MSW) management in Chittagong, Bangladesh. 970
Md. Rashidul Islam, Jobair Al Rafi and Md. Sajjad Hossain

Modeling and Performance Analysis of an Γ-Model Based DFIG Wind Turbine with Duel Crowbar Protection 976
Mehedi Azad Shawon, S. M. Imrat Rahman and Naheed Hossain

Electrical Properties of CSS Deposited CdTe Thin Films for Solar Cell Applications 982
Nipu Kuamr Das, Sheikh Alimur Razi, Kazi Sajedur. Rahman, Mahmud Abdul Matin Bhuiyan and Nowshad Amin

Layer thickness effect on power conversion efficiency of a P3HT:PCBM based organic solar cell 986
Raian Islam, Murad Abrar, Farhatul Hassan and Sifat Adnan

An Efficient CZTS Solar Cell from Numerical Analysis 989
E.M.K. Ikball Ahamed, Ashoke Kumar Sen Gupta, Muhammad Qamruzzaman and M. A. Matin

Modelling a Solar Charge Station for Electric Vehicle with Storage Backup 994
Muhammad Ali, Safayat Mohammad and Md. Minhajur Rahman

Comparison of Different Hybrid Renewable Energy Systems With Optimized PV Configuration to Realize the Effects of Multiple Schemes 998
Faishal Islam Bappy, Md Jahirul Islam, Amit Kumer Podder, Debopriya Roy Dipta, Hossain Mansur Resalat Faruque and Eklas Hossain

Thickness Dependency of Zinc Selenide (ZnSe) Thin Film Deposited By Vacuum Evaporation Method 1004
Md. Abu Sayeed, Hasan Khaled Rouf, Mohd. Rezaul Hasan and K.M.A. Hussain

Feasibility Analysis and Economic Impact Assessment of Microgrid Renewable Energy Systems in Burao, Somaliland 1009
Abdiaziz Mouse Abdi Miad, S. M. G. Mostafa, Mohammad Azim Uddin and Yasir Arafat

Lightning Effect on a Large-Scale Solar Power Plant with Protection System 1015
Mohammad Riajul Karim and Md. Raju Ahmed

Analysis of NCFET with Si Doped HfO2 as Ferro-Electric Material to Achieve Lower SS . 1020
Towfiqur Rahman, Antara Fairooz Khan and Nowshin Nawal

Large Magnetic Remanence Property Based a New Material Used in the Linear Generator For Wave Energy Extraction 1026
Selim Molla and Omar Farrok

Performance Enhancement of Multigate FinFETs by Using High-k Stack Oxide 1031
Emdadul Huq Minhaj, Muhammad Abdur Razzak, Md. Mohsinur Rahman Adnan and Md. Majharul Islam

Effect of Solar Concentration on Performance of Perovskite Solar Cell: A detailed balance study 1035
Md. Masud Rana, Shyamal Kumar Mandal and Md. Tariquzzaman

Details semiconductor losses calculation of single phase dual active bridge converter for fast charger station application 1040
Mohammad Jalal Uddin

Land-Use Detection Using Residual Convolutional Neural Network 1045
Shuvendu Roy and Sneha Paul

Pathological Voice Classification Using Deep Learning 1051
Shuvendu Roy, M. A. H. Akhand and Md Ijaj Sayim

An Efficient Method of Lane Detection and Tracking for Highway Safety 1057
Bubly Barua, Shuva Biswas and Kaushik Deb

Bangla Sign Language Conversation Interpreter Using Image Processing. 1063
Prosenjit Roy, S.M.Miraj Uddin, Md. Arifur Rahman, Md. Musfiqur Rahman, Mohammad Shahin Alam and Md.Saidur Rashid Mahin

Depression Analysis of Bangla Social Media Data using Gated Recurrent Neural Network 1068
Abdul Hasib Uddin, Durjoy Bapery and Abu Shamim Mohammad Arif

Designing an automatic workplace recommendation system for garment workers using fog computing 1074
Arunavo Dey and Md Nahid Newaz

Retinal Fundus Identification Utilizing Supervised and Unsupervised Nature of Deep Neural Network 1078
Md. Sanaullah Chowdhury, Faozia Rashid Taimy, Abdullah-Al Nahid, Md. Younus Ali and Ferdous Bin Ali Hridoy

Virtual Image from EEG to Recognize Appropriate Emotion using Convolutional Neural Network 1084
Md. Rabiul Islam and Mohiuddin Ahmad

EduNet: A Deep Neural Network Approach for Predicting CGPA of Undergraduate Students 1088
Md Rifatul Islam Rifat, Abdullah Al Imran and A. S. M. Badrudduza

Authorship Attribution in Bengali Literature using Convolutional Neural Networks with fastTexts word embedding model 1094
Hemayet Ahmed Chowdhury, Md. Azizul Haque Imon, Md. Saiful Islam and Syed Md. Hasnayeen

BHCDR: Real-Time Bangla Handwritten Characters and Digits Recognition using Adopted Convolutional Neural Network 1099
Muhammad Aminur Rahaman, Md. Mahin, Md. Haider Ali and Md. Hasanuzzaman

Rice Disease Identification and Classification by Combining Deep Convolutional Neural Network with Support Vector Machine 1105
Md. Jahid Hasan, Shamim Mahbub, Md. Shahin Alom and Md. Abu Nasim

Human Activity Recognition Using Smartphone Sensors: A Dense Neural Network Approach 1111
Dr. Abdul Kadar Muhammad Masum, Sadia Jannat, Erfanul Hoque Bahadur, Dr. Md. Golam Rabiul Alam, Shahidul Islam Khan and Mohammad Robiul Alam

GRU-based Attention Mechanism for Human Activity Recognition 1117
Md. Nazmul Haque, M. Tanjid Hasan Tonmoy, Saif Mahmud, Amin Ahsan Ali, Muhammad Asif Hossain Khan and Mohammad Shoyaib

Discretion Way for Bleeding Detection in Wireless Capsule Endoscopy Images 1123
Abdullah Al Mamun, Md. Sohag Hossain, Md. Galib Hasan and Md. Motaher Hossain

Easy Scheme for Ulcer Detection in Wireless Capsule Endoscopy Images 1129
Md. Sohag Hossain, Abdullah Al Mamun, Md. Galib Hasan and Md. Motaher Hossain

Investigation of EEG Signals for Brain Computer Interface 1134
Etu Podder, Md. Maniruzzaman and Alok Sarkar

Temporal Variables Disorder of The Gait Cycle in Parkinsons Disease 1140
Adharaa Neelim Dewanjee, Quazi Delwar Hossain and Mohammed Shahriar Arefin

Design and Development of a Portable Recording System for Simultaneous Acquisition of SCG and ECG Signals 1146
Tauhid Anwar, Mohammad Zubair, Mohammad Dehan Rahman, Nazmus Sakib, Md. Ahsan-Ul Kabir, Tasnuva Faruk and Md Kafiul Islam

A lossless speech signal compression technique 1152
Md. Atiqur Rahman, Most. Jannatul Ferdous, Md. Mamun Hossain, Md Rashedul Islam and Mohamed Hamada

Effects of Different Environmental noises and Sampling Frequencies on the performance of MFCC and PLP based Bangla Isolated Word Recognition System 1158
Oli Lowna Baroi, Azhar Niaz, Md. Jahidul Islam, Md. Shaikh Abrar Kabir, Emrul Islam and Md. Jakaria Rahimi

Design and Characterization of low loss Single Mode Fiber for Terahertz Signal Transmission 1164
Akash Bosu, Chaity Basak, Shumaia Sharmin, Khandakar Md. Tasrif Hossain, Sharmina Zaman Urmi and Mohammad Al Mahfuz

Classification of Intrinsic Mode Spectra of EEG Signal for Epilepsy Detection 1170
Ankush Sarker and Md.Masudur Rahman

Detection of Abnormal Electrocardiogram (ECG) Using Wavelet Decomposition and Support Vector Machine (SVM) 1175
Pratik Deb, Mohammad Nooruddin and Md. Shajahan Badsha

Epileptic Seizure Detection and Classification using Support Vector Machine from Scalp EEG Signal 1183
Sania Zahan and Md. Rabiul Islam

Identification of Autoregressive Systems in the Presence of Additive Noise Using the Matrix Pencil Method 1188
Md. Shamim Hussain

Detection Performance Analysis of Compressive Sensing in Cooperative Cognitive Radio Network 1194
Teethiya Datta, Anindyo Dip and Sk. Shariful Alam

Design of EEG Based Wheel Chair by Using Color Stimuli and Rhythm Analysis 1200
Nafiul Hasan, Md. Mahmudul Hasan and Md. Akramul Alim

Contrast Enhancement using Mode-Separate Histogram Equalization 1204
Sadia Khan Rupa, Md. Faishal Yousuf and Shanto Rahman

Cloud Based Automated Power Consumption Optimization, Power management and Appliance Control 1210
Juglul Hasan, Tanjim Ul Haque and Sabab Hasan

An IoT Based Green Waste Management System For Bangladesh 1215
Mazhar Ibna Zahur, Md. Saidur Rahman, Aysha Akther and Kazi Masudul Alam

Real-time Water Vehicle Monitoring and Tracking System for Bangladesh 1221
Dr. Mohammad Salah Uddin and Md.Naim Khan

Development of an Android Based Real Time Bus Tracking System 1226
Mohammad Nazmul Hasan and Md. Sharif Hossen

An IoT based mHealth Platform for Chronic Kidney Disease Patients 1231
Alaa Noor, Anonnya Banerjee, Mayeesha Fairooz Ahmad and Mohammed Nazim Uddin

Design and Implementation of IoT based Ideal Fish Farm in the Context of Bangladesh Aquaculture System 1237
Dr. Abdul Kadar Muhammad Masum, Md. Shahin Mia, Md. Kalim Amzad Chy, Shahidul Islam Khan, Ahmed Shan-A-Alahi and Dr. Md. Golam Rabiul Alam

Digital Forensic Investigation Framework for Internet of Things (IoT): A Comprehensive Approach 1243
Md. Jahidul Islam, Md. Mahin, Ayesha Khatun, Biplab Chandra Debnath and Sumaiya Kabir

Path and Partner Selection in Unstructured Resource- Constrained Environments Based on Node Psychology in the Context of IoT 1249
Tanusree Debi and Assaduzaman Asad

Bridge Strength Preservation by Automatic Traffic Density Control: an IoT Application . 1255
Aisha B. Rahman, Mohammed Shahriar Kamal and Ariful Islam

IoT Based Automated Solar Irrigation System Using MQTT Protocol in Charandeep Chakaria 1260
Mohammad Minhajul Islam, Rezawl Karim Reza, Md. Shahadat Hossain and Anik Nath

Towards IoT and ML Driven Cardiac Status Prediction System 1266
Md. Ittahad Uz Zaman, Shaira Tabassum, Md. Sabbir Ullah, Ashikur Rahaman, Samsun Nahar and A.K.M. Muzahidul Islam

IoT Based Smart Energy Management in Residential Applications 1272
Md. Rakibul Hasan, Eklas Hossain, Hossain Mansur Resalat Faruque and Tipu Sultan

IoT-based Portable ECG Monitoring System for Smart Healthcare 1278
Tamanna Shaown, Imam Hasan, Md.Muradur Rahman Mim and Dr. Md. Shohrab Hossain

A Complete Device to Call Candidates and Time Management in Viva-voice 1283
Mohammed Abdul Kader, Rajvi Sutra Dhar, Md. Saiful Islam, Naeemul Islam, Mohammad Abdur Razak and Yasir Arafat

Smart Water System at House Using Arduino Uno and C# Desktop Application to Reduce Water Wastage and Energy Loss 1288
Arif - Ul - Islam and Shamim H Ripon

Towards Developing an Intelligent Automated Water Pumping System for Dhaka City ... 1293
Kazi Md. Munim, Iyolita Islam, Moumita Sarker and Muhammad Nazrul Islam

Bluetooth Network Based Remote Controlled Home Automation System 1298
Md. Ismail Husain, Mahfuz Alam, Md. Golam Rashed, Md. Emdadul Haque, Mirza A. F. M. Rashidul Hasan and Dipankar Das

Automated Railway Maintenance System in Context of Bangladesh 1304
Nusrat Kabir Samia, Rifah Tamanna Usha, Anindita Kundu, Sadia Nowshin, Sharmin Sultana Sharmee, Sanjida Nasreen Tumpa, Major Md Fazle Rabbi and Col Md Mahboob Karim

Wireless Home Automation System Using IoT and PaaS 1309
Adiba Mahjabin Nitu, Md. Jahid Hasan and Md. Shahin Alom

Design and Implementation of an IoT Based Home Automation 1315
Md. Mohaiminul Islam, Md. Nahiyan Farook, Sk.Md. Golam Mostafa and Yasir Arafat

ndroid Based Home Automation Using AVR Microcontroller Through Wireless Communication 1320
Muzahidul Islam, Md Tarek Hossein, Md. Ilias Rahman, Md. Mohtasim Billah, Mohammad Al Mahfuz and Mohammad Hasibul Hasan Hasib

An Effective Combination of Microcontroller and PLC for Home Automation System 1324
Mohammad Hanif, Nur Mohammad and Bijoy Harun

Subject Oriented Data Partitioning A Proposed Data Warehousing Schema 1330
Abdul Moktadir and N. M. Istiak Chowdhury

Feature Selection and Classification of SpinalAbnormalities to Detect Low Back PainDisorder using Machine Learning Approaches 1336
Md. Shariful Islam, Md. Asaduzzaman and Mohammad Masudur Rahman

An Efficient Ensemble Method for Cancer Detection 1340
Md. Sirajum Munir Prince, Asib Hasan and Faisal Muhammad Shah

Big Data Mining in the Presence of Concept Drifting 1346
Tabassum Siraj, Efrana Jannat, Warafta Rasul, Meher Afroz Chowdhury, Swakkhar Shatabda and Dewan Farid

Efficient Implementation of Dynamic Array System Using MapReduce Framework 1352
K. M. Azharul Hasan, Mehnuma Tabassum Omar, Rahat Haider and Sk. Md. Masudul Ahsan

Develop a Model to Secure and Optimize Distributed File Systems for ISP Log Management 1358
Mohammad Shazzad Hossain, Pratik Kumar Bhowmik, Md. Abidur Rahman, Mahbub Uddin, Mohammad Abu Yousuf and Kazi Abu Taher

Feature Selection and Comparative Analysis of the Supervised Learning Model for Hyperspectral Image Classification 1364
Abu Sayeed, Md. Ali Hossain and Md. Rabiul Islam

Analysis of Performance-Improvement of Microstrip Antenna at 2.45 GHz Through Inset Feed Method 1369
A K M Baki Baki, Md. Nurur Rahman and Shawon Kumar Mondal

Ultra High Efficient 2x1 Graphene Patch Antenna Arrays for Single and Dual Band Operation 1375
Md. Abdul Kaium Khan, Md Ibrahim Ullah, Towqir Ahmed Shaem, Rifat Kabir and Mohammad Abdul Alim

Performance Analysis of Tx-Rx Isolated Distributed Antenna System Implementing In-band Full-duplex for Up-link Communication to Mitigate Self-interference in 5G 1381
Ruhin Chowdhury and A. K. M. Sharoar Jahan Choyon

Performance Analysis of a CPDM-QPSK Direct Detection Optical Transmission System under the effects of Cross-Polarization 1386
A. K. M. Sharoar Jahan Choyon, S M Raiyan Chowdhury and Ruhin Chowdhury

Design and Simulation of an UWB Vivaldi Antenna for Cancer Detection and Treatment 1390
Antora Hossain and Kamal Hosain

Hybrid-core Microstructure Fiber for Chemical Identification in Terahertz Regime 1395
Md. Ahasan Habib and Md. Selim Reza

Single Mode Flattened Dispersion Circular-Shaped Porous Core Photonic Crystal Fiber for THz Regime 1400
Md. Anwar Sadath, Sabbir Ahmed Khan, Md. Hafijur Rahman, Md. Tariqul Islam, Mohammad Saiful Islam and Shahadat Hossain

Optimization and energy management of hybrid renewable power generation using HOMER and FLC 1404
Nesar Uddin, Dr. Sayed Abu Nahian, Dr. Md. Saiful Islam and Jewel Sikder Joy

Eye Assistant 1410
Md. Amanat Khan Shishir, Fairuz Maesha Habib, Shahariar Rashid Fahim and Tanjila Farah

An Approach for Improving the Complexity of Longest Common Subsequence Problems using Queue and Divide-and-Conquer Method 1414
Ratul Bhowmick, Md Ibrahim Sadek Bhuiyan Mobin, Md. Sabir Hossain, Muhammad Kamal Hossen and Ahsan Sadee Tanim

Proficiency Assessment of Adaptive Neuro-Fuzzy Inference System to Predict Wind Power: A Case Study of Malaysia 1419
Md. Rasel Sarkar, Mst. Jesmin Nahar, Awatif Nadia, Md. Abdul Halim, S. M. Sajjad Hossain Rafin and Md. Mamunur Rahman

CASCADED H-BRIDGE MULTILEVEL INVERTER USING SVPWM MODULATION TECHNIQUE 1424
Robin Ray, Mashrur Shadh and Md. Shamim Reza

A Modified Nave Bayesian-based Spam Filter using Support Vector Machine 1429
Md. Sabir Hossain, Md. Zubair, Mohammad Obaidur Rahman, Muhammad Kamrul Hossain Patwary and Md. Golam Sarwar Rajib

COMPARISON OF LINEAR DISPLACEMENT MEASUREMENTS BETWEEN A MEMS ACCELEROMETER AND HC-SR04 LOW COST ULTRASONIC SENSOR 1436
Azmyin Md. Kamal, Salman Habib Hemel and Mohsan Uddin Ahmad

Students Evaluation in AGILE After Completing Software Engineering Course 1442
Sabbir M. Saleh and Supta Richard Philip

Improving Detection Accuracy for Imbalanced Network Intrusion Classification using Cluster-based Under-sampling with Random Forests 1448
Md. Ochiuddin Miah, Sakib Shahriar Khan, Swakkhar Shatabda and Dewan Farid

Thermal Characteristics of Hybrid Photovoltaic Solar Thermal System during Summer .. 1453
Md. Faisal Kader, Md. Awal Khan and Md. Habibul Huq

Seizure and Non-Seizure EEG Signals Detection Using 1-D Convolutional Neural Network Architecture of Deep Learning Algorithm 1460
Tanima Tasmin Chowdhury, Afrin Hossain, Celia Shahnaz and Shaikh Anowarul Fattah

An IoT based Smart Energy Meter for the Idle Use of Electricity: An Industry Perspective 1464
Dr. Abdul Kadar Muhammad Masum, Saquib Manzur, Md. Kalim Amzad Chy, Shahidul Islam Khan, Md. Golam Rabiul Alam and S.M. Taslim Reza

Data Mining Techniques for Predicting User Interest in Facebook Pages: A Comparison . 1470
Nusrat Jahan Farin, Morium Akter, Puthi Roy and Mohammad Shorif Uddin

Design of CMOS On-Chip Transformer Coupled Matching Network for Millimeter-Wave Amplifiers with Optimal Chip Area

Md. Sazzad Hossain[1], Minoru Fujishima[2], Takeshi Yoshida[2], Shuhei Amakawa[2], Md. Mamunur Rashid[1]
Dept. of Electrical and Electronic Engineering[1], Semiconductor Electronics & Integration Science[2]
The International University of Scholars, Bangladesh[1], Hiroshima University[2], Japan[2]
{mshossain.icd@, mrashiddm@}gmail.com[1], {fuji@, tyoshida@dsl. ,amakawa@} hiroshima-u.ac.jp[2]

***Abstract*— Transformer coupled matching network is one of the alternative topologies for millimeter-wave integrated circuit design. Its impedance (inter-stage) transforming network (ITN) have been examined by planar and stacked topology with dimensional changing where planar transformer gives comparatively better performance than the stacked one. One-to-one-turn transformer-coupled ITN has been evaluated by the physical properties on optimum dimensional changing. Where ITN parameters and its properties function have been set below self-resonant frequencies (SRF) at 60 GHz & 79 GHz amplifiers. The quality factors (Q_f)and insertion losses (IL)of transformer ITN are $Q_{f@60GHz}$>20, insertion loss IL_{60GHz}<4dB and $Q_{f@79GHz}$>22, IL_{79GHz}<3dB at their center frequency. Also, parasitic inter-winding capacitances between primary and secondary coil are minimum because of the phasor angle between these two-coil voltage goes near to 0°. At millimeter-wave, the parasitic resistances of transformer coupled ITN are coming down <3Ω form tee-type equivalent circuits of impedance matrices. Extracted Optimal values of capacitances, resistances and inductances help to increase mutual coupling co-efficient as well as a good quality factor. Proposed planar and stacked small transformers have been utilized on ITN matching section of two millimeter-wave amplifiers on 55nm CMOS process. Supplying with 0.9V, mm-wave transformer coupled amplifiers have been designed layout area to 0.77 mm^2 with 34.51 dB peak gain at 60 GHz and 0.46 mm^2 with 17.15dB peak gain and both contain 24GHz 3dB bandwidth. As a result, it is confirmed that transformer coupled planar and modified stacked transformer are effective for shrinking layout area and high gain wideband 3dB BW in millimeter-wave regime.**

Keywords—Millimeter-Wave Transformers; Matching Networks; Impedance Transforming Network (ITN); Stacked & Planar Transformers; Transformer Coupled Balun & Power Divider; Millimeter-Wave Amplifiers.

I. INTRODUCTION

Millimeter-wave amplifiers are one of the key element for transceiver design. There are several topologies of mm-wave amplifiers with different circuit techniques. The single ended design technique suffers from poor noise performance and high area consumption due to the bulky passive devices. But differential technique can solve several problems behind this. Some previous Inductive-Capacitive (L-C-L) coupling, multi-section coupled lines, metal insulator metal (MIM) capacitor transmission line face a larger layout area and comparatively more insertion loss due to matching with ITN and input-output side matching network. Compared with those coupling [1] [2] [3], Transformer coupling network with cross coupling neutralizing capacitor (CCC) (fig.1) is one of the circuit technique for a good mutual coupling, impedance matching, minimize insertion loss, single to differential conversion, power distribution and combination. It helps to reduce layout area, minimizing insertion loss, a good quality factor and impedance matching with optimal mutual coupling co-efficient and lessen parasitic inter-winding capacitance and evaluation of lower resistances and optimal capacitances of each coils which have been considered for designing mm-wave devices. Along with other passive components, an arbitrary transformer coupled network is also used for impedance matching with a fixed material property changes. A center-trapped transformer coil can be easily set with differential ended active element. The design of several type of transformer is a one part of this paper.

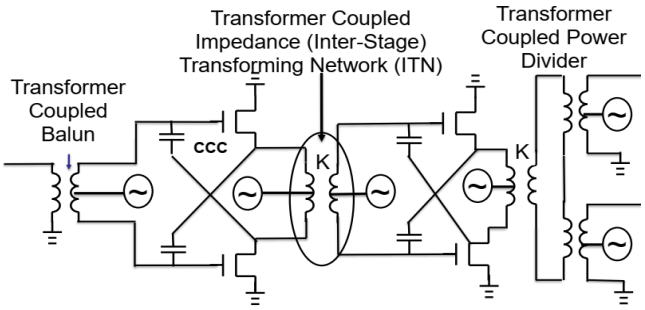


Fig. 1. Transformer coupled ITN with differential amplifier

Regarding the HFSS & Cadence layout of each coils, their relative location, geometric dimension, substrate shielding, transformation ratio, physical properties above. Similarly Transformer coupled balun and power divider provide much smaller amplitude & phase imbalance within larger bandwidth and consume less layout area rather than other topologies. Transformer practical circuit is little bit different from ideal equivalent circuit due to real time application. It contains several characteristics due to substrate capacitance (C_{SUB}), intra metallic capacitance (C_{Intra}), oxide capacitance (C_{OX}) which add several factors to degrade performance of passive elements and amplifier circuits (eqn. 1).

$$f_{prac}(L_P, L_S, R_P, R_S, C_P, C_S) = f_{ideal}(L_P, L_S, R_P, R_S, C_P, C_S) + Z_{SUB} + Z_{Intra} + Z_{OX} \quad \text{------ (eqn. 1)}$$

Maintaining optimal performances of mm-wave transformer and the physical properties evaluation with dimensional changing is one of the way to set a high-quality and small area on-chip transformers. In a brief, the goal of

this paper is the physical properties analyzing with On-Chip CMOS transformers with planar and stacked design. And finally realizing on-chip transformer coupled matching network for passive elements, millimeter-wave amplifiers design with higher gain, wideband 3dB BW and shrinking layout area at 60GHz and 79GHz on U and V band.

II. LITERATURE REVIEW

100 GHz 90nm CMOS amplifier uses constant planar transformer topology containing 11GHz 3dB BW, 11 dB gain, layout area is 0.87mm^2, 1.2V satisfying a good coupling co-efficient and set impedance transformation ratio by changing coil turns where higher SRF is important to minimize the loss [4]. Using a single steak of planar design reduces SRF and quality factor. Not even various dimensional planar and stacked topology have been described here. 60 GHz 65nm CMOS wideband differential driving amplifier measure wideband and high gain performance with multi-section coupled lines (Peak Gain: 20dB, 3dB BW: 10~15 GHz, 0.70mm^2) which faces quality factor degradation due to using RF grounded. Firstly, they regard transformer face defected ground effect on mutual coupling, increasing resistance and inductance and that will degrade the performance of amplifier at higher frequency [3]. From this paper, the loss has been minimized with higher SRF and having a good quality factors for transformer coupled inter-stage transforming network design. Transformer based 90nm CMOS at 60GHz amplifier (peak gain: 22dB, 0.27mm^2) stacked rectangular transformer has been visualized as coupled line segments. Topology is four-stage differential amplifier with 22dB gain at 60GHz. Here transformer is divided into three parts which includes two coupled lines and inductive lines to the terminals [5]. Here, physical properties with dimensional change of transformer isn't clarified using geometrical dimensional equation. Although it's tough to re-generate a general basement for transformer optimization but analysis aiding with some physical properties would be helpful to solve the different patterns.

III. PROPOSED MILLIMETER-WAVE TRANSFORMER COUPLED MATCHING NETWORK

A. *Millimeter-Wave Transformers*

Transformer is a passive element having primary and secondary impedance (L_p, L_s), primary & secondary resistance (R_s, R_p) and capacitance (C_p, C_s), parasitic coupling co-efficient, mutual coupling co-efficient (k_{imag}), turn ratio (n), quality factor (Q) and transformer efficiency (fig.4). Depending on whether the lateral or vertical magnetic coupling is used, transformer structures have two categories: planar and stacked. Several circuits (e.g., differential circuit) prefer transformers with a symmetrical layout.

Planar transformer topology (fig.2) contains top metal and thickness of 55nm CMOS process. There are many variations in planar transformer realization. Planar transformer usually occupies a large area. In multilayer process, several top metal layers can be strapped together to reduce resistance. The tradeoffs are a large terminal-to-substrate capacitance and a smaller self-resonant frequency [6]. But in mm-wave transformer coil has been chosen from top (TNA, Al, 2.8um) metal 12 only for a horizontal transformer design. The planar transformer is octagonal shaped having a space between primary and secondary coil maintaining design rules.

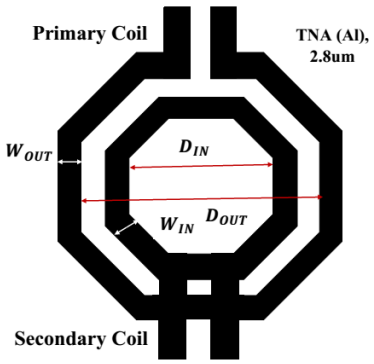


Fig. 2. Octagonal Planar Transformer with Top metal (TNA, Al, 2.8um)

A stacked transformer using two metal layers is shown in (fig.3) with utilizing vertical and lateral magnetic coupling, ideally self-inductance and coupling co-efficient are high. Millimeter-wave stacked transformer coil has been chosen from primary top (TNA, Al, 2.8um) metal 12 and Secondary metal 11 (LDA, Cu, 3.3um) without any vias (TWA) in a vertical design. The stacked transformer is octagonal shaped having a space between primary and secondary coil maintaining design rules. From fig.3, if the number of metal (below metal number-10) increases both in planar and stacked topology, it creates different properties. If several metals are used for stacked transformer, resistance decreases but capacitance increases. So, several mutual Intra & Inter metallic coupling inductances will generate and it will deteriorate the self-resonance frequency and insertion loss will be high enough because of using more thick metal. Because of some physical properties imbalances of multi turn planar and stacked transformer, achieving good quality factor, optimal mutual coupling and higher SRF, 1:1 turn is the reasonable choice of millimeter wave application arenas with their acute dimensional changing.

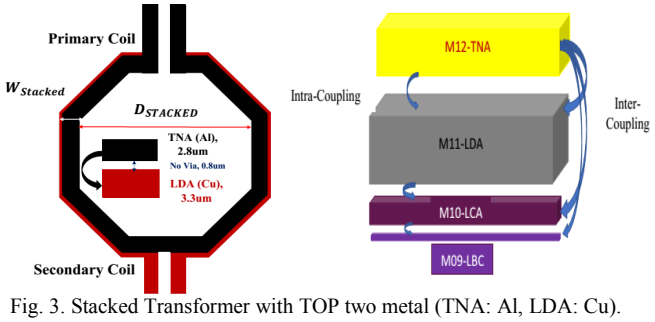


Fig. 3. Stacked Transformer with TOP two metal (TNA: Al, LDA: Cu).

B. *Ideal ITN Circuit Optimization*

Millimeter-wave Transformer based circuit can be represent into Tee & Pi type fourth order band pass filter and giving prime focus on mutual coupling co-efficient, self-inductances variation with an optimized goal, the matching section can be illustrated perfectly. Here, Primary, secondary & mutual inductance is gradually noted by $L_A = L_P - L_C$,$L_B = L_S - L_C$, $L_C = k_{imag}\sqrt{L_P L_S}$ and these two coils with 6 port networks (Port1_P, Port2_P, Port1_S, Port2_S, V_{ds}, V_{gs}). The Z parameters can be defined by eqn.2. From HFSS & Cadence MOMEMTUM cell, practical model of transformer with center trapped 6 port S-parameters has been extracted from simulation environment. The center trapped signal lines have related to supply source (gate & drain voltage).

$$\begin{bmatrix} V_P \\ V_S \end{bmatrix} = \begin{bmatrix} j\omega L_P & j\omega L_C \\ j\omega L_C & j\omega L_S \end{bmatrix} \begin{bmatrix} I_P \\ I_S \end{bmatrix} \quad \text{(eqn. 2)}$$

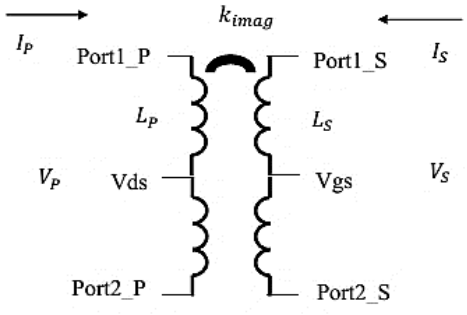


Fig. 4. Ideal circuit model of transformer

From 6 port S parameters MOM cell, V_{ds} & V_{ds} port have been taken ground initially to evaluate four port S matrices. Using theory of mode transformation [7] four port S matrix has been orthogonally transformed into even-even, even-odd, odd-even & odd-odd matrices, and odd-odd has been shown here (eqn.3).

$$S_{oo} = \begin{bmatrix} S_{o1o1} & S_{o1o2} \\ S_{o2o1} & S_{o2o2} \end{bmatrix} = \frac{1}{2}(S_{11} - S_{21} - S_{12} + S_{22}) \quad \text{(eqn. 3)}$$

Here, transformer Tee type equivalent circuit (fig. 5) has been represented from practical design cell. In Tee type transformation, it contains resistance (R_P, R_S), leakage inductance (L_{lecp}), self-inductance (L_A, L_B) and mutual inductance (L_C). Mutual Coupling Co-efficient k_{imag} will be calculated from Z matrices (eqn.4).

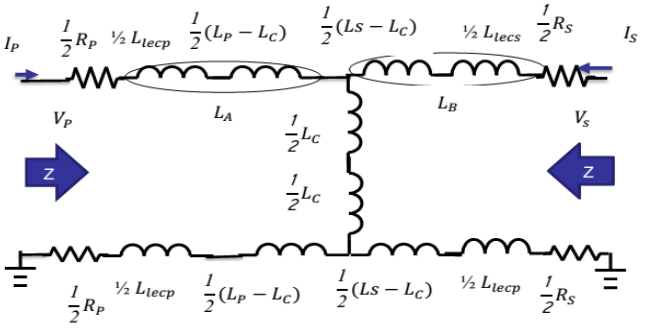


Fig. 5. Tee Type equivalent ideal transformer circuit

$$Z_{11} = a_{1t} + jb_{1t} = \frac{V_p}{I_p} = R_P + j\omega L_P;\ V_s = 0$$

$$Z_{12} = a_{2t}' + jb_{2t}' = \frac{V_p}{I_s} = j\omega L_C;\ V_p = 0$$

$$Z_{21} = a_{1t}' + jb_{1t}' = \frac{V_s}{I_p} = j\omega L_C;\ V_s = 0$$

$$Z_{22} = a_{2t} + jb_{2t} = \frac{V_s}{I_s} = R_S + j\omega L_S;\ V_p = 0 \quad \text{(eqn. 4)}$$

Extracted Parameter could be expressed in eqn.5,

$$L_A = \frac{im(Z_{11})}{\omega} + L_{lecp};\ L_B = \frac{im(Z_{22})}{\omega} + L_{lecp} \quad \text{(eqn. 5)}$$

$$R_P = a_{1t};\ R_S = a_{2t}$$

$$L_P = L_A + L_C;\ L_S = L_B + L_C;\ L_C = k\sqrt{(L_P L_S)}$$

$$k_{imag} = \sqrt{\frac{im(Z_{12} \text{ or } Y_{12}) * im(Z_{21} \text{or } Y_{21})}{im(Z_{11} \text{or } Y_{11}) * im(Z_{22} \text{or } Y_{22})}}$$

Pi type transforming network (fig.6) behaves as like band pass filter with RLC network. From Pi type equivalent circuit, mutual coupling co-efficient (k_{imag}) can be determined from two port Y matrices (eqn.6). And from here, parasitic capacitance (C_P, C_S) along with resistance (R_P, R_S) and self-inductance (L_P, L_S) can also be calculated from following equations and extracted parameters have been taken match with tee type transformation one.

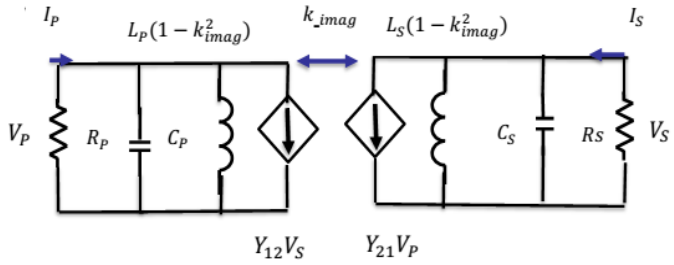


Fig. 6. Pi type equivalent ideal transformer circuit

$$Y_{11} = \mathbf{a_{1y} + jb_{1y}} = \frac{I_P}{V_P} = \frac{1}{R_P} + j\omega C_P + \frac{1}{j\omega L_P(1-k_{imag}^2)};\ V_S = 0$$

$$Y_{12} = \mathbf{a_{2y}}' + \mathbf{jb_{2y}}' = \frac{I_P}{V_S} = \frac{k}{j\omega L_S(1-k_{imag}^2)};\ V_P = 0$$

$$Y_{21} = \mathbf{a_{1y}}' + \mathbf{jb_{1y}}' = \frac{I_S}{V_P} = \frac{k}{j\omega L_P(1-k_{imag}^2)};\ V_S = 0 \quad \text{(eqn.6)}$$

$$Y_{22} = \mathbf{a_{2y} + jb_{2y}} = \frac{I_S}{V_S} = \frac{1}{R_S} + j\omega C_S + \frac{1}{j\omega L_S(1-k_{imag}^2)};\ V_P = 0$$

Primary & secondary inductance (L_p, L_S), Parasitic resistance and capacitance (C_P,C_S,R_P,R_S) and quality factor (Q_P, Q_S) have been expressed by these following equations.

$$L_P = \frac{-k_{imag}}{\omega b_{1y}'(1-k_{imag}^2)};\ C_P = \frac{1}{\omega}\left(b_{1y} - \frac{b_{1y}'}{k_{imag}}\right);\ R_P = \frac{1}{a_{1y}}$$

$$L_S = \frac{-k_{imag}}{\omega b_{2y}'(1-k_{imag}^2)};\ C_S = \frac{1}{\omega}\left(b_{2y} - \frac{b_{2y}'}{k_{imag}}\right);\ R_S = \frac{1}{a_{2y}}$$

$$Q_p = \frac{\omega L_p}{R_p};\ Q_s = \frac{\omega L_s}{R_s}$$

Cross coupling neutralized techniques for CCC means to cancel out the Gate-Drain Capacitance between a transistors at high frequency and increase transistor maximum frequency (f_{max}). And insertion loss (IL) can be calculated by maximum available gain (MAG) (eqn.7),

$$IL = -10\log_{10}(G_{max}) \quad \text{(eqn.7)}$$

$$G_{max} = 1 + 2(x - \sqrt{x^2 + x}), \text{ Where}$$

$$x = \frac{Re(Z_{11})Re(Z_{22}) - [Re(Z_{11})]^2}{[Im(Z_{12})]^2 + [Re(Z_{12})]^2}$$

$$\text{or}, x = 1/k_{imag}^2 Q_p Q_s.$$

C. *Practical ITN Circuit Optimization*

Planar and stacked designed transformer coupled ITN coils ware taken into vertical & horizontal view (fig. 7). In the oxide layer, these two types of coils have been set in an un-similar thickness but same dimensional area and same width. And the larger self-resonance frequency (SRF) and

good quality factor (Qf) have been determined from their simulated S matrices. So, from simulated result, stacked transformer SRF is lower than the planar one. On the other hand, stacked quality factor is better than planar transformer but doesn't have wideband characteristics. For wideband amplifier application, the trans-impedance of planar transformer give the broadband characteristic remain in 1GHz - 200GHz where stacked transformer contains in 1GHz – 70 GHz both having same structural area (fig. 8 & 9).

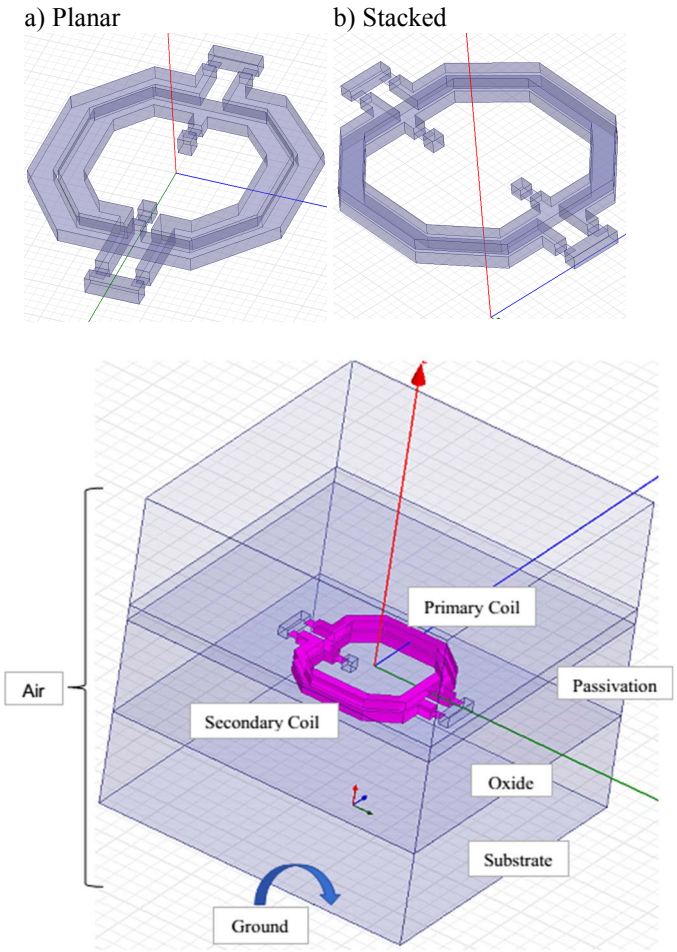


Fig. 7. (a & b) Planar and Stacked Transformer practical cell and final HFSS designed transformer cell

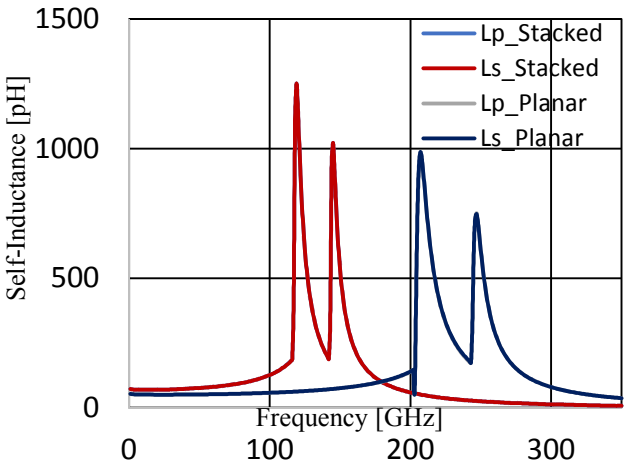


Fig. 8. Self-resonance frequency (SRF) of practical transformer cell

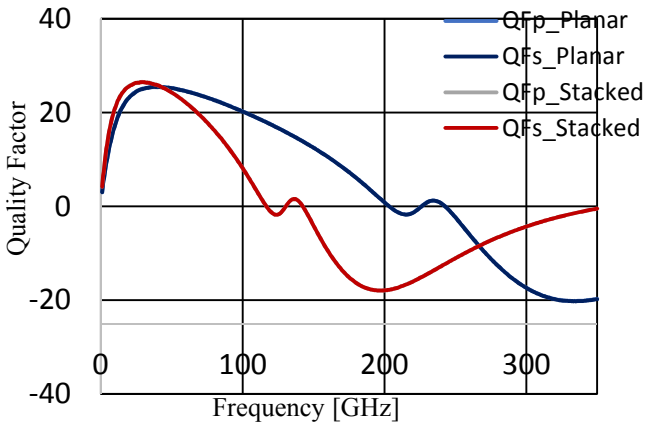


Fig. 9. Quality factor of practical transformer cell

Based on fixed dimensional planar and stacked transformer cell physical property, a geometrical boundary has been optimized. Upon this boundary and checking chip design rules, several planar and stacked transformers has been evaluated for amplifier design in mm-wave region. From (fig. 10, 11 and 12), a decision could be finalized with planar transformer because of its continuous good quality factor, lower insertion loss and higher SRF. On the other hand, a smaller stacked transformer can perform same characteristics as like planar beyond the edge of design rule. Such as 40um stacked one give same SRF as like 70um planar one.

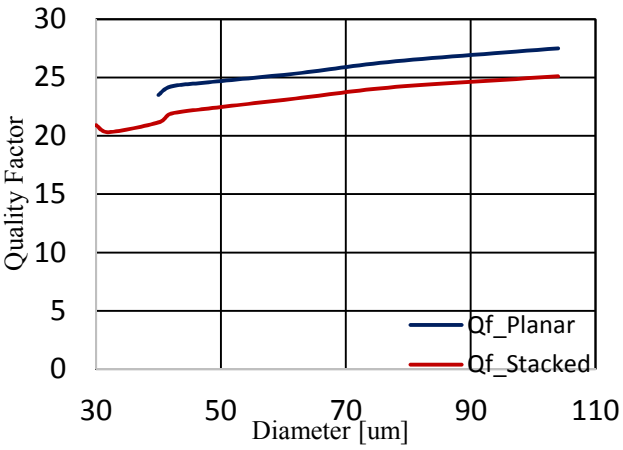


Fig. 10. Optimized practical transformer cells quality factor

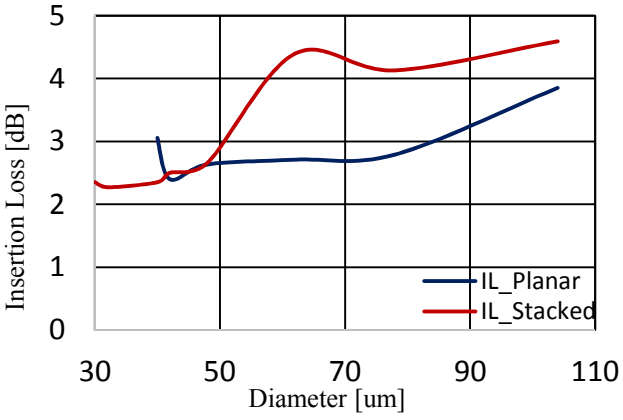


Fig. 11. Optimized practical transformer cells insertion loss

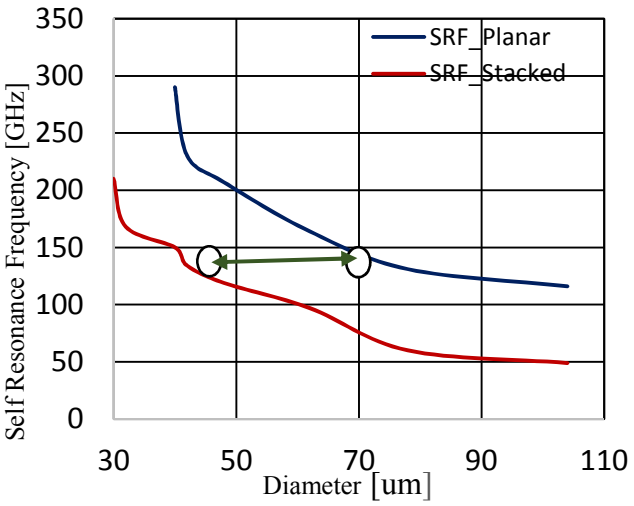


Fig. 12. Optimized practical transformer cells Self-resonance Frequency

IV. RESULTS OF TRANSFORMER COUPLED TECHNIQUE ON MILLIMETER-WAVE AMPLIFIERS

Transformer coupled ITN has been set with no ground wall and ground floor but XPWA to fix a good quality factor. Although forward (FC) and backward (BC) connector create a little bit more insertion loss into the circuit, but this phenomenon might be slightly change coupling co-efficient. And two differential pair transistors with fingers {(32X2), (64X2)} have been attached with periodical changing of material diameter and width of planar and stacked transformer (fig.13).

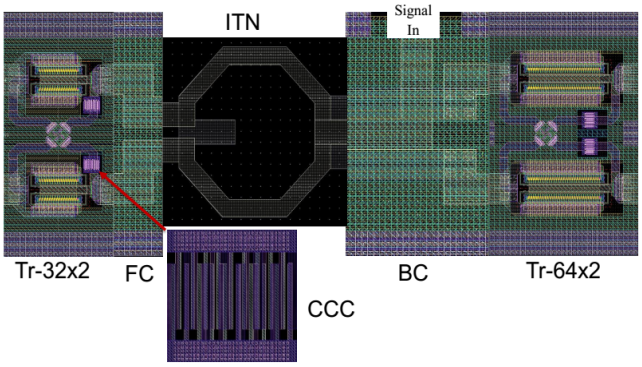


Fig. 13. Layout of two stage single amplifier with transformer coupled ITN

A. Transformer Coupled Amplifier at U and V Band

Here fig. (14 & 18), transformer coupled amplifier schematic at 60GHz and 79 GHz have been designed with Q_{f60GHz}>20dB, optimal IL_{60GHz} < 4dB and Q_{f79GHz}>22dB, IL_{79GHz} < 3dB and both amplifier satisfied optimal K_{imag}~0.4. Parasitic inter-winding capacitance angles are nearly to 0^0 which means minimum capacitance shows between two windings and $V_{CC} = 0$. Planar transformers are being geometrically designed to sustain itself at higher self-resonance frequency. But smaller stacked will provide good performance as like planar under a certain design rules. From (fig. 15 & 19), total layout area has been estimated into 1673 um^2 × 460 um^2 = 0.77 $mm^2$and 3dB BW 24GHz in U with a 34.51 dB peak gain (fig.16) at 60GHz. Again at 79GHz amplifier layout size is 1006 um^2 × 455 um^2 = 0.46 $mm^2$and 3dB BW 24GHz in V band with a 17.15 dB peak gain (fig.20). Transformer coupled amplifier at 60GHz and 79GHz comparison diagram have been viewed here (fig. 17 & 21). And these transformer coupled amplifiers have been compared with other transformer coupled and several topologies in case of peak gain, 3dB BW and less layout area. Here, a smaller dimensional stacked transformer with diameter and width changing provide similar characteristics as planar one in the millimeter-wave application.

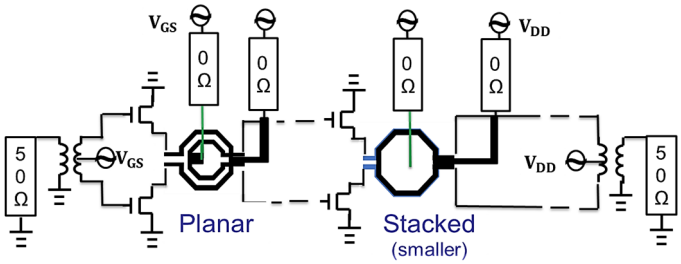


Fig. 14. Schematic of 60 GHz Transformer coupled amplifier

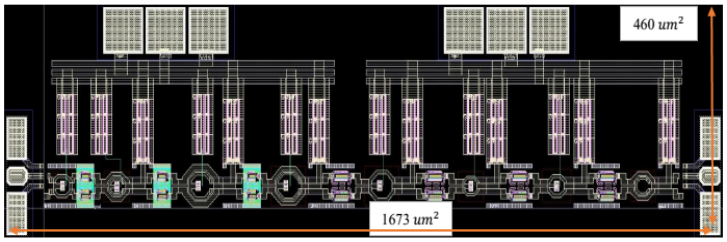


Fig. 15. Layout of 60 GHz Transformer coupled amplifier

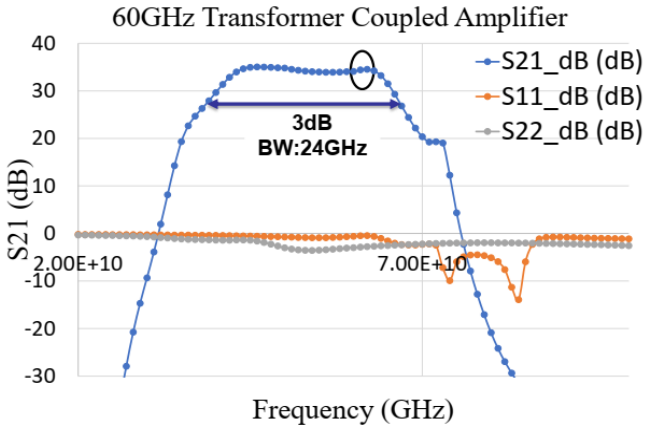


Fig. 16. 60 GHz Transformer coupled amplifier gain

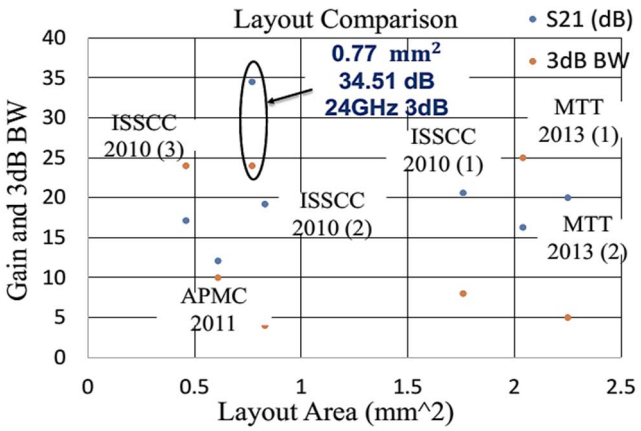


Fig. 17. 60 GHz Transformer coupled amplifier comparison diagram

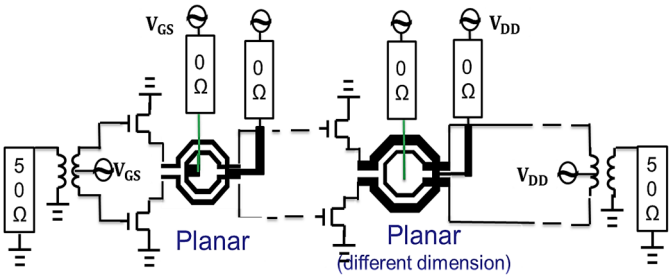


Fig. 18. Schematic of 79 GHz Transformer coupled amplifier

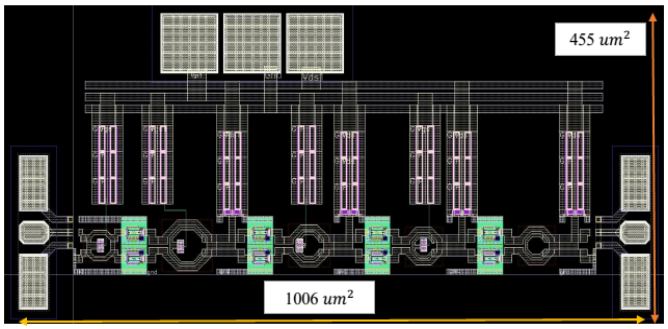


Fig. 19. Layout of 79 GHz Transformer coupled amplifier

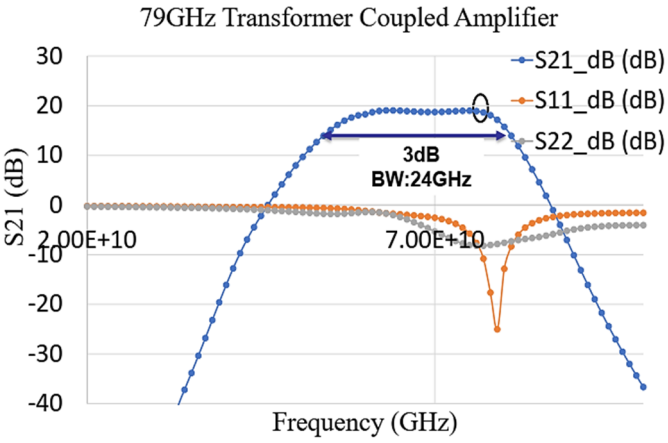


Fig. 20. 79 GHz Transformer coupled amplifier gain

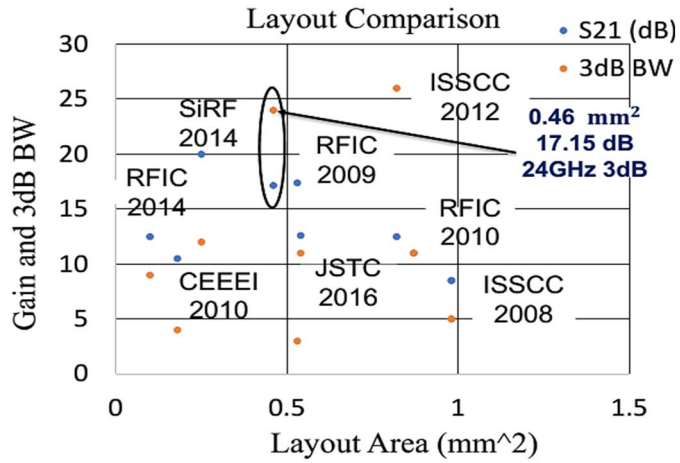


Fig. 21. 79 GHz Transformer coupled amplifier comparison diagram

VI. CONCLUSION

In this simulated work on cadence, high-quality and small-area on-chip transformers are proposed and optimized for differential amplifier. Proposed transformers are used as passive matching circuits and have advantage to compare with traditional matching circuits to reduce chip area. For millimeter-wave amplifiers, transformer-coupled planar and stacked type inter-stage transforming network (ITN) has been evaluated. As a result, wideband trans-impedances of fourth-order ITN is realized with a good quality-factor and low insertion-loss transformer by miniaturizing parasitic inter-winding capacitance and resistance with high self-resonant frequency.

ACKNOLEDGEMENT

The author would like to thank the member of Frontier Integrated System Laboratory, AdSM, Hiroshima University, along with MIE Fujitsu Semiconductor Inc. Japan, for cadence 55nm DDC CMOS IC design platform and their design rules.

REFERENCES

[1] Shinsuke Hara et. al, "Compact 141-GHz Differential Amplifier with 20-dB Peak Gain and 22-GHz 3-dB Bandwidth", IEICE Trans. Electron, Vol. E99-C, No.10, Oct. 2016.

[2] K. Katayama, S. Amakawa, K. Takano, T. Yoshida, M. Fujishima, "An 80-106 GHz CMOS Amplifier with 0.5V Supply Voltage", RFIC, 2017

[3] Haiwei Zhang, Quan Xue, "60 GHz CMOS wideband differential driving amplifier using multi-section coupled lines", IEEE MTTs, Vol. 25, No.9, Sept. 2015.

[4] Noel Deferm, Patrick Reynaert, "A 100 GHz Transformer-Coupled Differential Amplifier in 90nm CMOS", RFIC, 2010.

[5] Amin Hamidian, "High Performance Transformer based mm-wave CMOS PA", 8th EMICs, Nuremberg, Germany, 2013.

[6] Bernardo Leite, "Design and modelling of mm-wave integrated transformer in CMOS and BiCMOS technologies", Electronique. Universite Sciences et Technologies- Bordeaux I, 2011, France.

[7] Shuhei Amakawa, Noboru Ishihara, Kazuya Masu, "A thru-only de-embedding method for on-wafer characterization of multiport networks" TokyoTech, INTECH.

[8] C.Y. Law et. al, "A high gain 60GHz power amplifier with 20dBm output power in 90nm CMOS", ISSCC 2010.

[9] J. Lai et al, "A 1V 17.9dBm 60GHz Power amplifier in standard 65nm CMOS", ISSCC 2012.

[10] B. Martineau et al, "A 53-68 GHz 18dBm PA with 8-way combiner in standard 65nm CMOS", ISSCC 2010.

[11] Ryo Minami "A 60GHz CMOS power amplifier using varactor cross-coupling neutralization with Adaptive bias", APMC 2011.

[12] Yuan-Hung Hsiao et al, "mm-wave CMOS power amplifiers with high output power and wideband performances", IEEE Trans. MTTs, 2013.

[13] S. Aloui et.al, "High gain and linear 60GHz power amplifier with a thin digital 65nm CMOS", IEEE Trans. MTTs, 2013.

[14] Jeng-HanTsai,"A 69-81 GHz PAusing90nmCMOS",SiRF 2014.

[15] E. Kaymaksut, "E band transformer Doherty PA in 40nm CMOS", RFIC 2014.

[16] T. Suzuki et al, "60 & 77GHz PA in 90nmCMOS", ISSCC 2008.

[17] Seungwon Park et al, "A transformer-matched mm-wave CMOS PA",JSTC, October 2016.

[18] K.Y. Wang, "A 1V 19.3dBm 79GHz PA in 65nmCMOS", ISSCC 2012.

[19] Amit Vishnipolsky, "A compact power efficient transformer coupled differential W-band CMOS Amplifier", CEEEI Israel, 2010.

Enhancement of Global Bandwidth of Internet Users Employing Internet Exchange Point Leveraging VPN Technology

Al- Nahian Bin Emran
Department of CSE
Daffodil International University
Dhaka-1207, Bangladesh
email: nahian252@gmail.com

Aman Ullah Juman
Department of CSE
Daffodil International University
Dhaka-1207, Bangladesh
email: aman@amanullahjuman.me

Gazi Zahirul Islam
Department of CSE
Daffodil International University
Dhaka-1207, Bangladesh
email: zahir.cse@daffodilvarsity.edu.bd

***Abstract*—Not all internet service providers provide the same amount of bandwidth to their users. Based on bandwidth allocation, location and price service providers capped their users' bandwidth. As a result, we observe bandwidth limitation among the users. In this paper, we will implement a technique to bypass bandwidth and firewall restriction using virtual private network where we will use common IXP route to avoid the initial level of the bandwidth cap. In this project, we show that internet speed of international route i.e. global bandwidth can bypass, if there are multiple ISPs use a common IXP and at least one of those ISP provide robust global bandwidth. Usually, regional ISPs use common IXP to route their local traffic within the region. We utilize IXP to bypass global bandwidth cap of users who have better local bandwidth allocation. We show that using our proposed method global internet speed of a user can raise several times effectively.**

Keywords—IXP, VPN, BDIX, Global Bandwidth, Local Bandwidth, ISP

I. INTRODUCTION

It is undeniable that the Internet is a ground-breaking technology in this century. We can hardly think of a day without using the internet. There's hardly any sphere in human life left where the internet hasn't brought any change yet. From the outer space to the deepest trench of the world, the internet has brought the world to our finger's end. For the greater welfare of mankind, the Internet should reach every people's hand at their greatest conveniences. But there are too many deficiencies in our system that has been a hindrance to spread the internet at a convenient cost. Pointing out all of the deficiencies of the internet is not the main focus of this paper. In this paper, we came up with a solution to make the Internet handy. In our country Internet bandwidth is not well distributed with a well-maintained pricing list. Pricing of packages varies from one ISP to another which is making conundrum to the users. Peoples are not getting proper bandwidth by paying money which they yearn. Somewhere they are getting fair service for internet packages by paying decent amount of money and somewhere they cannot think of getting the same service with the same amount of money. So, focusing on the pricing variations, we will try to find a convenient solution to get superior global speed by using the Internet Exchange Point (IXP).

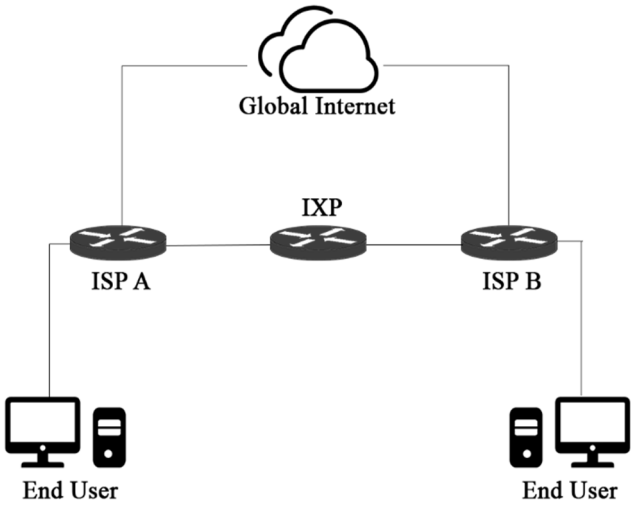


Fig. 1. Basic schematic diagram of an IXP

Internet exchange point was invented to affect the speed of the internet and give a whole new infrastructure of communication. A typical IXP schematic diagram is shown in *Fig. 1*. Here IXP servers as a common exchange point of the internet where Internet Service Providers (ISPs) are interconnected via IXP. ISPs are also connected to each other through global internet cloud [1]. An IXP interconnects ISPs in a region or country, allowing them to exchange domestic Internet traffic locally without having to send those data across multiple international gateways to reach their final destination [13], [14].

As IXP ensured us better local traffic, our goal is to use local traffic speed to get over with global bandwidth syndication and ensure a thorough distribution and maximized speed of global traffic with a decent amount of money. The details of the protocol will be explained in 'System Design and Implementation' section i.e. section IV.

TABLE I. AN ESTIMATE OF INTERNET USERS AND THEIR GROWTH

Year	Number of users	Population	% Penetration
2000	93,261	131,280,739	0.1 %
2005	345,372	142,929,979	0.2 %
2010	5,609,821	151,616,777	3.7 %
2014	15,271,441	159,077,513	9.6 %
2015	19,420,674	160,995,642	12.1 %
2016	21,439,070	162,910,864	13.2 %

The rest of the paper is organized as follows. We first discuss the background in section II then elaborate related works and motivation in section III. Section V contains performance evaluation and comparison. Constraints and recommendations are discussed in section VI. Finally, section VII concludes the paper.

II. BACKGROUND

Bangladesh is passing through some massive technological changes. From nineties' VSAT communication system to today's Internet accessibility through submarine cable, our country's communication pattern faced many challenges and successfully overcame those obstacles by adopting 3G, Wi-Max and LTE technologies. After introducing the submarine cable system, the number of local ISPs grew up drastically. Before the introduction of IXP, local traffic has to travel to other countries to reach its final destination. As such prior infrastructure was not able to provide a cost-effective service. The internet penetration of Bangladesh is one of the lowest in the world, only greater than North Korea, Myanmar and Sierra Leon which are exceptional cases [12]. In one of the press releases, the CEO of Grameenphone enlighten us with impressive information. He forecasts that by the year 2020, the total number of internet user of Bangladesh will be nearly 20 million. An estimate of internet users and their growth is shown in *Table I* [3], [4]. We can see in the year 2000, only 0.1% of the total population were using the Internet. By the year 2016 the number of Internet users increases to 13.2% which is 132 times of the year 2000. The BTCL, WiMAX operators, mobile operators and numerous ISPs are trying to reach every corner of the country to expand their services to make the internet a daily useful companion.

As Internet Exchange Point has profound impact over local traffic, we propose a method where local traffic can be used for elevated global traffic by creating a third wheel [2]. The third wheel will use Virtual Private Network (VPN) or proxy server. If a user has higher local bandwidth than global bandwidth, our method can establish an elevated global traffic route by counting little extra money to VPN providers.

For example, one internet provider is serving their customer with 1 Mbps Global bandwidth along with 20 Mbps of local bandwidth for 500 BDT. On the other hand, another provider is providing his customer with 5 Mbps of global bandwidth and 30 Mbps of local bandwidth for the same amount of money. In this case, we observe both ISPs are providing sufficient bandwidth for local traffic, but their global bandwidth is not satisfactory yet. Again, there is a huge difference in global bandwidth between two ISPs. Now the first customer who is currently getting 1 Mbps global bandwidth from his ISP could get up to 5 Mbps global bandwidth from the second ISP using their VPN server.

III. RELATED WORKS AND MOTIVATION

Bangladesh Telecommunication Regulatory Commission (BTRC) has been working deliberately for making internet handy to everyone. BTRC had taken a quintessential step to make a revolutionary change in existing infrastructure since the day Bangladesh was connected to the international submarine cable 'SeaMeWe-4'. As a result of simultaneous effort, Bangladeshi ISPs had been connected to its own IXP known as Bangladesh Internet Exchange (BDIX), thereby Bangladeshi users enjoying tremendous internet speed for local traffic. Without IXP local traffic has to travel to other countries to reach its final destination. So IXP has been a revolutionary technology in the world of communication as it changed the whole communication pattern of the Internet. There has been a lot of researches and projects for deployment of IXP.

"Internet Exchange Point: A Context of Bangladesh" is one of the finest research paper based on IXP. In this paper, the authors described the role of IXP and also explained how a whole infrastructure of a county could be changed through the setup of IXP [1]. The authors also enlightened us how a cost-effective model could be achieved by deployment of Internet Exchange Point. In 2004, Russell Southwood published an outstanding article upon the request of the International Telecommunication Union (ITU) and the International Development Research Centre (IDRC) that addresses two issues of direct relevance to most developing countries [5]. One of the factors inhibiting the growth of Internet usage in these countries is the high cost of bandwidth and the usage of this international bandwidth to exchange data at a local and national level. The other being: the need to setup an IXP to accelerate the growth of local traffic and save incurring international bandwidth cost.

The author of [6] focuses on the problems and opportunities for IXP deployment in Africa. The paper includes two distinct case studies for Kenya and Bangladesh focusing on their local Internet traffic, cost of bandwidth and IXP deployment issues. Another paper titled "Economics of Internet Exchange Points" by Mouhamet Diop published in 2009 gives an in-depth analysis of Internet transit, internet peering ecosystem and the economic impact of an IXP in a country. The paper also discusses various issues concerning the creation and operation of an Internet Exchange and tries to come up with an appropriate business model for it [7].

IV. SYSTEM DESIGN AND IMPLEMENTATION

In this section we will delineate a general IXP system then we will describe and implement our proposed method elaborately.

A. Apparatus and Mechanism of Internet Exchange Point

Internet Exchange Point acts as a switch that routes traffic among its associated ISPs. It saves the bandwidth cost and time of large intermediate portion of internet flow from source to destination. It is mandatory that all intended ISPs must be connected thorough a common IXP. In *Fig. 2* Tier 1 providers

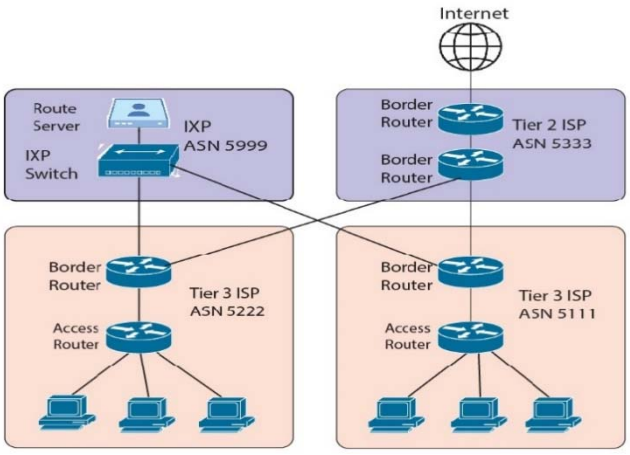


Fig. 2. A typical network with an IXP in the middle

TABLE II. TOP 20 WEBSITES BROWSES BY BANGLADESHI USERS

1.	google.com	11.	bd24live.com
2.	youtube.com	12.	teletalk.com.bd
3.	facebook.com	13.	bdjobs.com
4.	prothomalo.com	14.	blogspot.com
5.	jugantor.com	15.	jagonews24.com
6.	yahoo.com	16.	banglanews24.com
7.	kalerkantho.com	17.	24livenewspaper.com
8.	google.com.bd	18.	crickbuzz.com
9.	wikipedia.org	19.	bdnews24.com
10.	bd-pratidin.com	20.	daraz.com.bd

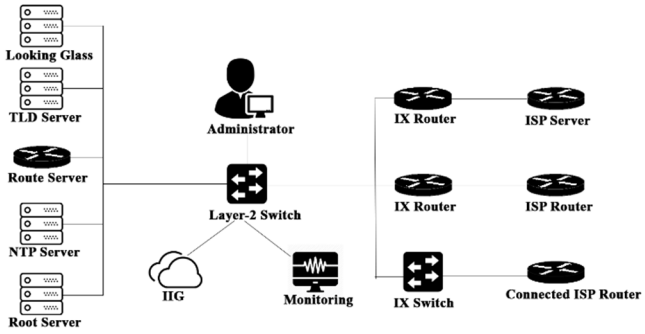


Fig. 3. Schematic diagram of an IXP

are the backbone of international traffic. Tier 2 is a thorough distributor of international traffic to Tier 3 by peering agreements. Tier 3 buys internet bandwidth from Tier 2 and provides internet connectivity to corporate and residences. Throughout the whole journey of traffic, IXP reduces the usage of global bandwidth by filtering out local traffic which previously uses the global route. Thus, the usage of international bandwidth reduces dramatically. If we investigate the top 20 websites searched by Bangladeshi users, we can see our Bangladeshi websites such as Prothom Alo, Kalerkantha, Jugantor, Bdjobs etc. are ruling in top 20. *Table II* shows top 20 websites which are surfed most of the time by Bangladeshi users [7]. The traffic of local websites need not travel through Tier 1 as this traffic will consume only local bandwidth using IXP. Thus, the Internet system using IXP could provide a faster connection in our country which could be implemented by any other country of the world.

For the establishment of an IXP, we require some elements such as root server, route server, NTP server, looking glass etc. [9]. These are illustrated in *Fig. 3*.

B. *Implementation Requirement*

The first and foremost requirement is both ISPs must be connected to the common IXP. For the third wheel that helps us to accelerate the global bandwidth, we need a VPN or proxy server. For our project, we rent a Linux VPS (Virtual Private Server) from the company 'Xeon BD' where we installed a VPN server. We have determined the result using our method after connecting to VPN server. Speedtest.net [11] is the website that we use to measure the acceleration of bandwidth with a Singaporean server named 'Singtel'.

C. *Implementation with an IXP in the middle*

Our aim is to maximize the global bandwidth leveraging VPN technology utilizing local bandwidth. A user who is getting poor global bandwidth from his current ISP can be benefitted by an IXP to boost up his global bandwidth. For that the user must be subscribed to another ISP that is providing higher global bandwidth from his current ISP and both ISPs must be connected to each other through the common IXP. Thus, a user getting a higher global bandwidth provided by his new ISP sitting in the network of his old ISP's and ISP to ISP traffic through IXP utilizes local bandwidth. The user has to take the services of a VPN or proxy server from the new ISP. If new ISP does not provide these services, then the user has to setup the VPN server by himself. When a user accesses the internet via a VPN server, he acts as a DHCP client to the VPN server. After connecting VPN server traffic between current ISP to VPN server (new ISP) will use user's local bandwidth portion. Now user's global traffic will be routed through the VPN server. *Fig. 4* illustrated the scenario.

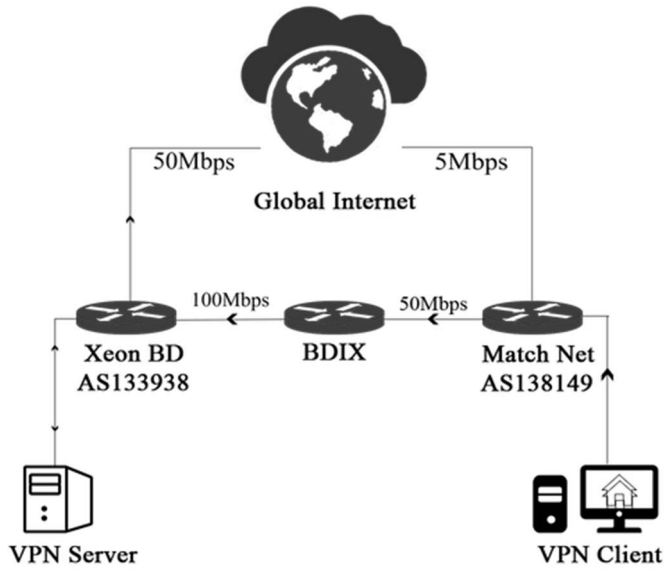


Fig. 4. A typical network of the proposed method with an IXP in the middle

TABLE III. DATA COLLECTED FROM DIFFERENT ISP

No	ISP Name	Package Specification	Package Cost	Obtained Global Internet Speed
01	Match Net	5 Mbps (Global Bandwidth) 50 Mbps (BDIX Bandwidth)	500 BDT	3.94 Mbps
02	ICC Communication	4 Mbps (Global Bandwidth) 30 Mbps (BDIX Bandwidth	700 BDT	3.96 Mbps
03	Explore Online	10 Mbps (Global Bandwidth) 100 Mbps (BDIX Bandwidth)	500 BDT	9.68 Mbps
04	Stargate Communication	2 Mbps (Global Bandwidth) 70 Mbps (BDIX Bandwidth)	500 BDT	2 Mbps

TABLE IV. DATA COMPARISON BEFORE AND AFTER USING PROPOSED METHOD

No	Source ISP	Obtained Global Internet Speed (Previous data)	Obtained Global Internet Speed (Applying VPN over IXP)	Improvement
01	Match Net	3.94 Mbps	30.55 Mbps	7.75 times faster
02	ICC Communication	3.96 Mbps	27.48 Mbps	6.94 times faster
03	Explore Online	9.68 Mbps	41.43 Mbps	4.27 times faster
04	Stargate Communication	2 Mbps	16.14 Mbps	8.07 times faster

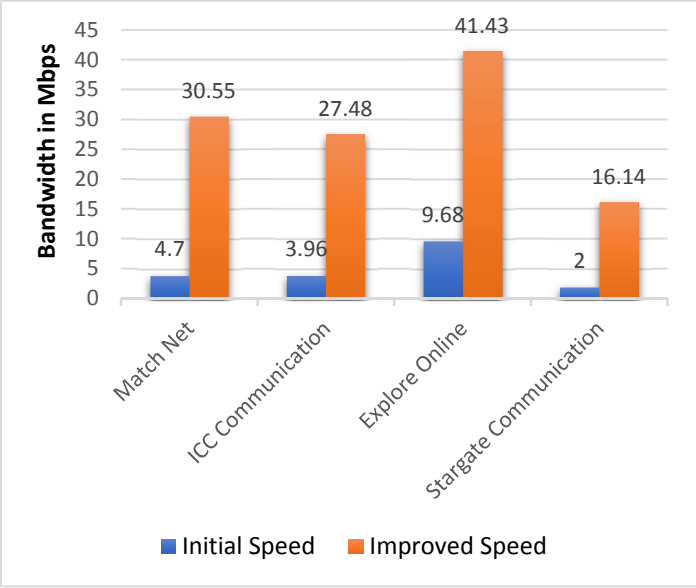


Fig. 5. Improvement of global bandwidth

Initially, we were taking internet services from an ISP named 'Match Net' located at Kolabagan, Dhaka that providing global bandwidth 5 Mbps along with 50 Mbps local bandwidth. To implement our technique, we rent a Linux VPS from another data center named 'Xeon BD' located at Mohammadpur, Dhaka. Xeon BD provides us 50 Mbps global bandwidth along with 100 Mbps local bandwidth for 400 BDT.

In VPS we installed the SoftEther VPN server which supports multiple types of VPN. In the server, we enabled L2TP, IPSec, SSTP, and OpenVPN. As VPS does not contain built-in DHCP server, we utilize virtual NAT function where we enable SecureNAT. Whenever a client connects the VPN server, the server will provide a private IP address from SecureNAT pool. In client end, we use SoftEther VPN Client application which uses SSTP (Secure Socket Tunneling Protocol) by default to connect VPN server.

V. PERFORMANCE EVALUATION AND COMPARISON

In this section, we will evaluate the performance of our proposed method and numerically determine the improvement of global bandwidth sitting at different locations of different ISP's in Dhaka City.

Table III shows different internet package information provided by some renowned ISP of capital city Dhaka. The data of the last column i.e. Obtained Global Internet Speed of *Table III* is measured by us using Speedtest.net website with the help of a Singaporean server 'Singtel'. We have gathered information of four ISPs' namely Match Net, ICC Communication, Explore Online and Stargate Communication. Their global bandwidth ranges from 2-10 Mbps, local bandwidth ranges from 30-100 Mbps, package cost ranges from 500-700 BDT and our measured global speed ranges from 2-9.68 Mbps.

Table IV shows the improvement of global bandwidth after using our method. To acquire higher global bandwidth, we connected our VPN server which hosted on Xeon BD from primary ISPs' user end. Again, we have tested global bandwidth by using Speedtest.net. It is noted that Xeon BD is providing a pleasing global bandwidth of 50 Mbps which is substantially larger than other ISPs'.

Examining *Table IV*, we can easily realize the differences in global speed using and without using an IXP and can perceive the effects of our proposed method. Now sitting at each of the primary ISP's network we are getting an enlarged global bandwidth which is around 4 to 8 times higher than the allocated bandwidth. *Fig. 5* summarizes the output of the

system graphically. The blue bars represent the global bandwidth which is given by the ISPs and the orange bars represent the results after employing our method.

Analyzing cost benefit, we can show that the bandwidth price will be significantly reduced after implementing VPN technology. According to the *Table III* and *Table IV*, we roughly estimate that after employing VPN, our cost increases around 1.75 times while speed increases 6.75 times. Thus, we must acknowledge the tremendous benefit of employing IXP leveraging VPN technology.

VI. Constraints And Recommendations

There are some constraints and limitations to implement the project using our intended method. Our method will work if and only if the ISPs are belong to a common IXP. For example, if one ISP only belongs to BDIX and another one only belongs to NOVOCOM NIX then the method will be void. Another thing is that local bandwidth of all ISPs must be good enough otherwise a bottleneck would occur. Lastly, users who want to experience an elevated global bandwidth must have to buy a VPN server from the ISP that providing a robust global bandwidth.

VII. Conclusion

The Internet has been a tremendous gateway for acquiring knowledge as well as a constant companion for entertainment. Our goal is to provide a method that can overcome bandwidth syndication, capping of bandwidth and provide every user a pleasant experience of using Internet with a maximized global bandwidth without paying large amount money.

In this article, we effectively illustrated a project that boosted global bandwidth using IXP. So our method can be used to resolve the syndication problem, deficiencies of internet speed and give comfort to the users for the next generation communications. In future, we will implement the project in a more flexible way so that any non-technical person can easily setup the system.

References

[1] Md. Shakilur Rahman, Muhtasim Iqbal and Nova Ahmed, "Internet Exchange Points: A Context of Bangladesh", 14th International Conference on Computer and Information Technology (ICCIT 2011), December 2011.

[2] M. Z. Ahmad and R. Guha, "Studying the effect of internet exchange points on internet link delays," Proceedings of the Communications and Networking Symposium (CNS), SCS SpringSim 2010, Orlando, FL, April 2010.

[3] Internetlivestats.com, "Bangladesh Internet Users", March 2018. [Online]. Available: http://www.internetlivestats.com/internet-users/bangladesh.html. [Accessed: 20- Nov- 2018].

[4] Cia.gov, "Central Intelligence Agency", 25 March, 2018. [Online]. Available: https://www.cia.gov/library/publications/resources/the-world-factbook/geos/bg.html. [Accessed: 29- Oct- 2018]

[5] Russell Southwood, "Via Africa Creating local and regional IXPs to save money and bandwidth", Global Symposium for Regulators by International Telecommunication Union, 2004.

[6] Andrew McLaughlin, "Internet Exchange Points Their Importance to Development of the Internet and Strategies for their Deployment - The African Example", Global Internet Policy Initiative, May, 2004.

[7] Mouhamet Diop, "Economics of Internet Exchange Points", Afrinic 10 Cairo, May 2009.

[8] Alexa.com, "Top sites in Bangladesh", January 8, 2018. [Online]. Available: https://www.alexa.com/topsites/countries/BD [Accessed: 15- Aug- 2018].

[9] Btrc.gov.bd, "BTRC NIX Licensing Guidelines (Draft)", March 16, 2018. [Online]. Available: www.btrc.gov.bd/sites/default/files/nix_guide0.pdf. [Accessed: 20- Oct- 2018].

[10] Bdix.net, "Bangladesh Internet Exchange (BDIX)", April 05, 2018. [Online]. Available: http://www.bdix.net. [Accessed: 30- Oct- 2018].

[11] Speedtest.net, "Speedtest by Ookla - The Global Broadband Speed Test", January 06, 2018. [Online]. Available: http://www.speedtest.net. [Accessed: 29- Dec- 2018].

[12] Wikipedia.com, "Telecommunications in Bangladesh", 2018. [Online]. Available: https://en.wikipedia.org/wiki/Telecommunications_in_Bangladesh.html. [Accessed: 18- Dec- 2018].

[13] K. Xu, Z. Duan, Z. L. Zhang and J. Chandrashekar, "On properties of internet exchange points and their impact on as topology and relationship," Networking, vol. 3042, 2002.

[14] Y. He, G. Siganos, M. Faloutsos and S. Krishnamurthy, "Lord of the links: a framework for discovering missing links in the internet topology," IEEE/ACM Trans. Networking, vol. 17, no. 2, pp. 391–404, 2009.

Comparison of Performance of Cognitive Radio Network under Single and Multi-User Scenario

Md Abul Kalam Azad
Department of Computer Science and Engineering,
Jahangirnagar University,
Savar, Dhaka-1342, Bangladesh
e-mail: makazad@juniv.edu

Md. Imdadul Islam
Department of Computer Science and Engineering,
Jahangirnagar University,
Savar, Dhaka-1342, Bangladesh
e-mail: imdad@juniv.edu

Muhammad R. A. Khandaker
School of Engineering and Physical Sciences,
Heriot-Watt University, Edinburgh,UK
e-mail: m.khandaker@hw.ac.uk

Jugal K. Das
Department of Computer Science and Engineering
Jahangirnagar University,
Savar, Dhaka-1342, Bangladesh
e-mail: cedas@juniv.edu

***Abstract*—In this paper, we analyze the performance of single user and multi-user communication model in LTE CRN links. The performance of the two models is compared in terms of outage probability and symbol error rate (SER) under small-scale fading including combining schemes. The impact of the number of antenna elements on combining schemes and two different fading environments have been investigated in terms of the above performance metrics. The performance of the communication system deteriorates under multi-user scenario because of inclusion of interference but the multi-user model represents more realistic scenarios of wireless links. We found that SER and outage probability of multi-user model comparable with single user case under multiple antennas. The analytical results are also verified by simulation with high confidence level.**

Keywords— LTE CRN, relay gain, outage probability, SER and MRC.

I. INTRODUCTION

Cognitive radio network (CRN) technologies have emerged with a handful of potentials to address spectrum scarceness in wireless communication systems [1]. CRN is often termed as *brain-empowered wireless communications* since the underlying technology enables unlicensed users gain control of the underutilized frequency band licensed to other users. However, with minor abuse, CRN has been adopted in long term evolution (LTE) systems, in which the primary users (PUs) can set up direct links with the evolved node B (eNB) but the secondary users (SUs) need one or more PU(s) as intermediary nodes. The SUs' ability to establish links with the home evolved node B (HeNB) directly remains unaffected.

In [1], the secrecy outage performance of a co-operative cognitive radio network is determined under Rayleigh fading channels. The system model in [1] consists of a secondary source, a decode-and-forward (DF) relay, a secondary destination and a passive eavesdropper. The outage probability is determined taking both signal to noise ratio (SNR) and transmission rate as the random variable. The secrecy outage performance is plotted against 'interference to noise power ratio in dB' under maximal ratio combining (MRC) scheme considering both power control and no power control. Similar analysis can be found for single hop links under multiple antenna systems in [2]. In cooperative cognitive radio, the concept of two-hop wireless links has also been introduced. For example, the outage performance of amplify-and-forward (AF) relay with multiple co-channel interference is considered in [3] under Nakagami-*m* fading channels. In [3], the product of fading factor and transmitted power is considered as a random variable following the pdf of Nakagami-*m*. The analytical expressions of outage probability have been derived using both pdf and cdf of Nakagami-*m* along with the end to end SNR. Finally, the outage probability is plotted against the threshold SNR taking the number of interferes in the first and second hops into consideration. It has been shown in [3] that the outage probability increases with increase in both threshold SNR and the number of interferers of any hop, both analytically and through simulations. In co-operative cognitive radio, more energy consumption is unavoidable to ensure accuracy of detection. Hence a tradeoff is desirable between these two parameters. In [4], a new scheme known as the energy-aware optimal relay selection (EAORS) is used, where the number of active relay station is optimized. The authors consider the weighted normalized sum of energy corresponding to transmission and sensing as the objective function. Finally, the variation of weighted energy consumption is plotted against the number of SU relays taking transmission distance, SNR and modulation schemes as parameters.

Another new cooperative CRN is proposed in [5] where primary transmitter (PT) sends message to primary receiver (PR) directly but the secondary transmitter (ST) and receiver (SR) listen to the message. When PR receives the message successfully and send the ACK then ST and SR listen to it and remove the message from buffer. If PR fails to decode the data packet, then SR will decode the data packet and redirect it to ST. If ST is able to decode it then the secondary system (SS)

will work in co-operative mode. The long-term average throughputs (LATs) and outage probability is considered as the key parameters in determining the performance of the network. The combination of above four nodes is found under a different protocol in [6]; where the authors derived closed-form solution of bit error rate (BER) and outage probability for PS and SS. A similar co-operative model has also been considered in [7] under Rayleigh fading channels where two SUs are used in co-operative mode to send the message in secondary destination. Another new concept of co-operative CRN is found in [8]. Several SUs closely located form a mesh connection in a CRN called coalition, used in co-operation mode to enhance throughput of the system. A five-stage algorithm is applied to reduce average probability of false alarm and to increase average throughput. Both performance metrics have been plotted against the number of users and a comparison is made with non-cooperative CRN.

However, all these works deal with the performance of CRN or dual hop wireless links under either SNR, signal to interference (SIR) or signal to interference and noise ratio (SINR) model. To the best of our knowledge, none of these works presents the relative performance of two models together.

In this paper, we analyze the performance of two models in terms of outage probability and symbol error rate (SER) under small scale fading and combining schemes. The impact of the number of antenna elements on combining schemes and two different fading environments (e.g., Chi-square and Nakagami-*m*) have been investigated in terms of the above performance metrics. Extensive simulation results demonstrate that the single user based model shows the best performance as compared with the multi-user model. Changing the modulation scheme, combining scheme and number of antenna elements in the multi-user model can provide comparable results under the single user model.

The rest of this paper is organized as follows. Section II deals with system model of CRN under single and Multi-SU Scenario with the concept of dual hop wireless link with full mathematical and statistical explanations. Section III provides results based on the analyses of section II, and finally, section IV concludes the paper.

II. System Model

We consider a dual-hop LTE CRN network consisting of a macro eNB, home eNBs (i.e., femto BS), PUs and SUs, as shown in Fig. 1. The performance of a traditional CRN depends on the utilization of traffic channels by SUs under the constraints of misdetection and false alarm. In LTE networks, however, the situation is more complicated where the macro and femto cells coexist. Here a SU can communicate with femto BS directly but depends on Pus for communicating with the macro BS, as shown inFig.1. Thus the PUs work as repeaters and the concept of dual-hop wireless links is used to measure the performance of the network. In this paper, we consider three models of the dual hop wireless links to observe the relative performance of LTE networks where PUs help the SUs to reach thee NB of the macro cell. We begin the analysis with the most simplified scenario, in which a single SU communicates with the eNB with the aid of a PU. Then we consider a multi-SU scenario where multiple SUs communicate with the eNB simultaneously. In the multi-SU scenario, each SU experiences interference from other users' transmission. Hence we analyze the performance of SNR based link model for the single SU scenario, and both SINR as well as SIR based models for the multi-SU scenario in the dual-hop LTE CRN links.

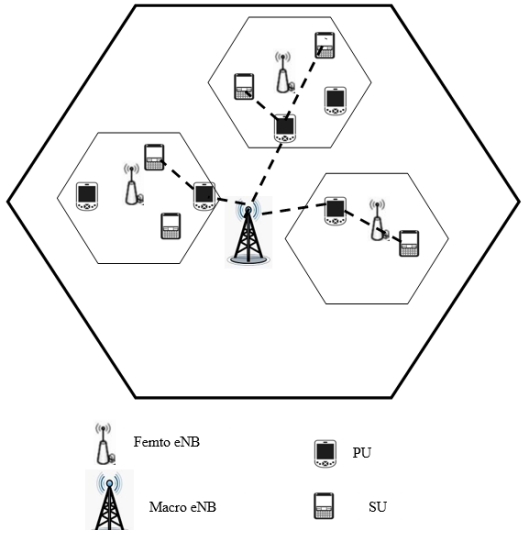


Fig.1. Multi-user model of cognitive radio network

A. SNR Method for the Single SU Scenario

The link between the SU and the macro eNB is established through a PU which acts as a relay. The PU performs an amplify-and-forward (AF) relaying operation for simplicity reasons. Therefore, 2-hop wireless link of AF relaying is applicable in this case. Let us assume that the SU transmits the symbol S to PU; the PU amplifies it by a factor β, and then sends it to eNB. The other parameters involved in this model are defined as:

$S \rightarrow$ Symbol transmitted by the SU,
$P \rightarrow$ Transmit power
$\gamma_1 \rightarrow$ SNR of the link between the SU and the PU,
$d_1 \rightarrow$ Distance between the SU and the PU,
$g_1 \rightarrow$ Path gain between the SU and the PU,
$\gamma_2 \rightarrow$ SNR of the link between the PU and the eNB,
$d_2 \rightarrow$ Distance between the PU and the eNB,
$g_2 \rightarrow$ Path gain between the PU and the eNB.

The signal received at the PU when S is transmitted by the SU can be expressed as [9-10],

$$r_{PU} = \sqrt{Pd_1^{-\varepsilon}}\, g_1 S + n_1; \quad (1)$$

Where P is the SU's transmit power, n_1 is the complex additive white Gaussian noise (AWGN) at the PU, and ε is the path-loss exponent. The PU performs linear operation on the received signal r_{PU} and transmits the amplified signal to the eNB. Thus the received signal at the eNB is given by

$$r_{eN\beta} = \sqrt{Pd_2^{-\varepsilon}}\, g_2 \beta r_{pu} + n_2$$

$$= \sqrt{Pd_2^{-\varepsilon}}\, g_2\beta\left\{\sqrt{Pd_1^{-\varepsilon}}\, g_1 S + n_1\right\} + n_2$$

$$= P\sqrt{(d_1 d_2)^{-\varepsilon}}\, g_1 g_2 S\beta + \sqrt{Pd_2^{-\varepsilon}}\, g_2\beta n_1 + n_2 \quad (2)$$

;where $n_2$is the complex AWGN at the eNB.

Now, the SNR of the link between the SU and the PU is given by

$$\gamma_1 = \frac{\left(\sqrt{Pd_1^{-\varepsilon}}\, g_1\right)^2}{n_1^2} = \frac{Pd_1^{-\varepsilon} g_1^2}{n_1^{/}}; where, n_1^{/} = n_1^2,$$

and the SNR of the link between the PU and the eNB is given by

$$\gamma_2 = \frac{\left(P\sqrt{(d_1 d_2)^{-\varepsilon}}\, g_1 g_2 \beta\right)^2}{n_2^2}$$

$$= \frac{P(d_1 d_2)^{-\varepsilon} g_1^2 g_2^2 \beta^2}{n_2^{/}}; \text{where}, n_2^{/} = n_2^2. \quad (3)$$

Thus the effective SNR experienced by thee NB in the second phase can be expressed as

$$\gamma_{eq} = \frac{\left\{P\sqrt{(d_1 d_2)^{-\varepsilon}}\,(g_1 g_2)\beta\right\}^2}{\left(\sqrt{Pd_2^{-\varepsilon}}\, g_2\beta n_1\right)^2 + n_2^2}$$

$$= \frac{P^2 (d_1 d_2)^{-\varepsilon} g_1^2 g_2^2 \beta^2}{Pd_2^{-\varepsilon} g_2^2 n_1^2 \beta^2 + n_2^2} \quad (4)$$

Rearranging the terms in (4), we obtain

$$\therefore \frac{1}{\gamma_{eq}} = \frac{Pd_2^{-\varepsilon} g_2^2 \beta^2 n_1^{/} + n_2^{/}}{P^2 (d_1 d_2)^{-\varepsilon} g_1^2 g_2^2 \beta^2}$$

$$= \frac{n_1^{/}}{Pd_1^{-\varepsilon} g_1^2} + \frac{n_2^{/}}{P^2 (d_1 d_2)^{-\varepsilon} (g_1 g_2 \beta)^2}$$

$$= \frac{1}{\gamma_1} + \frac{1}{\gamma_2}. \quad (5)$$

Similarly,the effective SNR in the downlink i.e.,eNB to SU link, through the PU as a repeater can be modeled.

B. SINR Method for the Multi-SU Scenario

In this subsection, we analyze the more complicated scenario of multiple SUs communicating with the eNB simultaneously. The actual communication is done in two phases: (i) in the multiple access (MAC) phase, both SU and eNB transmit their symbolsS_1 and S_2, respectively, to the PU, (ii) in the broadcast (BC) phase, the PU amplifies the combined received signal by a factor of β and transmits the amplified signal to the eNB and the SU discussed in [11]. In this case, S_2 creates interference to the SU, while S_1 creates interference to the eNB. Hence the SINR based link model is more appropriate in the context of practical network compared to the SNR model presented in the previous subsection.

In the MAC phase, the received signal at the PU is given by

$$y_{pu} = \sqrt{Pd_1^{-\varepsilon}}\, g_1 S_1 + \sqrt{Pd_2^{-\varepsilon}}\, g_2 S_2 + n_1, \quad (6)$$

Where $n_1$is the complex AWGN at the PU with mean zero and variance σ_p^2. The SNR between SU and PU of the transmitted signal S_1 is given by

$$\gamma_1 = \frac{\left(\sqrt{Pd_1^{-\varepsilon}}\, g_1\right)^2}{n_1^2} = \frac{Pd_1^{-\varepsilon} g_1^2}{n_1^{/}}, n_1' = \sigma_p^2.$$

In the BC phase, the PU broadcasts the amplified signal to both the SU and the eNB. Thus the received signal at the eNB is defined as

$$y_{eN\beta} = \sqrt{Pd_2^{-\varepsilon}}\, g_2 y_{pu}.\beta + n_2$$

$$= \sqrt{Pd_2^{-\varepsilon}}\, g_2\beta\left\{\sqrt{Pd_1^{-\varepsilon}}\, g_1 S_1 + \sqrt{Pd_2^{-\varepsilon}}\, g_2 S_2 + n_1\right\} + n_2$$

$$= P\sqrt{(d_1 d_2)^{-\varepsilon}}\, g_1 g_2 \beta S_1 + Pd_2^{-\varepsilon} g_2^2 \beta S_2 + \sqrt{Pd_2^{-\varepsilon}}\, g_2\beta n_1 + n_2 \quad (7)$$

Wheren_2 is the zero-mean complex AWGNat the eNB with variance σ_e^2. Therefore, the SINR of the transmitted signal S_1 at thee NB is given by

$$\gamma_{eNB} = \frac{\left(P\sqrt{(d_1 d_2)^{-\varepsilon}} g_1 g_2 \beta\right)^2}{(Pd_2^{-\varepsilon} g_2^2 \beta)^2 + \left(\sqrt{d_2^{-\varepsilon}} g_2 \beta \sigma_p\right)^2 + n_1'}$$

$$= \frac{P^2 (d_1 d_2)^{-\varepsilon} g_1^2 g_2^2 \beta^2}{P^2 d_2^{-2\varepsilon} g_2^4 \beta^2 + Pd_2^{-\varepsilon} g_2^2 \beta^2 n_1^{/} + n_2^{/}}, \quad (8)$$

Where$n_2' = \sigma_e^2$. Rearranging the terms in (8), we have

$$\therefore \frac{1}{\gamma_{eN\beta}} = \frac{P^2 d_2^{-2\varepsilon} g_2^4 \beta^2 + Pd_2^{-\varepsilon} g_2^2 \beta^2 n_1^{/} + n_2^{/}}{P^2 (d_1 d_2)^{-\varepsilon} g_1^2 g_2^2 \beta^2}$$

$$= \frac{d_2^{-\varepsilon} g_2^2}{d_1^{-\varepsilon} g_1^2} + \frac{n_1^{/}}{Pd_1^{-\varepsilon} g_1^2} + \frac{n_2^{/}}{P^2 (d_1 d_2)^{-\varepsilon} (g_1 g_2 \beta)^2}$$

$$= \left\{\left(\frac{d_2}{d_1}\right)^{-\varepsilon}\left(\frac{g_2}{g_1}\right)^2\right\} + \frac{1}{\gamma_1} + \frac{1}{\gamma_2}$$

$$= \frac{1}{\gamma_I} + \left(\frac{1}{\gamma_1} + \frac{1}{\gamma_2}\right)$$

$$= \frac{1}{\gamma_I} + \frac{1}{\gamma_{eq}}, \quad (9)$$

,where $\gamma_I \triangleq \left(\frac{d_2}{d_1}\right)^{\varepsilon}\left(\frac{g_2}{g_1}\right)^{-2}$. It can be clearly seen from (9) that the effective end-to-end link experiences an additional drop in SNR due to the amplification of the interference at the relay node (i.e., PU).Note that we consider a flat fading channel environment in this paper so that the combined equivalent SINR can be considered as a random variable. Let us now assume that for successful decoding at the receiving end, the required SINR threshold is $\gamma_{t\square}$. Thus we have

$$\gamma_{eq} = \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2 + c} \geq \gamma_{th}$$

$$\Rightarrow \gamma_1 \geq \frac{\gamma_{th}(\gamma_2 + c)}{\gamma_2 - \gamma_{th}}$$

,

Where c is a random variable that represents the combined SINR. Accordingly, the probability density function (PDF) of γ_{eq} is given by

$$f_{\Gamma_{eq}}(\gamma_{th}) = \int_{\gamma_{th}}^{\infty} \frac{d}{d\gamma_{th}} \left\{ \frac{\gamma_{th}(\gamma_2 + c)}{\gamma_2 - \gamma_{th}} \right\} f_{\Gamma_1}\left(\frac{\gamma_{th}(\gamma_2 + c)}{\gamma_2 - \gamma_{th}} \right) f_{\Gamma_2}(\gamma_2)\, d\gamma_2$$
$$= \int_{\gamma_{th}}^{\infty} \frac{\gamma_{th}(\gamma_2 + c)}{(\gamma_2 - \gamma_{th})^2} f_{\Gamma_1}\left(\frac{\gamma_{th}(\gamma_2 + c)}{\gamma_2 - \gamma_{th}} \right) f_{\Gamma_2}(\gamma_2)\, d\gamma_2 \quad (10)$$

and the cumulative distribution function (CDF) of γ_{eq} can be derived as

$$F_{\Gamma_{eq}}(\gamma_{th}) = 1 - \int_{\gamma_{th}}^{\infty} \left\{ 1 - F_{\Gamma_1}\left(\frac{\gamma_{th}(\gamma_2 + c)}{\gamma_2 - \gamma_{th}} \right) \right\} f_{\Gamma_2}(\gamma_2)\, d\gamma_2 . \quad (11)$$

Now the SINR outage probability of the link is defined as

$$P_{out}(\gamma_{th}) = 1 - F_{\Gamma_{eq}}(\gamma_{th}) \quad (12)$$

In this paper, the combined SINR is considered as a random variable that follows Rayleigh and Chi-square distribution with v degree of freedom.

C. Combining Scheme

In CRN, several MSs can be used as repeaters to achieve diversity gain. While diversity is a powerful technique to overcome the effects of flat fading by combining multiple independent fading paths, diversity typically entails some penalty in terms of rate, bandwidth, complexity, or size. In practice, both selection combining (SC) and MRC significantly reduce the impact of fading. However, both SC and MRC offer different levels of performance against complexity. The concept of virtual MIMO with combining schemes can be used at the receiving end for multi-user systems. In this section, we consider the SC and MRC schemes under the virtual MIMO setup.

Under SC, we aim at deriving the cdf,
$\gamma_{sc} = \max(\gamma_1, \gamma_2, \gamma_3, \cdots, \gamma_N)$,
where $\gamma_1, \gamma_2, \gamma_3, \ldots, \gamma_N$ are the SNR provided by repeater 1, 2, 3, …, N, respectively.
The probability,

$$P(\gamma_i \le \tau) = \int_{-\infty}^{\tau} f_{\Gamma_i}(\gamma_i) d\gamma_i = \Psi_i(\tau). \quad (13)$$

Therefore,

$$P(\gamma_1, \gamma_2, \gamma_3, \ldots, \gamma_N \le \tau) = \prod_{i=1}^{N} P(\gamma_i \le \tau) = \prod_{i=1}^{N} \Psi_i(\tau).$$

Thus the cdf of γ_{sc} is defined as

$$F_{\Gamma_{sc}}(\gamma_{sc}) = \prod_{i=1}^{N} \Psi_i(\gamma_{sc}). \quad (14)$$

Under the MRC scheme, we can use upper bound of SNR of dual hop link as considered in [12-13]. Accordingly, the received SNR at the i-th node is defined as

$$\gamma_i = \frac{\gamma_{SR_i}\gamma_{RD_i}}{\gamma_{SR_i} + \gamma_{RD_i}} \le \min(\gamma_{SR_i}, \gamma_{RD_i}) = \gamma_{u_i} \quad (15)$$

where γ_{u_i} is the upper band of the end-to-end SNR of i-th relayed link.
Therefore, the combined SNR of the virtual MIMO system $\gamma_{mrc} = \sum_{k=1}^{N} \gamma_{uk}$ is the random variable with Chi-square pdf with N degree of freedom such that

$$f_{\Gamma_{mrc}}(\gamma_{mrc}) = \frac{1}{(N-1)!} \frac{\gamma_{mrc}^{N-1}}{\gamma_{av}^{N}} \exp\left(-\frac{\gamma_{mrc}}{\gamma_{av}} \right), \quad (16)$$

;where γ_{av} is the average value of γ_{mrc}.

III. Results

In this section, we evaluate the performance of the proposed analytical results through numerical simulations.

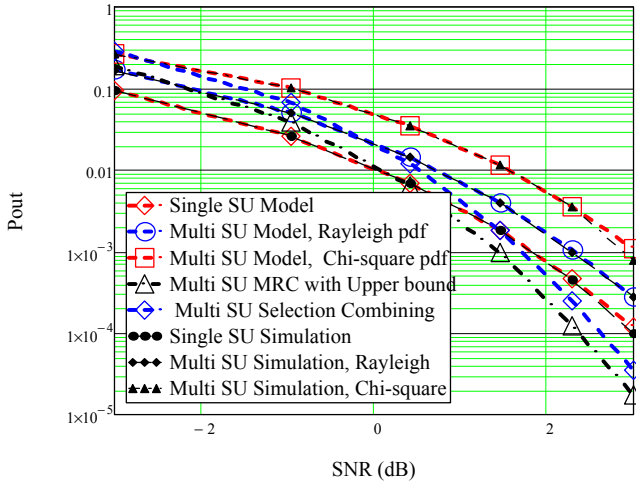
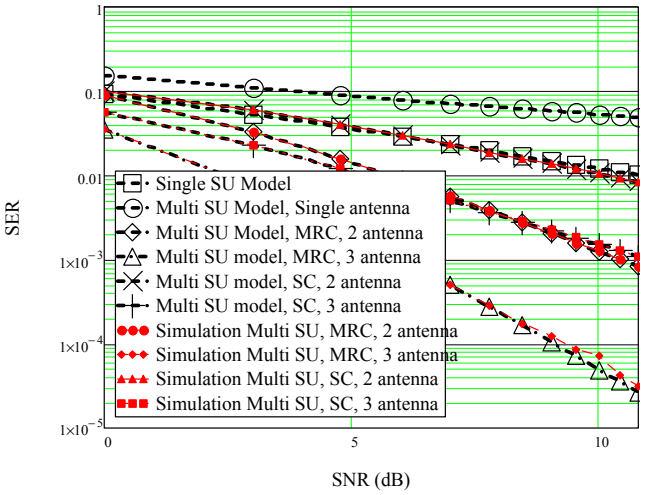


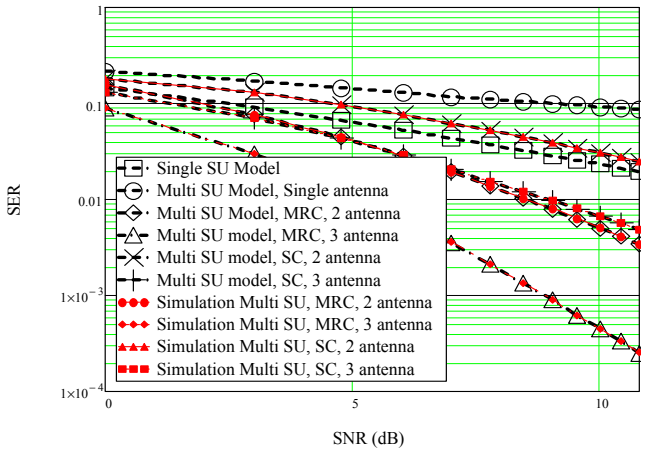
Fig.2. Comparison of outage probability of analytical and simulation work.

Here, we consider two antenna elements at the receiving end for both MRC and SC schemes under Rayleigh fading environment. Fig.2 shows the variation of outage probability against instantaneous SNR of the received signals for five different cases. Results in Fig. 2 reveal that the single user model yields the better performance at low SNR since the model ignores the interferences. However, as the SNR increases beyond 0.1dB, the MRC under multi-user case (upper bound) becomes comparable with this model. On the other hand, the performance of the SC scheme seems a little bit worse than that of the MRC scheme, but better than the other two fading models. The performance of the Chi-square pdf case depends on the degree of freedom v. For v = 2, the Rayleigh and Chi-square fading become comparable under multi-user model but for larger value of v the performance of the Chi-square fading environment deteriorates heavily. In this paper we use Monte

Carlo simulation generating 10^6 random variable of Rayleigh and Chi-Square distribution to evaluate outage probability. The analytical and simulation results are found very closed.



(a) QPSK modulation



(b) 16-QAM modulation

Fig.3. Comparison of SER for different modulation schemes.

Finally, Fig.3(a) and 3(b) show the SER performance of the models against SNR in dB for QPSK and 16-QAM modulation schemes, respectively. Similar to the previous figures of this section, the performance of the single user case is better than the multi-user case, provided SC and MRC schemes (multi-user case model) show better performance than single user case when the number of antenna elements $N = 3$. If the relative performance of SC and MRC are compared for the same number of antenna elements, the MRC performs far better than the SC under the multi-user model as visualized in Fig.3.

IV. Conclusions

Recent literature deals with performance of CRN or dual hop wireless links under either SNR or SINR model. This paper presents the relative performance of both models in a simpler manner. The SINR model or multi user case represents the real-life scenario of wireless link between the two models considered in the paper, and provides desired results using adaptive modulation schemes and combining schemes of LTE. Here we consider Rayleigh and Chi-square pdf random variables in both hops of a dual-hop link. The dissimilar fading models, for example, the first hop under Rayleigh fading and the second hop under Nakagami-*m* fading, can also be used in future to get the scenario of dense urban areas.

References

[1] Pratik Chakraborty and Shankar Prakriya, 'Secrecy Outage Performance of a Cooperative Cognitive Relay Network,' IEEE Communications Letters, vol.21, no.2, pp.326-329, Feb'2017.

[2] Xin Jiang, CaijunZhong, Xiaoming Chen, Trung Q. Duong, Theodoros A. Tsiftsis and Zhaoyang Zhang, 'Secrecy Performance of Wirelessly Powered Wiretap Channels,' IEEE Transactions on Communications, vol. 64, no. 9, pp.3858-3871, September 2016.

[3] Daniel Benevides da Costa and Michel DaoudYacoub, 'Outage Performance of Two Hop AF Relaying Systems with Co-Channel Interferers over Nakagami-*m* Fading,' IEEE Communications Letters, vol. 15, no. 9, pp.980-982, September 2011.

[4] XiaorongXu, JianrongBao, Haiyan Cao, Yu-Dong Yao, and Sanqing Hu, 'Energy-Efficiency-Based Optimal Relay Selection Scheme With a BER Constraint in Cooperative Cognitive Radio Networks,' IEEE Transactions on Vehicular Technology, vol. 65, no. 1, pp.191-203, January 2016.

[5] Eun-Yeong Park, Moon-Gun Song, Wonseok Choi, and Gi-Hong Im, 'Maximization of Long-Term Average Throughput for Cooperative Secondary System With HARQ-Based Primary System in Cognitive Radio Network,' IEEE Communications Letters, vol. 20, no. 2, pp.356-359, February 2016.

[6] Naveen Gupta and Vivek Ashok Bohara, 'An Adaptive Subcarrier Sharing Scheme for OFDM-Based Cooperative Cognitive Radios,' IEEE Transactions on Cognitive Communications and Networking, vol. 2, no. 4, pp.370-380, December 2016.

[7] RaikelBordón, Samuel Montejo Sánchez, Samuel BaraldiMafra, Richard Demo Souza, JoãoLuizRebelatto, and Evelio Martin Garcia Fernandez, 'Energy Efficient Power Allocation Schemes for a Two-User Network-Coded Cooperative Cognitive Radio Network,' IEEE Transactions on Signal Processing, vol.64, no.7, pp.1654-1667, April 2016.

[8] Mohammad Tahir, M.H. Habaebi, M. Rafiqul Islam, 'Coalition Formation for Cooperative Spectrum Sharing in Cognitive Radio Wireless Networks using Gale Shapley Algorithm,' 2015 IEEE 6th Control and System Graduate Research Colloquium, pp.43-48, Aug. 10 - 11, UiTM, Shah Alam, Malaysia.

[9] Baofeng J, Bingbing Xing, Kang Song, Chunguo Li, Hong Wen, and Luxi Yang, 'Performance Analysis of Multihop Relaying Caching for Internet of Things under Nakagami Channels,' Wireless Communications and Mobile Computing, pp.1-9, Volume 2018, Article ID 2437361, 2018

[10] Abu Sayed Md. Mostafizur Rahaman, Jesmin Akhter, Md. Rafsan Jani, Md. Imdadul Islam, 'Determination of Array Gain of Single Hop to Achieve the Performance of a 2-Hop Wireless Link,' Journal of Computer and Communications, 6, pp.84-98, 2018

[11] Neng Ye ,Hangcheng Han, Lu Zhao and Ai-hua Wang, 'Uplink Nonorthogonal Multiple Access Technologies Toward 5G: A Survey,' Wireless Communications and Mobile Computing, Volume 2018,pp. 1-26, Article ID 6187580, 2018

[12] Md. Zahurul I. Sarkar; Tharmalingam Ratnarajah, 'Bounds on the secrecy capacity with diversity combining techniques,'2012 IEEE Wireless Communications and Networking Conference (WCNC) Pages: 2847 - 2851, Year: 2012.

[13] Diana Cristina González, Daniel Benevides da Costa and José Cândido Silveira Santos Filho, 'Distributed Suboptimal Schemes for TAS/SC and TAS/LS in Fixed-Gain AF Relaying Systems,' IEEE Transactions on Wireless Communications, vol. 13, no. 11, pp.6041-6053, November 2014.

Impact of SUs under FC in Co-operative Cognitive Radio Network in PUEA Environment

Fahmida Hossain
Department of Computer Science and Engineering
Jahangirnagar University
Dhaka-1342, Bangladesh
fahmida.4323@gmail.com

AnzumanYeasmin Snigdha
Department of Computer Science and Engineering
Jahangirnagar University
Dhaka-1342, Bang;adesh
anzuman.yasmin@gmail.com

Md. Imdadul Islam
Department of Computer Science and Engineering
Jahangirnagar University
Dhaka-1342, Bang;adesh
imdad@juniv.edu

***Abstract*— In this paper we determine the performance of a co-operative Cognitive Radio Network (CRN) under Additive White Gaussian Noise (AWGN). The impact of the number of received samples on probability of detection (P_D) and false alarm (P_{FA}) are shown graphically to determine the optimum number of samples for perfect detection. Next we show the impact of number of SUs under FC on the performance of co-operative CRN and determine the optimum value of threshold signal to noise ratio (SNR) at FC to combat the effect of PUEA. Analytical results are verified by simulation and found confidence level above 99%.**

Keywords—false alarm, misdetection, NP theorem, threshold SNR, FC

I. INTRODUCTION

This In MAN (Metropolitan Area Network) or WAN (Wide Area Network) limited number of channels are allocated to a huge number of users. The channel allocation is done using MAC sub-layer. Sometimes all channels are found busy but most of the time few channels are found idle or unused. In conventional mobile cellular network, there is no explanation about these unused channel. CRN gives the real explanation of these unused channel. Several works are done on CRN and few of them are mentioned in this paper. In [1] the analytical and simulation results of probability of detection and false alarm of a co-operative cognitive radio network are compared under both AWGN and Rayleigh fading environment. The research work of [2] extends the performance of a CRN using the combination of membership functions of SNR, SIR (signal to interference ratio) and SINR (signal to noise plus interference ratio) which improves the detection hypothesis at the expense of process time and complexity of Fuzzy model. Authors of [3], revealed the threshold gain of the adaptive beam former along the line of sight (LOS) between the transmitter and the receiver. The paper [4], introduces a new CSS scheme in the presence of a malicious PUEA based on multi-level hypothesis testing (MLHT),generalizing from binary hypothesis testing to MLHT, they partition the decision space to four decision options and apply minimum Bayes cost criteria to determine the channel status. The Simulation results are provided to indicate the performance improvement of the proposed MLHT method against PUEA, compared with the conventional method. The research work of [5] states the issue of delay caused by a band-selection decision process that directly affects the security and performance. The model cluster-based distributed cooperative spectrum sensing is proposed here that cluster heads (CHs) exchange control information with other CHs and ordinary nodes, significantly reduced the delay, sensing, convergence, routing, in band-selection process. The simulated results show the improved performance of cognitive radio networks in terms of delay, packet loss ratio and bandwidth usage as compared to cluster-based cooperative spectrum sensing model. The opportunity for primary user emulation attacker is minimized as the overall delay is reduced. According to the concept of [6], a proposed algorithm is projected to distinguish primary user signal from others, through an energy-efficient localization mechanism and channel parameter variance, which is simulated in MATLAB environment and it is found to be an effective mechanism to authenticate primary user and mitigate the primary user emulation attack. As explained in [7], a detection method according to the characteristics of wireless channel. Using the multipath delay characteristics of the fading channels, PUEA can be detected by the secondary users (SUs). Simulation results illustrate that the proposed PUEA detection method achieves high detection probability in terms of low false positive probability. Authors of [8] stated an adaptive and realistic PUE attack detection technique is proposed in this paper inspired by the physical features widely used as the fingerprint of variant electronic devices. It leverages the PU transmission features that attackers are not able to mimic. In this work, the transmission power is selected as one of the hard-to-mimic features due to the intrinsic discrepancy between PUs and attackers, while considering constraints in real implementations. Their experimental results verified the effectiveness and correctness of the proposed mechanism. Authors in [9] propose a PUEA detection method that exploits the radio channel information to detect the PUEA in the cognitive radio network. In the proposed method, the uniqueness of the channel impulse response (CIR) between the secondary user (SU) and the signal source is used to determine whether the received signal is transmitted by the primary user (PU) or the primary user emulator (PUE). Numerical results show that the proposed PUEA detection method performs well although the difference in channel characteristics between the PU and PUE is small. The research work of [10] provides a survey for resource allocation in underlay CRNs as for conventional wireless networks, the main target of resource allocation (RA) is to efficiently utilize the available resources that lead to spectrum under-utilization, and in this regard cognitive radio networks (CRNs) have received great attention due to their potential to improve the spectrum utilization. Finally [11] introduces cognitive radio network in order to dynamic the spectrum therefore to increase the spectrum efficiency, decreasing costs and increasing number of wireless users, cognitive radio network solved bad utilizations and in advance it improved more characteristics for the spectrum. For large value of degree of freedom, the Chi-square pdf becomes Gaussian pdf, which makes the analysis of CRN easier. Above concept motivates us to work in this paper. In this paper, our main contribution is to estimate the threshold SNR at FC to make the probability of false alarm and misdetection below 4% at the same time to ensure the probability of detection above 96%. Finally the probability of PUEA is also

dropped below 4% under fine tuning of threshold SNR both analytically and simulation.

The rest of paper is organized like: section II provides basic theory and statistical analysis of detection of presence or absence of PU, section III gives analytical and simulation results based on analysis of section II and finally, section IV concludes entire analysis.

II. SYSTEM MODEL

A. Basic of two Hypothesis model

The received signal of the SU on a traffic channel, $x(n)$, is expressed under two hypothesis model like,

$$x(n) = \begin{cases} w(n) & ; H_0 \\ s(n) + w(n) & ; H_1 \end{cases} \quad (1)$$

where $w(n)$ is additive white Gaussian noise (AWGN), H_0 is the hypothesis of the absence of PU signal and H_1 denotes the hypothesis of the presence of the PU signal. To determine the detection performance we first note that the test statistics $T(x)$ as,

$T(x) = \left(\frac{1}{N}\right)\sum_{n=0}^{N-1} x[n]$; is Gaussian under each hypothesis. The means of the $T(x)$ is,

$$E(T(x); H_0) = E\left(\frac{1}{N}\sum_{n=0}^{N-1} w[n]\right)=0 \quad (2)$$

Similarly, $E(T(x);H_1)= A$ for hypothesis H_1

The variance of test statistics,

$$var(T(x); H_0) = var\left(\frac{1}{N}\sum_{n=0}^{N-1} w[n] \; ; H_0\right)$$

$$= \frac{\sigma^2}{N}$$

$$= var(T(x); H_1) \quad (3)$$

Therefore, $$T(x) \sim \begin{cases} N\left(0, \frac{\sigma^2}{N}\right) \text{under} H_0 \\ N\left(A, \frac{\sigma^2}{N}\right) \text{under} H_1 \end{cases} \quad (4)$$

We have the probability of false alarm as,

$$P_F = P_r\{T(x) > \gamma \; ; H_0\} = Q\left(\frac{\gamma}{\sqrt{\sigma^2/N}}\right) \quad (5)$$

Next the probability of detection becomes,

$$P_D = P_r\{T(X) > \gamma \; ; H_1\} = \frac{1}{\sqrt{2\pi}} Q\left(\frac{\gamma - A}{\sqrt{\sigma^2/_N}}\right) \quad (6)$$

The relation between P_F and P_D is expressed as,

$$P_D = Q\left(Q^{-1}(P_{FA}) - \sqrt{\frac{NA^2}{\sigma^2}}\right) \quad (7)$$

B. Co-operative spectrum sensing

To improve the detection accuracy the received signal, several SUs are combined at FC to take the decision about presence of PU like fig.1. Fusion rule may be AND logic, OR logic, MRC scheme etc. If N SUs send the signal to FC; where each SU sends M samples then the number of samples received at FC is MN. In energy detection, the detection random variable becomes, , follows Chi-square pdf. When

$$Y = \sum_{k=1}^{MN} |x(k)|^2$$

MN is large enough then Chi-square pdf approximated by Gaussian pdf. If each sample $x(n)$ follows $N(0, 1)$ then mean and variance of Y is found in [4].

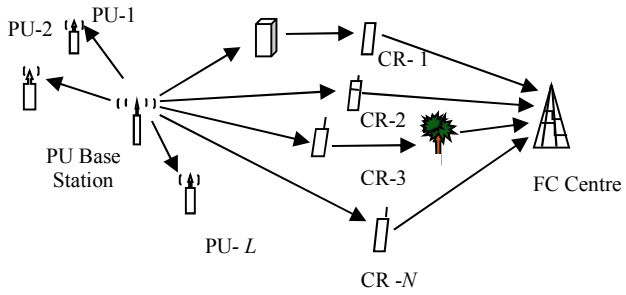


Fig. 1. Cooperative spectrum sensing

C. Co-operative spectrum sensing

PUEA presents a serious threat on the performance of CRN; which actually imitate the PUs to confuse the SUs. The real scenario of PUEA is shown in fig.2 can be expressed with three hypothesis model like,

$$Y \sim \begin{cases} H_0 : only\ noise \\ H_1 : PU + noise \\ H_2 : PUEA + noise \end{cases} \quad (8)$$

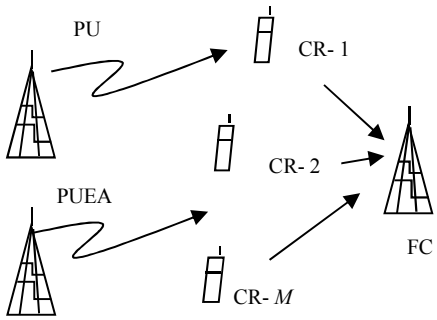


Fig. 2. Primary user emulation attack in cognitive radio networks

The PUE attackers or malicious user creates signals of similar power of PU, hence in energy detection technique it is difficult to recognize a PUE attacker. In feature-based detectors the attacker will emulate the corresponding feature of the PU but dynamic change of feature can be a remedial measure, which is beyond the scope of the paper.

D. Simulation algorithm

This subsection provides Monte Carlo simulation in determining P_f and P_D. The steps of algorithm are:

1. Input variables under H_0: $N\rightarrow$ number of users under FC, $M\rightarrow$ number of samples/symbol, $m\rightarrow$ number of random variable in simulation, $n\rightarrow$ length of random variable to take average, $T\rightarrow$ average of n random variable, $\mu_0 = MN\rightarrow$ the mean of T, $\sigma_0^2 = 2MN\rightarrow$ variance of T, $\gamma\rightarrow$ SNR, $\gamma_{max}\rightarrow$ maximum SNR.
2. Take an array of length m with initial value of each elements 0, as T = zeros(m,1);
3. Generate n random number with normal pdf of mean μ_0 and variance of σ^2 :
   x = sqrt (σ^2)*randn(n,1)+μ_0 and take mean of x.

4. Repeat step 3,m times and store the mean of x each time in an array T like:

for i = 1:m
{
$x = sqrt(\sigma^2)*randn(n,1) + \mu 0$
T (i) = mean (x);
}

5. Select the range of SNR, γ as γ = *0 : interval:* γ_{max};Find the number of elements of the array T greater than $\gamma(i)$ for each value of SNR,$\gamma(i)$ designated as:γ_{num}; The probability of false alarm under simulation is the ratio of length of γ_{num} and m.

for i =1: length (gamma)
{
γ_{num} = *find (T>gamma (i));*
PFA_sim(i) = length (γ_{num})*/m;*
}

6. Evaluate the analytical probability of false alarm, $PFA\_th = 0.5*\text{erfc}((\gamma-\mu_0)/(\text{sqrt}(2*\sigma_0^2/n)))$;
7. Under hypothesis H_1 the variables are:The mean SNR$\rightarrow \bar{\gamma}$, $\mu_1 = MN(\bar{\gamma}+1)$, $\sigma_1^2 = 2NM(\bar{\gamma}+1)^2$. Repeat step 1 to 5 under the variables of hypothesis H_1 and determine the probability of detection from step 5.
8. Evaluate the analytical probability of detection, $PD\_th = 0.5*\text{erfc}((\gamma-\mu_1)/(\text{sqrt}(2*\sigma_1^2/n)))$;
9. Under hypothesis H_2 the variables are: The mean SNR$\rightarrow \bar{\lambda}$, $\mu_2 = MN(\bar{\lambda}, +1)$, $\sigma_2^2 = 2NM(\bar{\lambda}, +1)^2$.
10. Repeat step 1 to 5 under the variables of hypothesis H_2 and determine the probability of detection from step 5.
11. Evaluate the analytical probability of detection, $PD\_th = 0.5*\text{erfc}((\gamma-\mu_2)/(\text{sqrt}(2*\sigma_{12}^2/n)))$;

III. RESULTS

First of all we plot the probability of false alarm against threshold SNR in dB (normalized with respect to μ_0) taking the number of samples as a parameter. The P_f decreases with increase in threshold SNR and attains to 0 at -1dB or above. The decrement of P_f occurs at higher value SNR for larger number of samples. Fig.3 shows the graphs for M = 10, 15, 20 and 25 samples. Similar analysis is shown in Fig.4 for the case of probability of misdetection (P_{md}). The increment of P_{md} is found at larger value of threshold SNR. Comparing fig.3 and 4 we found that taking threshold SNR/μ_0 within 0 to 5 dB we can ensure both P_f and P_{md} below 5% only for M>20. A tradeoff is necessary between proper detection (or spectrum sensing) of signal and the number of samples (related to process time).The results of fig.3 and 4 are for non-co-operative spectrum sensing case.

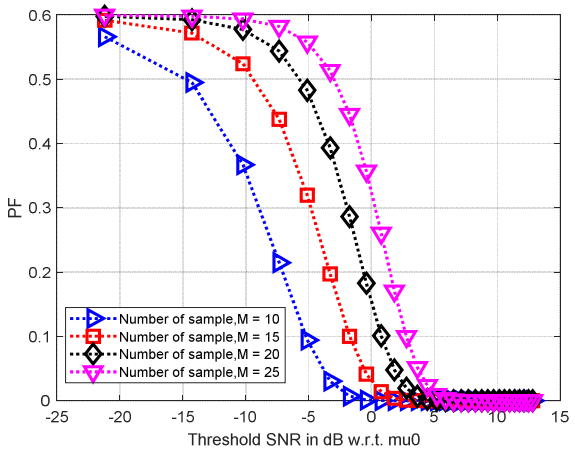


Fig. 3. Variation of probability of false alarm against the threshold SNR in dB

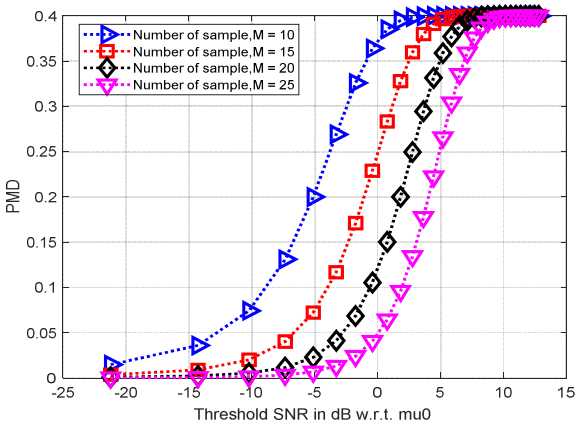
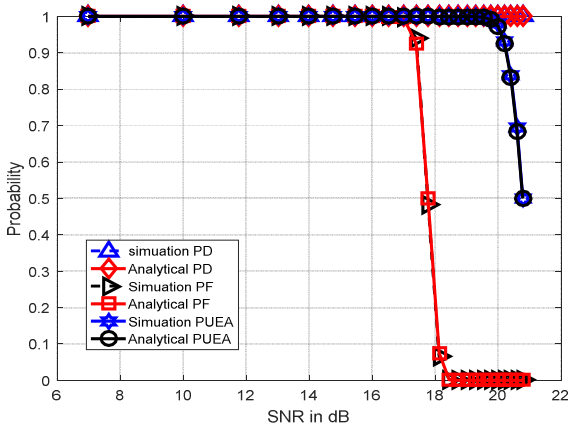
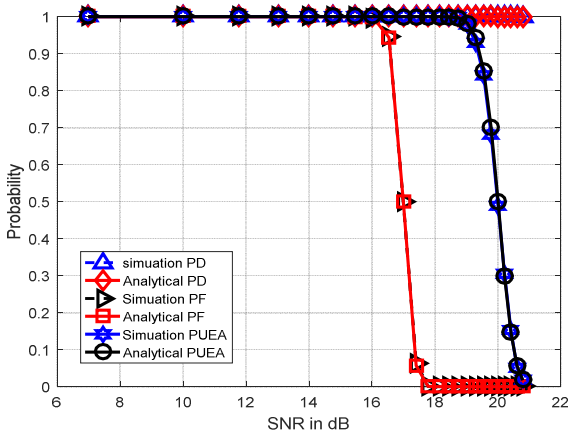


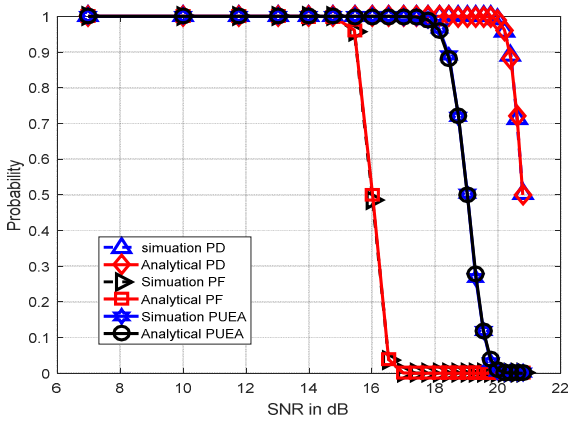
Fig. 4. Variation of probability of misdetection against the threshold SNR in dB



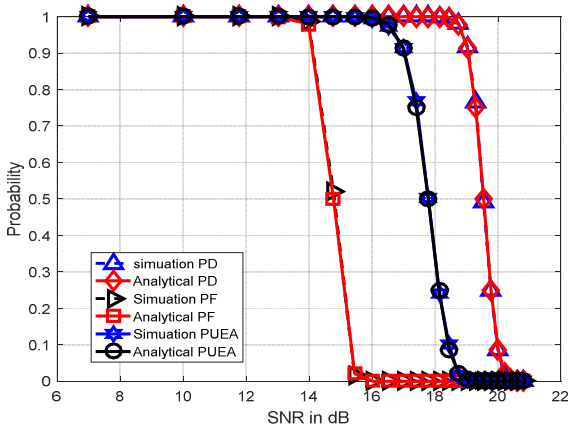
(a) $N = 6$



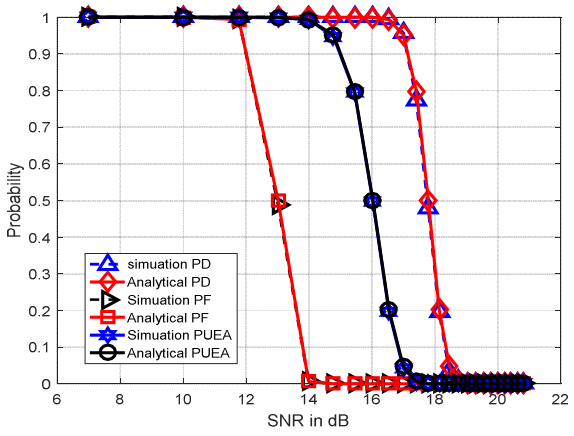
(b) $N = 5$



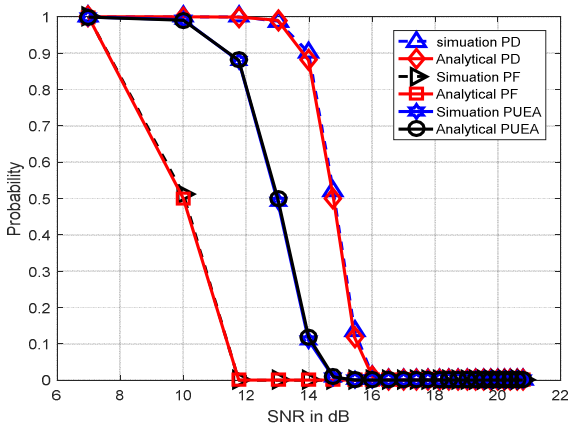
(c) $N = 4$



(d) $N = 3$



(e) $N = 2$



(f) $N = 1$

Fig. 5. Impact of users under a FC with PUEA

Next part of the result section deals with co-operative spectrum sensing case. A large number of co-operative SU under a FC improves the SNR at FC to take the correct decision. Fig. 5 (a) – (f), shows the comparison of analytical and simulation results under three hypothesis model: H_0 (no user), H_1 (only PU), H_2 (only P_{UEA}). Each figure provides three pairs of graph on P_D, P_F and P_{PUEA} under both analytical and simulation cases. Our simulation provides very closed results with that of analytical (the confidence level is above 98%). Because of power consumption, the presence of attacker will be small even transmitting power is also smaller compared to PU. Here we assume $P(H_3) = 0.2$ and $P_{PUEA} = P(H_3).P\{\gamma \geq \gamma_{th}\}$.The relation among the number of users N under FC, P_{PUEA} and threshold SNR is shown in table I. Both the simulation and analytical result reveals that P_D is very closed to 100% and P_F to 0% for $N > 4$.

TABLE I. IMPACT OF SUs ON CO-OPERATIVE CRN

Number of SU under FC (N)	Threshold SNR (γ_{th}) in dB	P_{FA} in %	P_D in %	P_{PUEA} in %
1	14.2	1.4	96.3	8
2	16.8	0.68	98.6	2
3	18.4	0	99.2	1.2
4	19.6	9	100	0.5
5	21.3	0	100	0
6	22.4	0	100	0

When both PUEA and PU exist simultaneously then $\gamma >> \gamma_{th}$ and SU can easily recognize the occupancy of the traffic channel, therefore we ignore the fourth hypothesis of [4].

IV. CONCLUSION

In this paper we evaluated performance of CRN based on probability of detection, probability of primary user emulation attack and that of false alarm varying threshold SNR. We determine the appropriate threshold SNR to maintain P_f below 4% and P_d above 96% under both non-cooperative and cooperative CRN. In future we will evaluate the performance of the same network under different types of small scale fading environment instead of awgn case. Even the concept of PUEA under four hypothesis model can be analyzed under Chi-square pdf.

REFERENCES

[1] Abul Kalam Azad, Jugal K. Das, Imdadul Islam, 'Signal Detection of Co-Operative Cognitive Radio Network under Neural Network,' Journal of Computer and Communications, vol. 6, no. 9, pp.60-72, September 2018

[2] Md Abul Kalam Azad, Sanjit Kumar Saha, Md. Imdadul Islam, Jugal Krishna Das, 'Detection of Primary User at Fusion Center of a CRN Using Fuzzy-Logic Rules,' International Journal of Computer Science and Information Security, vol. 16, no. 8, AUG 2018, pp.84-92

[3] Abu Sayed Md. Mostafizur Rahaman, Jesmin Akhter, Md. Rafsan Jani and Md. Imdadul Islam, 'Determination of Array Gain of Single Hop to Achieve the Performance of a 2-Hop Wireless Link,' Journal of Computer and Communications, vol. 6, no. 7, pp.84.98, July 2018

[4] Morteza Sharifi,Abbas Ali Sharifi,Mir Javad Musevi Niya,'Coo perative spectrum sensing in the presence of primary user emulation attack in cognitive radio network: multi-level hypotheses test approach,' Springer Science + Business Media New York 2016,Wireless Networks(The Journal of Mobile Communication, Computation and Information),vol. 24, Issue 1, pp.61–68, January 2018.

[5] Muhammad Ayzed Mirza,Mudassar Ahmad,Muhammad Asif H abib,Nasir Mahmood,C. M. Nadeem Faisal, Usman Ahmad, 'cluster-based distributed cooperative spectrum sensing model against primary user emulation (PUE) cyber attacks,'Springer Science, The Journal of Supercomputing, vol. 74, Issue 10, pp.5082–5098, 2018.

[6] Rukhsar Sultana, Muzzammil Hussain, 'Mitigating Primary User Emulation Attack in Cognitive Radio Network Using Localization and Variance Detection,' Springer Science, Proceedings of First International Conference on Smart System, Innovations and Computing, vol. 79, pp.433-444, January 2018.

[7] Yongcheng Li,Xiangrong Ma,Manxi Wang,Huifang Chen,Lei X ie, 'Detecting Primary User Emulation Attack Based on Multipath Delay in Cognitive Radio Network', Springer Science, Smart Innovations in Communication and Computational Sciences, vol. 669, pp.361-373,June 2018.

[8] Qiao-mu Jiang, Hui-fangChen, Lei Xie, Kuang Wang, 'On detecting primary user emulation attack using channel impulse response in the cognitive radio network', Springer Science, The Journal of Frontiers of Information Technology & Electronic Engineering, vol. 18, Issue 10, pp.1665–1676, October 2017.

[9] Qi Dong, Yu Chen, Xiaohua Li, Kai Zeng, Roger Zimmermann, 'An Adaptive Primary User Emulation Attack Detection Mechanism for Cognitive Radio Networks', Springer Science, International Conference on Security and Privacy in Communication Systems; SecureComm 2018: Security and Privacy in Communication Networks, pp.297-317, December 2018.

[10] M. E. Tanaband, W. Hamouda, 'Resource allocation for underlay cognitive radio networks: A survey,'*IEEECommun. Surveys Tuts.*, vol. 19, no. 2, pp.1249_1276, 2nd Quart., 2017.

[11] Omer I, Hamid K, Mohamed M, 'Overview on Cognitive Radio Network Review,' International Journal of Engineering Sciences Paradigms and Researches, vol. 33, no. 20, pp.7-13, 2016.

Sleep-based DBA Algorithm for Energy Efficiency by End Points Collaboration in EPON

Kamrun Nahar Runa, Monir Hossen, and Shuvashis Saha
Dept. of Electronics and Communication Engineering
Khulna University of Engineering & Technology, Khulna – 9203, Bangladesh
E-mail: kamrunnahar204@gmail.com, mhossen@ece.kuet.ac.bd, anikkuet@gmail.com

***Abstract*— Ethernet passive optical network (EPON) offers large bandwidth that requires less power consumption. The EPON systems require dynamic bandwidth allocation (DBA) algorithm to share a medium for different users with fairness and quality of services. In energy efficient DBA (EE-DBA) algorithm, the optical network units (ONUs) enter the power saving mode in lower traffic condition. However, the main limitation of the existing EE-DBA algorithm is that it requires to keep optical line terminal (OLT) in active state for all the time. That is why some energy is wasted in the OLT side. For mitigating the excess energy consumption by the OLT, we propose a sleep-based DBA (SB-DBA) algorithm for EPON system. In the SB-DBA scheme, both the OLT and ONUs enter into the power-saving mode for reducing overall energy consumption in the network. Simulations are used to evaluate the performance of the SB-DBA scheme. Comparative analysis of simulated results proves that the energy consumption of the proposed SB-DBA scheme is less than the existing EE-DBA scheme with similar end-to-end packet delay.**

Keywords— SB-DBA, EE-DBA, power-saving mode, energy efficiency.

I. INTRODUCTION

Ethernet passive optical network (EPON) is one of the promising access network architecture that can effectively reduce the energy consumption. The EPON system has three major components, namely optical line terminal (OLT) which forms the base station, optical network units (ONUs) perform as the terminating unit of the network, and the optical splitter which splits the downstream optical signal into multiple ONUs.

The energy savings for EPON is sub categorized into two classes; such as energy savings for OLT and energy savings for ONUs. In the energy savings for ONUs category, each ONU is entered into the power savings mode when there is no data packet to transmit. In the energy savings for OLT category, the OLT is entered into the power savings mode when there is no data packet to receive. Actually, power saving method in EPON depends on the availability of upstream and downstream data traffic [1]. In the EPON systems, based on the principle of multi-point control protocol (MPCP) the OLT knows the time of data transmission for ONUs. This paper uses the predefined time slot of each ONU and then additional sleep control signal is not required.

In reference [2], a new green bandwidth allocation (GBA) framework has been introduced where hybrid sleep mode for ONUs and batch mode transmissions for the OLT and ONUs are used. The sleep mode for ONU is determined emerged on the maximum delay whereas in batch mode transmission, the ONU can reduce the overhead time during the switching. A batch mode transmission at the ONUs and OLT is still the outlay of the Quality of Service (QoS) losses due to packet delay. To overcome this problem, a new analytical model has been presented in [3], in which the downstream packet delay is analyzed at the ONU sleep time by considering the limited service discipline with the overlapped time window in upstream transmission.

In an analytical model, the ONU receiver running speed and the expended energy is not sufficient condensed. To overcome this problem in reference [4], the authors studied the downstream burst transmission (DBT) for energy saving.. So, the receiver data processing follows the on and off control. The pre-scheduled data is arrived in the downstream buffer because it enters the power on mode. The downstream data packets are destined for other ONUs by entering the power off mode. In a sleep-controlled rule, the handshaking between the OLT and ONUs is involved. In reference [5], a sleep-controlled rule using semi-Markov chain method has been proposed. This proposed scheme doesn't necessitate the handshaking to attain high energy savings efficiency and high bandwidth utilization. However, the absent of downstream data traffic is detected for ONU in the scheduling cycle followed by the ONUs into sleep mode.

In a sleep-controlled rule, the semi-Markov chain model for the receiving ONU is defined only in the downstream traffic whereas any downstream traffic is not lost because of ONU's sleep time. A sleep-aware dynamic bandwidth allocation (SDBA) algorithm for maximizing energy efficiency has been proposed in reference [6]. Here, reflecting the multiplication of the polling cycle as much as possible for enlarging the ONU sleep time. The ONU is allowed for sleep for allocating the enough bandwidth to transmit traffic.

In a minimum sleeping period of an ONU, to decrease the contrary effects of frequent mode switching a sleep mode aware (SWA) algorithm has been proposed which consists of two kinds of ONUs such as legacy ONUs (L-ONUs) and next generation ONUs (NG-ONUs). The L-ONUs are assigned for

large sleep overheads and the NG-ONUs are employed for small sleep overheads. This algorithm also uses a novel sleep period prediction scheme. As a result of frequent switching process, the ONUs cannot have plenty sleep time and the energy efficiency is narrowed even under light traffic loads. To exceed this problem in [7], the authors signified the energy efficiency dynamic bandwidth allocation (EE-DBA) algorithm. In this scheme, two predefined sleep times are used such as maximum and minimum sleep times. The maximum sleep time is calculated for OLT when the accumulated packet and delay can assure the shortcoming of buffer status and end-to-end packet delay. The power saving mode is not allowed for ONUs when there is no sleep time.

In the EE-DBA scheme, the OLT always stays at the active mode rather than considering if there exist any Report message or not, consequence into the devour of power of OLT at all the time. To overcome this problem, in this paper, we propose a sleep-based DBA (SB-DBA) algorithm for the EPON systems. In the proposed SB-DBA scheme, we use sleep time for both OLT and ONUs based on the upstream and downstream data traffic. So, both the ONUs and OLT can reach the power saving mode simultaneously for minimizing the energy consumption of the EPON systems. A ONU steps into the power saving mode when there are no data packets in the buffers of that ONU. Similarly, the OLT also steps into the power saving mode when there is no Report message from an ONU. Performances of the proposed SB-DBA scheme are evaluated in terms of total energy consumption and energy savings ratio by numerical simulation. The simulation results confirm that the performance of the proposed SB-DBA scheme outperforms the existing EE-DBA scheme.

The paper organization is as follows. Section 2 explains the principle of the proposed SB-DBA scheme. Section 3 presents the performance analysis for numerical simulation. In section 4 and 5, the simulation results are explained and concluded the paper, respectively.

II. Principles of the Proposed SB-DBA Schemes

In this section, we 1st we explain about the principle of the existing EE-DBA scheme then the principle of the proposed SB-DBA scheme. Fig. 1 shows a basic network architecture of EPON system where *N* ONUs and single OLT are connected through a passive splitter in tree topology.

A. Proposed SB- DBA algorithm

In the existing EE-DBA scheme [8], the OLT is always in awake mode whether there is Report message or not from the ONUs to receive. Such continuous active state of the OLT leads to the misuse of power which is the major drawback of

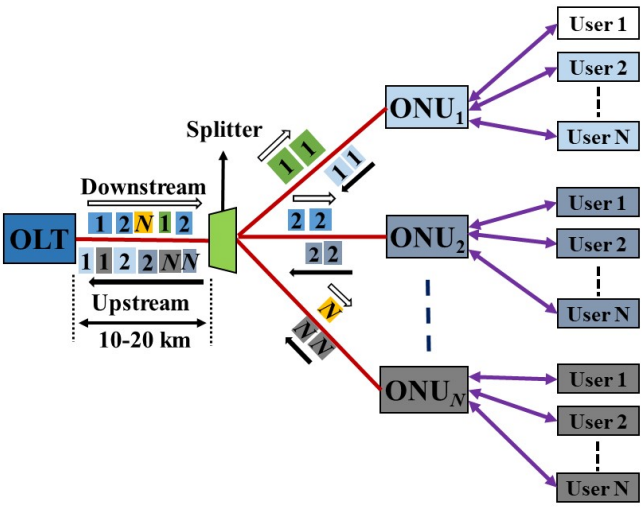


Fig. 1. Tree based EPON architecture.

the existing EE-DBA scheme. For solving such problem, in the proposed SB-DBA scheme, we encourage the OLT to enter into power savings mode in addition to the ONUs.

Figure 2 shows the data transmission scenario of the proposed SB-DBA scheme using two ONUs and single OLT. Based on the DBA principle the packet distributions in the Figure prove that the downstream and upstream transmissions

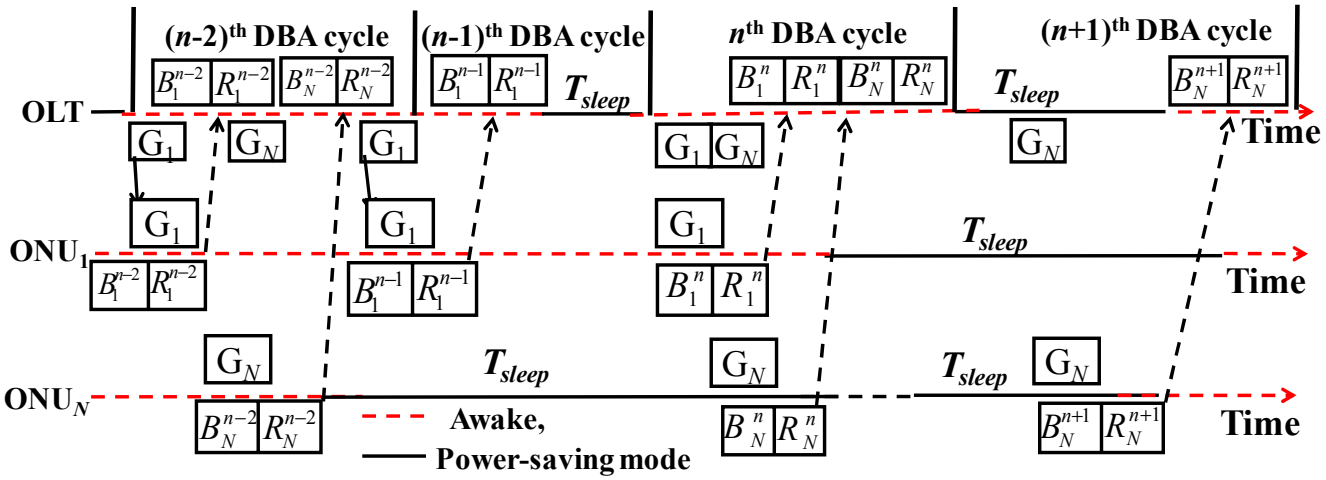


Fig. 2. Illustration of awake and power saving modes of the proposed SB-DBA scheme.

have no fixed bandwidth. Here, ONUs transmit the Report message to the OLT and the OLT checks the circumstance and delivers Gate messages to the ONUs and receives the upstream data packets. So, an ONU enters into the awake mode if it has data packets in the buffer. In contrast, the ONU ignites the power saving mode if there is no data packet in the buffer. From the Fig. 2, it is clear that the OLT may receive the data packet from the two ONUs or one ONU or none of the ONUs. If the OLT receives the data packet from both the ONUs, then the OLT remains in fully awake mode this case is illustrated in the $(n\text{-}2)^{th}$ DBA cycle of the Fig. 2. If the OLT receives the data packets from one ONU, then the OLT does not remain in awake mode for whole the DBA cycle. In this condition, sometime the OLT is in awake mode and sometime in power saving mode. If the OLT does not receive any Report message from both the ONUs, then the OLT conforms into fully power saving mode.

B. Methodology of the Proposed SB-DBA Scheme

Figure 3 shows the flowchart of the proposed SB-DBA algorithm. In the initial phase of individual DBA cycle, OLT checks the current buffer status of all the ONUs based on the received Report messages. If the condition $W_R = W_{max}$, where, W_R is the requested window size of an ONU in a time cycle, and W_{max} is the maximum window size of an ONU, is satisfied, then the ONU is fully awake mode and so the OLT also remains in awake mode because the ONU transmits the data packets and the OLT has to receive those data packets. In this condition, awaken time period for the ONUs and OLT are calculated using the following equation.

$$Awake\ \ time\ = \frac{W_R}{W_{\max}} \times 100\ \% \tag{1}$$

If the ONU has no data packet in the buffer, then it remains in sleep mode and the OLT also enters into the power saving mode.

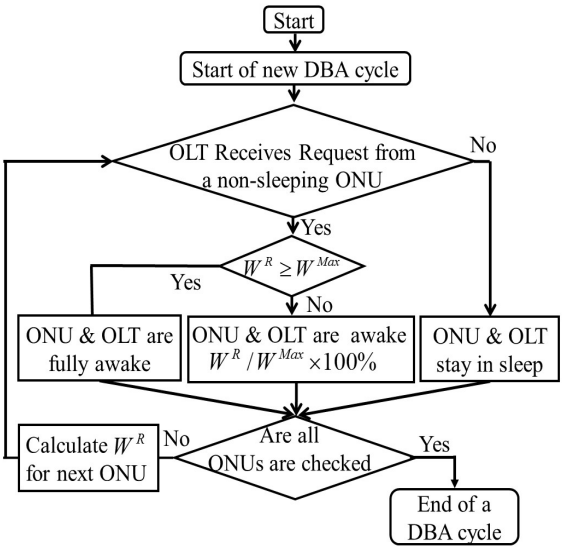


Fig. 3. Flowchart of the EE-DBA scheme.

III. PERFORMANCE ANALYSIS

The main features of the proposed SB-DBA scheme are making the EPON system energy efficient by employing active and power saving modes for both the OLT and ONUs. In this section, we provide detail analysis of the energy consumption in both active and sleep mode condition mode in the proposed SB-DBA scheme.

A. Energy Consumption in Active Mode

Let, W_{max} is the maximum window size for each ONU then following equation is used to calculate the W_{max}.

$$W_{\max} = (T_{cycle} - N * T_g) / N \tag{2}$$

where, T_{cycle} is the maximum cycle time, N is the number of ONUs, and T_g is the guard time.

If W_R is the requested window size by an ONU. Then equation (3) is used to calculate the W_R^i of an ONU *i.*

$$W_R^i = n^i (E_o + P_l) + R_l \tag{3}$$

where, E_o is the Ethernet overhead, P_l is the packet length, R_l is the length of Report message, and n^i is the number of packets in the buffer of an ONU *i.*

Following formula is used to calculate the active time of an ONU *i* T_{act}^{ONUi} in a time cycle.

$$T_{act}^{ONUi} = \begin{cases} 0 & \text{if } W_R = 0 \\ W_R & \text{if } W_R \leq W_{\max} \\ W_{\max} & \text{if } W_R > W_{\max} \end{cases} \tag{4}$$

Then, the active time of the OLT T_{act}^{OLT} in a time cycle is calculated as follows:

$$T_{act}^{OLT} = \sum_{i=1}^{N} T_{act}^{ONUi} \tag{5}$$

Equation (6) is used to calculate the total energy consumption in active mode by the OLT and ONUs in a time cycle *j*.

$$E_{act}^{Total(j)} = \left(T_{act}^{OLT} + \sum_{i=1}^{N} T_{act}^{ONUi} \right) P_a \tag{6}$$

where, $E_{act}^{Total(j)}$ is the total energy consumption in active mode in a time cycle *j*, and P_a is the power consumption by the OLT or ONUs in active mode.

B. Energy Consumption in Sleep Mode

As equation (4) is used to calculate the active time of an ONU *i* in a time cycle then following equation is used to calculate the sleep time of an ONU *i* T_{sleep}^{ONUi} in a time cycle.

$$T_{sleep}^{ONUi} = T_{cycle} - T_{act}^{ONUi} \quad (7)$$

Then formula (8) is used to calculate the total sleep time of all the ONUs and OLT in the EPON system.

$$E_{sleep}^{Total(j)} = \left(\sum_{i=1}^{N} T_{sleep}^{ONUi} + (T_{cycle} - T_{act}^{OLT}) \right) P_s \quad (8)$$

where, $E_{sleep}^{Total(j)}$ is the energy consumption in the sleep mode of a time cycle j, and P_s is the power consumption by the OLT and ONUs in the sleep mode.

Total energy savings in a time cycle is calculated using formula (9).

$$E_{saving}^{Total(j)} = (N+1) T_{cycle} P_a - E_{sleep}^{Total(j)} \quad (9)$$

where, $E_{saving}^{Total(j)}$ is the total energy savings in the proposed SB-DBA scheme in a time cycle j.

IV. Simulation Result

In this section, we have compared the system performance of the proposed SB-DBA scheme with the existing EE-DBA scheme in terms of energy consumption, energy savings in terms of percentage and packet delay of the EPON system. All performance analyses were done numerically for different range (0 to 1.0) of non-uniform normalized offered load (NOL)with cycle time variation from 1 ms to 3 ms and the data rate in the range of 0.5–2 Gbps. In the simulations, we have considered a tree topology based EPON with 16 ONUs and one OLT. We also have considered the power consumption in active and sleep modes of OLT and ONUs are 3.85 W and 1.7 W, respectively [8]. The simulation parameters are summarized in Table I.

Table I Simulation Parameters.

Symbol	Description	Quantity
N	Number of ONUs	16
T_{cycle}	Maximum cycle time	1 – 3 ms
T_g	Guard time	5 µs
E_o	Ethernet overhead size	576 bits
R_l	report message size	304 bits
P_l	Maximum Packet length	1500 bytes
R_n	Transmission data rate	0.5 to 2 Gbps
P_a	Power consumption of OLT/ONU in active mode	3.85 W
P_s	Power consumption of OLT/ONU in sleep mode	1.7 W
D	OLT and ONUs distance	10 – 20 Km

Fig. 4 presents a comparative results of the energy consumption vs. NOL for the OLT between the proposed SB-DBA and the existing EE-DBA schemes for 2ms cycle time and different data rates of 1 and 2 Gbps. From the results it is clear that the energy consumption by the OLT in the proposed SB-DBA scheme is far better than the existing EE-DBA scheme. Actually, the existing EE-DBA scheme requires

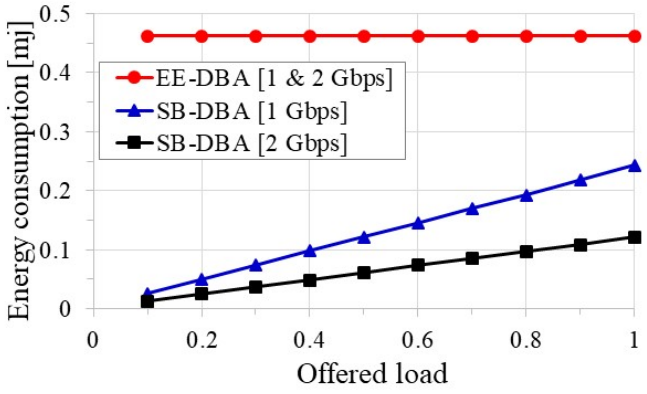


Fig. 4 Comparison of energy consumption for the OLT and maximum time cycle of 2 ms.

almost constant energy consumption for any value of offered load and data rate. The main reason of this result is that in the EE-DBA scheme energy saving technique has not been considered for the OLT. In contrast, the proposed SB-DBA scheme provides less energy consumption for the OLT than the existing EE-DBA scheme for any value of the NOL and data rate. Moreover, with increase of data rate the energy consumption in the proposed scheme is reduced. In the proposed SB-DBA scheme, at the NOL of 1.0 the maximum values of the energy consumptions are 0.24 and 0.12 mj for 1 and 2 Gbps data rates, respectively. Whereas, the energy consumption in the existing scheme is 0.462 mj for any value of NOL and data rate.

Fig. 5 represents a comparison of the percentage of energy savings for the OLT among the proposed SB-DBA and the existing EE-DBA schemes for data rates of 1 and 2 Gbps. The proposed scheme shows the highest energy savings for the OLT whereas the existing scheme has no energy savings. At the data rate of 2 and 1Gbps, the SB-DBA scheme saves 97% and 95% energy, respectively, at the lowest NOL of 0.1.

Fig. 6 compares the overall energy consumption, i.e., summation of energy consumptions by OLT and ONUs, between the proposed SB-DBA and existing EE-DBA schemes for 2 ms cycle time. The proposed SB-DBA scheme provides less energy consumption than the existing EE-DBA scheme for any value of NOL and data rate. This is because

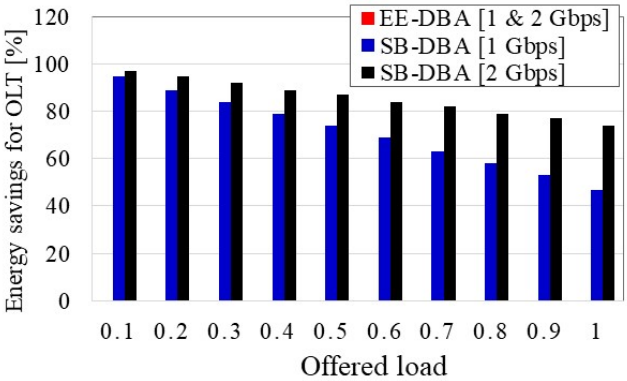


Fig. 5 Percentage of energy savings for OLT and 2 ms cycle time.

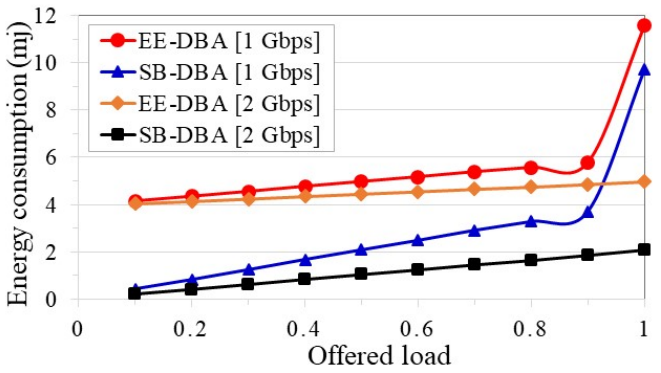


Fig. 6. Comparison of overall energy consumption.

that the proposed scheme considers the energy saving technique for both the OLT and ONUs. The maximum values of energy consumptions in the proposed scheme for 1 and 2 Gbps data rates are 9.8 and 2.1 mj whereas those of the existing scheme are 11.5 and 5 mj at the maximum NOL of 1.0. So, it is clear that the performance improvement in the proposed scheme is more significant at higher data rates and lower NOL.

Fig. 7 provides a comparison of the percentage of overall energy savings between the proposed SB-DBA and existing EE-DBA schemes for different data rates of 1 and 2 Gbps. At the data rates of 1 and 2 Gbps the proposed SB-DBA scheme saves the highest energy that is 98% whereas that of the existing scheme are 19%, and 64%, respectively, at the lowest offered load of 0.1.

From the analysis of above-mentioned results, it is found that the proposed SB-DBA scheme provides better performance than the existing EE-DBA scheme. However, the end-to-end packet delay of the proposed SB-DBA scheme is almost similar to the existing EE-DBA scheme. Figs. 8(a) and 8(b) show the comparison of end-to-end packet delays for the proposed SB-DBA and existing EE-DBA schemes for 2 ms cycle time and data rate of 1 Gbps. Here, we see that the delay of the proposed scheme is similar to that of the existing scheme.

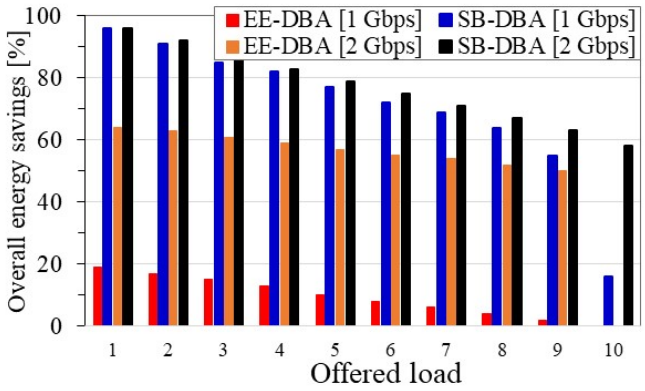
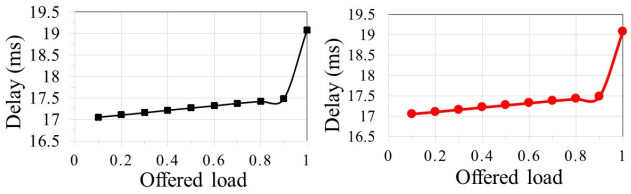


Fig. 7 Percentage of overall energy savings.



(a) SB-DBA (b) EE-DBA

Fig. 8 End-to-end packet delay of SB-DBA and EE-DBA schemes.

V. Conclusion

In this paper, we have proposed a new DBA scheme called sleep-based DBA (SB-DBA) scheme for reducing the energy consumption in the PON system. The performances of the proposed SB-DBA scheme are evaluated in terms of energy consumption, percentage of energy savings and end-to-end packet delay. The proposed SB-DBA scheme provides better energy efficiency than that of the existing EE-DBA scheme while maintaining the same delay performance. The proposed DBA scheme provides 98% energy savings for the 2 ms time cycle for the data rate and offered load of 2 Gbps and 0.1, respectively, whereas the existing EE-DBA scheme provides 64% energy savings.

References

[1] M. M. Rahman, M. Hossen, and R. Bushra, "Control message scheduling algorithm for improving throughput and jitter performance of the MHSSR DBA scheme of PON", IEEE Int. Conf. on Elec., Comput. and Commun. Eng. (ECCE), Cox's Bazar, Bangladesh, February 16-18, 2017.

[2] R. Dhaini, P. H. Ho, and G. Shen, "Energy efficiency in Ethernet passive optical networks: For how long can ONU sleep?" University of Waterloo, Tech. Rep., 2011.

[3] S. Chen, A. R. Dhaini, P. H. Ho, B. Shihada, G. Shen, and C. H. Lin, "Downstream-based scheduling for energy conservation in green EPONs," J. Commun., vol. 7, no. 5, pp. 400–408, 2012.

[4] Y. Luo, M. Sui, and F. Effenberger, "Energy-efficient next generation passive optical network supported access networking, J. Opt. Switch. Net., vol. 14, no 1, pp. 43–52, 2014.

[5] J. Zhang, M. T. Hosseinabadi, and N. Ansari, "Standards-compliant EPON sleep control for energy efficiency: design and analysis," J. Opt. Commun. Net., vol. 5, no. 7, pp. 677–685, 2013.

[6] D. P. Van, L. Valcarenghi, M. Chincoli, and P. Castoldi, "Experimental evaluation of a sleep-aware dynamic bandwidth allocation in multi- ONU 10G-EPON testbed," J. Opt. Switch. Net., vol. 14, no. 1, pp. 11–24, 2014

[7] A. Dixit, B. Lannoo, D. Colle, M. Pickavet, P. Demeester, "Energy efficient dynamic bandwidth allocation for Ethernet passive optical networks: overview, challenges,

and solutions," J. Opt. Switch. Net., vol. 18, no. 2, pp. 169–179, 2015.

[8] M. T. Rahama, M. Hossen, and M. M. Rahman, "A routing protocol for improving energy efficiency in wireless sensor networks," IEEE International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT), MIST, Dhaka, Bangladesh, September 22-24, 2016.

Server Placement for Serving VIP Clients in Case of Link Failure

Shafayat Ahmed, Sourav Kumar Mondal, Sadia Sharmin and Md. Saidur Rahman
shafayat309@gmail.com, shomoyerotithi@gmail.com, sadiasharmin.ss@gmail.com, saidurrahman@cse.buet.ac.bd
Department of Computer Science & Engineering
Bangladesh University of Engineering & Technology

***Abstract*— Designing and operating a reliable network which is robust against network failures is crucial for internet service providers. Especially, a network must be sufficiently connected even if link failures occur. Since link protection is costly, it is important to place the servers in such a way that a network suffering from link failures can still provide connectivity to some special clients. In this paper, we focus on the placement of servers assuring that special clients named as VIP clients remain connected to the servers even if links fail. We first set some criterion for serving the VIP client so that they get priority in case of link failure. We then propose an approximation algorithm that considers the trade-off between path length and alternative paths from the servers. We compute scores for various server placement combinations and try to provide the best suitable placement of servers. We have experimented with different types of synthetic graphs and tested our approach in case of link failures. Results indicate the effectiveness of our approach to maintain a sustainable network.**

Keywords — Link protection, Connectivity, Server Placement, VIP Clients, Average Distance, Approximation algorithm.

I. INTRODUCTION

Network Service Providers always want to provide service which is reliable and stable. When any link drops, the network may be disconnected or has to get connected via different route. This is called link failure. Critical links are those links whose failures drastically destroy the connectivity and must be protected. Such a failure prevents a large number of users from access to the servers. In case of link failures if there are no alternative paths from server to the client then connection is disrupted and can happen due to failure of routers, cutting of fibers, congested connection lines and so on. Moreover there are some clients among the network whom we can identify as valuable or VIP clients and failure to provide service to these clients can have severe damage to a country or society. For example banks, hospitals, universities, civil defense, security service and so on. Therefore, this clients need to be taken special care while serving network connections and given back-up in case of failure.

Moreover our working approach of placing service providers across a network for serving valuable clients can serve various other applications. Like where one shopkeeper can set up his food court considering school and universities as VIP clients, a pharmacy can be set up considering availability of hospitals around and how easy it is to get to the pharmacy from the hospitals etc.

A. Related Work

Designing a fully robust network is not simple. In this section, we are going to overview some of the recent approaches regarding network connectivity and link failure.

In 2004 Markopoulou et. al showed that 70 percent of unintentional failures, excluding failures caused by maintenance, are single-link failures [1]. The problem of determining a set of edges to meet a required edge connectivity or vertex-connectivity by adding edges to a graph was addressed in [2], [3], [4]. Approaches for determining a set of edges to meet a required diameter by adding edges to a graph was described in [5], [6].

From the viewpoint of link protection, researchers have investigated the smallest number of links to be protected so that the diameter of a network resulting from failures of non-protected links is less than or equal to a threshold in [7] and [8]. It was proved that this problem is NP-hard [7], and an algorithm was presented to solve the problem in polynomial time when the number of failed links is restricted. Imagawa and Hiroyoshi [9] defined and analyzed to find the links to be protected so that, even if any non-protected links fail, the number of the nodes connected with a server in the resulting network is not less than a threshold. Irie, Kurimoto and Miwa showed an approach to find the links to be protected so that, even if any non-protected links fail, all connected components include specified servers in a resulting network [10].

B. Notations

A graph $G = (V, E)$ consists of V, a nonempty set of vertices (or nodes) and E, a set of edges. Each edge is an unordered pair of vertices. An edge between two vertices u and v is often denoted by (u, v). We also denote the set of vertices of a graph G by $V(G)$ and the set of edges of G by $E(G)$. Path is a non-empty graph $P = (V, E)$ of the form $V = \{x_0, x_1, ..., x_k\}$ and $E = \{(x_0, x_1), (x_1, x_2), ..., (x_{k-1}, x_k)\}$, where the x_i are all distinct. The number of edges in a path is called its length. Two vertices u and v of graph G are said to be connected if there is a path in G whose end vertices are u and v. In a graph G, the distance between two vertices u and v, denoted by $d(u, v)$, is defined to be the length of a shortest path joining u and v in G. The diameter of G , denoted by $D(G)$, is the maximum distance over all pairs of vertices in G.

II. Our Approach

In our work, instead of protecting links or selecting critical links or minimizing cost of protecting links we concentrate on serving significant clients more reliably by placing servers in a robust way. In a network, all clients are not of same significance. We propose an algorithm for deploying a pair of servers in a network for serving the VIP clients more reliably. We focus on the properties for optimal selection of servers for a client. Our algorithm considers some trade-offs between fitness scores of server pairs and time complexity. We experimented on various types of networks like dense, sparse, small and large and demonstrate the effectiveness or our algorithm by dropping random links.

A. Evaluation Criteria for the Fitness of Server Pairs

We consider the following three criteria for deciding a server pair for serving VIP clients.

- **Number of alternative paths from a server to client:** If a server has multiple routes to a client, then if any edge drops which we say link fails, then it can still provide connectivity to that client using alternative routes. More the alternative paths, greater the chance of surviving link failures.
- **Distinct paths from a server pair to a client:** If the paths from server pairs to a specific client has less number of common edges then the chance of sustaining link failure is increased. So while choosing between multiple server pairs the pair with distinct routes to the clients will get advantage.
- **Distance from server to client:** Sometimes it is worth choosing long distance paths for distinct and multiple alternate paths rather than having short and less disjoint path. So while placing servers this is also taken in consideration.

B. Algorithm

Suppose we are given a network, considering it as a Graph G(V,E), there is a set of VIP clients C. We will select a pair of server according to our algorithm which will ensure better connectivity to those VIP clients.

First we need to find all the routes to the VIP clients from all other nodes of the network. But finding all the routes between two nodes in a graph is an NP-hard problem [11]. Therefore, instead of computing paths from all other nodes to VIP clients, in our experiments we selected only nodes with higher degrees. Because it is more likely that the higher degree nodes will have better chances of providing routes to the VIP clients as they are more likely to have alternate paths. Algorithm 1 illustrates the server placement procedure using all path listing procedure $allpaths$ and server ranking procedure $evaluationval$.

Algorithm 1 Algorithm for Server Placement

Input: A graph $G = (V, E)$, VIP client set C.
Output: Server Pair Set V'

```
V' ← φ
Bestval ← −1
int e ← 0
for i ← V, i ∉ C do
    for j ← V, j ∉ C do
        e ← evaluationval(i, j)
        if e > Bestval then
            V' ← (i, j)
            Bestval ← e
        end if
    end for
end for
return V'
```

In Algorithm 2, we take all server pair combinations and calculate the evaluation value for each pair. Calculating evaluation value is done considering the selected priorities. The evaluation function takes two server nodes, VIP clients and the graph as inputs. It calculates combinations of paths to a VIP client and generates an evaluation value.

Algorithm 2 Evaluation Value Computation ($evaluationval(i, j)$)

Input: A graph $G = (V, E)$, VIP client set C, Server pair (i, j).
Output: $Evaluation\ Value$

```
D ← |E(G)| − log(|E(G)|) + 1
Evaluationval ← 0
for each k ∈ C do
    for allpaths(i, k) do
        for allpaths(j, k) do
            Evaluationval = Evaluationval + D − commonedge − log(averagedistance)
        end for
    end for
end for
return Evaluationval
```

The evaluation value calculation is quite simple. We prioritize having alternate paths over distance and common paths using Algorithm 3. For every combination of paths, their evaluation value keeps adding to overall evaluation value calculation of the server pair candidate. This is ensured with the help of a constant D. Algorithm 3 is nothing but Breadth First Search (BFS) operation. We will use BFS approach to find all the paths from a particular node to another node.

The number of common edges is subtracted for particular path pair to a VIP client. It can be noted that for certain path pair highest number of common edges possible is the number of total edges in the network.

Again, we considered distance for the server pair candidate using the concept of **average distance** into action. We took the path length of both of the server to VIP node

Algorithm 3 Find All Paths ($allpaths(i, j)$)

Input: A graph $G = (V, E)$, pair(src, des).
Output: All paths from src to des, $src \in V, des \in V, src \neq des$

```
1: Allroutes ← φ
2: bool * visited
3: int * path
4: int * pathindex ← 0
5: for each v ∈ V do
6:    visited[v] ← false
7: end for
8: for each u ∈ adj[i] do
9:    visited[u] ← true
10:   path[pathindex] ← u
11:   pathindex + +
12:   if (u == j) then
13:      Add path to Allroutes
14:   else
15:      for each v ∈ adj[u] do
16:         if (visited[v] == false) then
17:            Proceed from v
18:         end if
19:      end for
20:   end if
21: end for
22: return Allroutes
```

considering common edges only once and call it average distance. The logarithm value of this average distance is subtracted because the larger the graph gets the impact of distance minimizes. Logarithm works perfectly to reflect the situation. The calculated average distance can be highest upto the logarithm value of the total number of edges in the network. So to obtain a positive value for every possible alternate pair of paths we added a constant value D with the evaluation value.

$$D \leftarrow |E(G)| - \log(|E(G)|) + 1 \tag{1}$$

In our experiments, we did not take all the alternative paths but only the paths with shortest distance and second & third shortest length paths into consideration for reducing complexity.

Algorithm 1 selects the server pair with highest evaluation value after trying all possible server pairs.

III. Time Complexity and Optimization

In this section we go through the time complexity of our algorithm and optimization techniques.

A. Time Complexity

Lets assume the given resources of a network graph G are,

- n : number of vertices.
- m : number of edges.
- c : number of VIP clients

Now in our algorithm we need to find all the routes between server candidates to the VIP clients. This is a time consuming operation and is of order $O(n!)$.

The main algorithm needs to check all possible pair of servers which takes $O(n^2)$ time. So, overall the algorithm becomes
$O(n^2) * O(c * (n!^2))$ = $O((n^3) * (n!^2))$

B. Optimization

Our proposed algorithm takes exponential amount of time, which is not suitable for practical applications. Therefore we employ some optimization techniques to improve the situation.

1) Reduce Candidate Nodes: Finding all the routes from the network to a VIP client takes an incredible amount of time. So for general purpose we try to cut down number of candidate nodes. From the evaluation criteria we know that the more the number of alternative paths the greater the chance of being selected as server to ensure a robust network. Therefore we look for the candidate nodes which have the highest degree or close to the highest degree. This prohibits a significant amount of nodes from being a candidate for server nodes. This speeds up the algorithm to a greater extent and in general shows reasonable result which we present in later sections.

2) Reduce number of routes: If we have to find all the routes, then it will take exponential amount of time. For this problem, we try the shortest and close to the shortest routes only. It helps to reduces computation time. For computing the shortest route from a server candidate to a VIP client, we use famous shortest path finding algorithm of Dijkstra [12].

3) More Number of Servers: Our main approach is to select optimum server positions for 2 servers. If the network is bigger and need more servers to serve the purpose, then we divide the network into subnetworks of smaller size and run our approach for all those subnetworks.

Suppose, we have a network of 10000 nodes and 100 VIP clients in it and need to place 20 servers. Then we divide the network into 10 subnetworks of 1000 nodes and then run our algorithm for each of the subnetworks.

IV. Performance Assessment

In this section, we discuss the performance assessment of our proposed algorithms.

A. Performance Assessment Method

Our performance assessment approach consists of three consecutive steps :

- Randomly dropping some links.
- Checking connectivity of the graph of resulting network.
- If disconnected already, checking if all the VIP clients are connected to at least one server.

1) Random Link Failures: To test the effectiveness of our solution we randomly drop some links from a sample network graph. We do this step by step. Slowly we increase the percentage of failed links from 5% to 100% and check

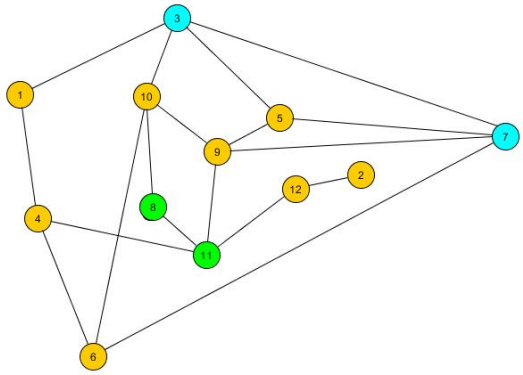


Fig. 1. A sample network graph with two VIP client node 8 and 11

if the servers previously chosen by our proposed algorithm can provide connectivity to all the VIP clients.

In Figure 1, green colored nodes are VIP clients and blue colored nodes are chosen as servers by our algorithm. Number of edges of this graph is 18. There are 2 VIP clients: node 8 and node 11. Node 3 and node 7 are chosen as servers.

2) Checking Connectivity of Resulting Network: In this step, we check for disconnectivity in the resulting network graph after random link failures. In a connected graph, each node must be reachable from all the other nodes. In Figure 2, VIP client nodes are 8 and 11, our previously selected servers are node 3 and 7. Let random failures occur for links $(3, 5), (3, 7), (4, 6)$ and $(4, 11)$. We can see, as number of connected component is 1 and the graph is not disconnected. There is still a route between previously selected server node 3 and VIP client node 8, represented by a bold edge in figure..

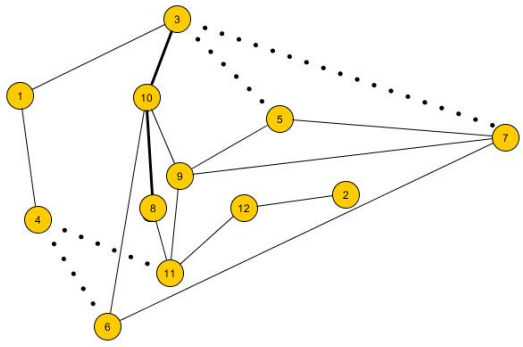


Fig. 2. An existing route between proposed server node 3 to VIP client node 8

3) Checking Service Reachability to VIP Clients: If the resulting network graph is disconnected after random link failures, then there is a possibility that any one or more of the VIP clients are no more receiving service from any server. This can happen if there is any VIP client left out in such a subgraph that has no server node in it. For this scenario, we check reachability from server nodes to all VIP clients. If any VIP client has connection with at least one server, we can declare that the client is getting uninterrupted service. We use BFS (Breadth First Search) for this reachability test.

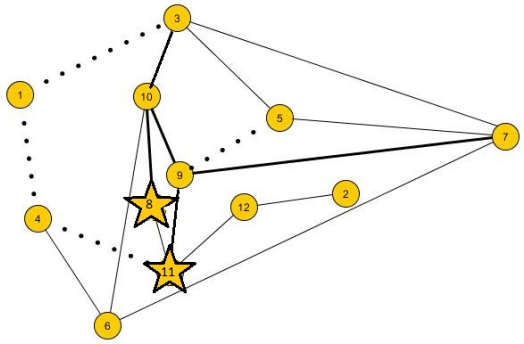


Fig. 3. All VIP clients have reachability to servers

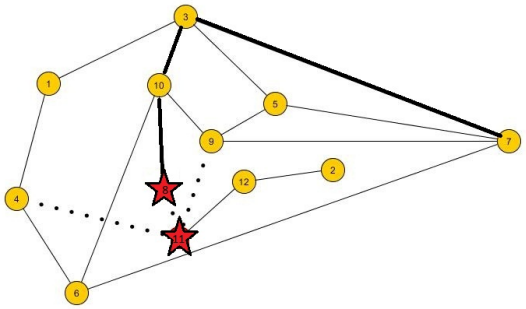


Fig. 4. VIP client 11 has no reachability to any server

In Figure 3, we can see that VIP clients are 8 and 11, selected servers are 3 and 7. 8 and 11 both have server connectivity. But for same network graph with different set of failed links, we can see in Figure 4, VIP client 11 is not reachable from any server.

B. Performance Evaluation for Variable Network Size and Density

In this section, we present the results of our simulation experiments done on sparse and dense graphs of variable sizes. In a sparse network there are more chances of disconnectivity and server failures. The result of our experiments agrees with this claim.

1) Small Networks: Our small networks consist of 10 to 99 nodes and 2-5 VIP clients. We increase the percentage of edge drop by 5% in each iteration and simulate our test 10 times. The results show that the success rate drops with the change of density of edges in the graph but the drop is significantly lower if the maximum degree is higher i.e. the graph is dense.

2) Large Networks: The large random networks consist of 100 to 1000 nodes. We increase the percentage of edge drop by 5% in each iteration and simulate our test 10 times. Results show the drop of success rate with the increase of edge drop in the graphs.

For networks with number of vertices 1000 and 10 VIP clients Figure 5 shows that success rate drops rapidly with

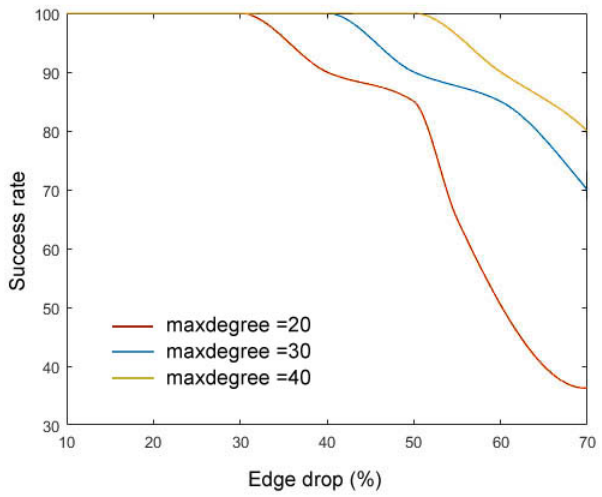


Fig. 5. Success Rate vs Edge Drop (%), Number of Vertices =1000, VIP Clients=10

increasing link failures for networks with nodes having smaller maximum degree. On the other hand networks with maximum degree 40 provide connectivity to all VIP clients even with 50% random link failures.

3) Computation Time for Different Network Size and Density: The server selection for a dense network should take more time than for a sparse network. Again, computation on large network takes much more time than for a small network. The following table shows the time taken by our algorithm to select required number of servers in networks of variable size and density.

TABLE I

PERFORMANCE EVALUATION REGARDING TIME NEEDED FOR VARIABLE NETWORK SIZE AND DENSITY

Number of Vertices	Maximum Degree	Number of Edges	Computation Time(s)
50	10	123	11.06
	15	190	11.347
	20	317	12.184
100	10	290	31.009
	20	546	31.173
	30	802	32.055
500	10	1132	979.487
	20	2547	968.113
	30	3755	967.567
1000	20	4460	4202.850
	30	6848	4208.063
	40	9377	4210.78

In Table I we can see that for the same number of nodes with the increase of edge density, the computation time doesn't change that much. But with the change of network size, time increases moderately fast. With respect to the complexity of the algorithm we can say our optimization approach has quite a success for large graphs.

C. Performance Evaluation for Variable Number of VIP Clients

In the following set of experiments, we present the results of our simulation experiments done on network graphs of variable sizes with variable number of VIP clients.

1) Small and Large Networks: We have run simulation on small and large networks that consist of 10 to 1000 nodes. The percentage of edge drop is increased by 5% in each iteration and each test is run 10 times. Number of VIP clients are varied from 5 to 15. Table II presents the results for the networks with 1000 vertices.

TABLE II

PERFORMANCE EVALUATION FOR 1000 VERTICES' NETWORK

Number of VIP Clients	Edge Drop(%)	Success Rate(%)
5	40	100
	50	90
	60	85
	70	70
	80	45
10	30	100
	40	90
	50	85
	55	70
	60	40
15	20	100
	30	90
	40	80
	45	70
	50	45

From Table II we can see that the success rate drops rapidly when we have to serve more VIP clients. Figure 6 shows it graphically. Here the network has 1000 nodes and maximum degree is 20. When the number of VIP clients is 1% of the total network size, the selected servers serve them

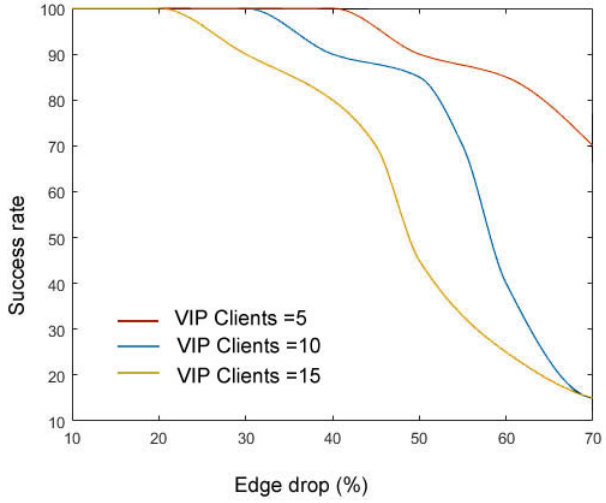


Fig. 6. Change of Success Rate vs Edge Drop (%) Varying Number of VIP Clients and Network Size=1000

2) Computation Time for Different Network Size and Density while varying the number of VIP Clients: The following table presents time taken by our algorithm to select servers in networks of variable sizes and densities.

TABLE III
PERFORMANCE EVALUATION REGARDING TIME NEEDED FOR VARIABLE NUMBER OF VIP CLIENTS AND NETWORK SIZE

Number of vertices	Number of VIP clients	Computation time (s)
10	2	0.204
	3	0.439
	4	0.464
50	5	11.068
	10	23.158
	15	40.051
100	5	42.229
	10	83.713
	15	135.37
500	5	712.063
	10	1406.290
	15	2074.17
1000	5	2138.56
	10	4202.85
	15	6223.172

Table III shows the increase of computation time with the increase of the number of VIP clients for fixed sized networks. For large networks computation time increases significantly with the increase of Number of VIP clients than for smaller networks.

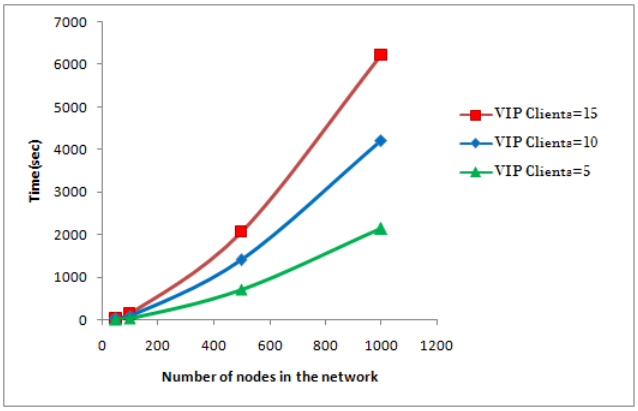


Fig. 7. Computational time needed for different networks with varying the network size for fixed number of VIP Clients

Figure 7 organizes the computational time needed to select servers for the same number of VIP clients while increasing the network size, which tends to be a polynomial function of the number of nodes in the networks.

V. CONCLUSION

Protecting links is always costly and selection of the links to protect is hard. Placing the servers in right positions is important so that clients having significance in a network always finds an alternate route in case of link failures. We have defined the criteria for selecting servers for serving VIP clients, provided an algorithm that can place a server pair in quite a large network and still give a reasonable choice. We have also shown how the algorithm reacts with various change of properties of the network. Moreover we have experimented our solution by link failures at random basis and shown its robustness for keeping the VIP clients connected to the service providers. We have analyzed to reduce the complexity of our algorithm, set various criterion to optimize it.

Last but not the least experiments with larger, weighted networks are yet to be explored. Moreover we have experimented with synthetic graphs only. Performance of our strategy in case of real network will also help to strengthen our claim.

REFERENCES

[1] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C.-N. Chuah, and C. Diot, "Characterization of failures in an ip backbone," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 4. IEEE, 2004, pp. 2307–2317.

[2] A. Frank, "Augmenting graphs to meet edge-connectivity requirements," *SIAM Journal on Discrete Mathematics*, vol. 5, no. 1, pp. 25–53, 1992.

[3] G. Kortsarz, R. Krauthgamer, and J. R. Lee, "Hardness of approximation for vertex-connectivity network design problems," *SIAM Journal on Computing*, vol. 33, no. 3, pp. 704–720, 2004.

[4] T. Ishii and M. Hagiwara, "Augmenting local edge-connectivity between vertices and vertex subsets in undirected graphs," *Lecture notes in computer science*, vol. 2747, pp. 490–499, 2003.

[5] T. Ishii, S. Yamamoto, and H. Nagamochi, "Augmenting forests to meet odd diameter requirements," *Discrete Optimization*, vol. 3, no. 2, pp. 154–164, 2006.

[6] V. Chepoi and Y. Vaxes, "Augmenting trees to meet biconnectivity and diameter constraints," *Algorithmica*, vol. 33, no. 2, pp. 243–262, 2002.

[7] K. Imagawa, T. Fujimura, and H. Miwa, "Approximation algorithm for finding protected links to keep small diameter against link failures," in *Intelligent Networking and Collaborative Systems (INCoS), 2011 Third International Conference on*. IEEE, 2011, pp. 575–580.

[8] T. Fujimura and H. Miwa, "Critical links detection to maintain small diameter against link failures," in *Intelligent Networking and Collaborative Systems (INCOS), 2010 2nd International Conference on*. IEEE, 2010, pp. 339–343.

[9] K. Imagawa and H. Miwa, "Detecting protected links to keep reachability to server against failures," in *Information Networking (ICOIN), 2013 International Conference on*. IEEE, 2013, pp. 30–35.

[10] D. Irie, S. Kurimoto, and H. Miwa, "Detecting critical protected links to keep connectivity to servers against link failures," in *New Technologies, Mobility and Security (NTMS), 2015 7th International Conference on*. IEEE, 2015, pp. 1–5.

[11] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York, NY, USA: W. H. Freeman & Co., 1979.

[12] G. T. Heineman, G. Pollice, and S. Selkow, *Algorithms in a nutshell: A practical guide*. " O'Reilly Media, Inc.", 2016.

DDoS Mitigation and Intrusion Prevention in Content Delivery Networks using Distributed Virtual Honeypots

Md. Mahmudur Rahman*, Shanto Roy†, and Mohammad Abu Yousuf ‡
Institute of Information Technology, Jahangirnagar University, Dhaka-1342, Bangladesh*†‡
Dept of IT, University of Information Technology and Sciences, Dhaka-1212, Bangladesh*
Dept of Computer Science and Engineering, Green University, Dhaka-1207, Bangladesh†
Email: dewmahmud*@gmail.com, shantoroy†@ieee.org, yousuf‡@juniv.edu

***Abstract*—Content Delivery Networks(CDN) is a standout amongst the most encouraging innovations that upgrade performance for its clients' websites by diverting web demands from browsers to topographically dispersed CDN surrogate nodes. However, due to the variable nature of CDN, it suffers from various security and resource allocation issues. The most common attack which is used to bring down a whole network as well as CDN without even finding a loophole in the security is DDoS. In this proposal, we proposed a distributed virtual honeypot model for diminishing DDoS attacks and prevent intrusion in securing CDN. Honeypots are specially utilized to imitate the primary server with the goal that the attack is alleviated to the fake rather than the main server. Our proposed layer based model utilizes honeypot to be more effective reducing the cost of the system as well as maintaining the smooth delivery in geographically dispersed servers without performance degradation.**

keywords- Network Security, Virtual Honeypot, Distributed Denial of Service (DDoS), Intrusion, Prevention, Mitigation.

I. Introduction

Denial-of-Service (DoS) attack can be a variety of attacks that are used to stop the system from functioning briefly or indefinitely to prevent authentic users from accessing it [1]. Pinging the system with a load of spam requests results the system to be overloaded and eventually crash due to the process power of the system is restricted. It is accomplished through diverse ways, botnets and buffer overflow vulnerability are the foremost famed among them.

The scale, frequency and sophistication of denial-of-service (DoS) and distributed denial-of-service (DDoS) attacks will continue to grow as businesses and other organizations move more and more online. Using malware to remotely control thousands of computers and direct them to overwhelm a specific target with bogus requests, DDoS attacks cause a CDN network to come slow or unavailable to users, potentially resulting in lasting damage to reputation, productivity and profitability. Apart from data loss, the DDoS attacks cost them a lot of time and money. Although there are many ways of preventing such attacks, new DoS attacks are being constantly introduced as a result of hacking advances.

Honeypot is a system [2] that imitates the primary system by attracting potential intruders who want to gain unauthorized access. It is used to assess the movements and footprints of hackers and to correct protection of the system to avoid future attacks [3]. Generally, it consists of a computer, applications and data that acts and behaves like a decoy. Here we address a method how distributed virtual honeypot pillared with a verification system is used to safeguard a CDN network with greater efficiency and lower cost. A basic architecture of secure CDN network is presented in **Figure 1**.

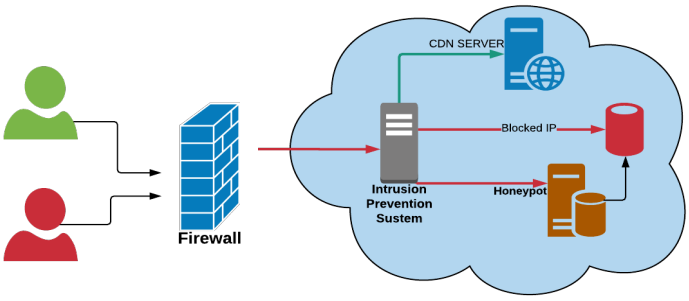


Fig. 1. DDos attack mitigation using Distributed Virtual Honeypot

The primary objectives of our work are as follows:

- To make a honeypot model for mitigating DoS attacks and prevent intrusion in Content Delivery Network.
- To propose an algorithm that offers more efficiency in detecting intrusion in the system.
- To use simulation to apply our proposed model.
- To compare results with existing system.

II. Literature Overview

A. Background

A firewall is designed to keep the attackers away from the network, while honeypots are designed to trick the hackers into attacking the system. CDN systems are very useful to control and optimize all network services that find the shortest distance for its users . On the other hand, servers can sometimes be a prime target for DoS or DDoS attacks in such architectures.

Several protective models with honeypots have been suggested for DDoS.

Tiwari, R., Jain, A.[4] clarified honeypot innovation and its weaknesses in the honeypot framework and proposes an appropriated framework which cure the current insufficiency in the centralized control system to enhance network security. A.Zakari and A.A.Lawan.[5] they identified three requirement for using honeypot that is deceptiveness, robustness and intelligence.

Y. Yang, H.Yang, and J.Mi [6] mentioned the distributed deployment for honeypot system. It widens the network area and mark the package to identify real attacks sources and has certain expansion ability. Hrishikesh Arun Deshpande [7] contributed by using virtual honeypot in preventing distributed denial of service attacks.

Nishat, Doh and Chae [8] suggested a technique against DDoS attacks that uses SDN-based CDNi networks to safely provide services with a multi-defense strategy. They combine ALTO servers to track a very large network to provide a bird's eye view. It also mark path map in the ALTO server to trace the request packets and has a protection switch to protect the main servers. But they have not mention how the legitimate user will get more efficient service.

Concerning DoS assaults on CDN, I might want to feature the research embraced and the thoughts proposed to manage this appalling attack. Resmi, An and Chezian, R.M.[9] proposes interruption prevention and discovery in Content Delivery Networks. There has additionally been research beforehand attempted into the utilization of honeypots to relieve DoS attack in IoT frameworks[10]. Though at the present point, this examination stays theoretic with just theoretic models being proposed.

To diminish DDoS attacks Sherif M. Khattab et al [11] proposed roaming honeypots. Here a pseudo-random schedule changes the location of the current active server and the honeypot, shared between the servers and the authorized client. Therefore legitimate clients always send their service request to active servers whereas attack request may reach the honeypot. The source address of any request that hits a honeypot is black listed so that all future requests from this source are subsequently dropped.

Quang Duy La et al [12] proposes game theoretic model in which both the attacker and defender endeavor to mislead one another. The defender utilizes honeypots to mislead and trap the aggressors. This research is an expnasion of the work done by N. Garg et al [13] that examines the viability of utilizing honeypots for double dealing in systems following an game theoretic methodology.

As made reference to beforehand these works have been limited to theoretic models. With this in mind, we deploy a honeypot based security system for a CDN in this paper, to block DDoS attacks from malicious attackers and also collect information on the attacker so that future attacks might be prevented.

III. Distributed Honeypot System

A. Honeypot in CDN

Honeypot is the second layer in our CDN security. A honeypot's most huge esteem depends on the data that it acquires and can quickly caution on. Information that enters and leaves a honeypot permits security staff to assemble data that isn't accessible from an IPS. An aggressor's keystrokes can be logged amid a session, regardless of whether encryption was utilized to set up it. Additionally, any endeavors to get to the system can trigger quick alerts. Sometimes many attackers can get into the main server without being detected by the Intrusion Prevention System(IPS). At that time they cannot really find the real server rather they will be observed by our decoy server honeypot that analyses their motive and blocks them if they fall in the danger zone . Otherwise if the analysis found to be safe then they can access the CDN server. The system flowchart diagram is shown in **Figure 2**.

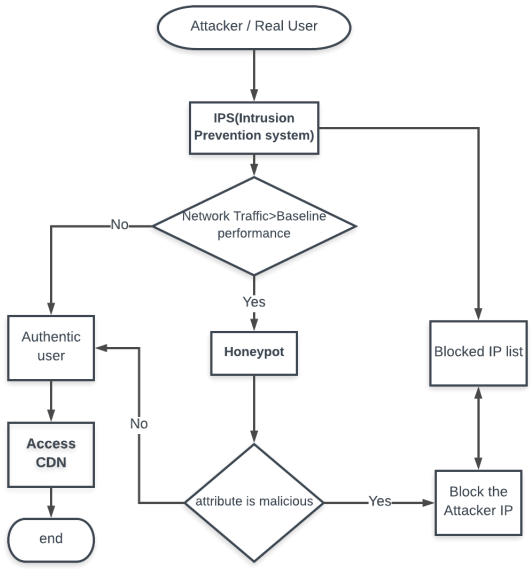


Fig. 2. System flowchart

As honeypot presents honey data as a decoy for the attackers, it is used for both intrusion detection and prevention purposes.

a) Intrusion Detection System: Intrusion Detection System (IDS) causes data frameworks to manage attacks. This is refined by gathering data from an variety of frameworks and system sources. The data gathered is broke down for conceivable security issues. An IDS assembles and breaks down data from different zones inside a system to distinguish conceivable security break [14].

b) Intrusion Prevention System: In our proposed system, we have maintained an intrusion prevention system (IPS) which is a cyber attack prevention technology that analyses the flow of network traffic to detect and prevent vulnerability. Vulnerability feats usually take the form of malicious inputs

into a target application or service used by attackers to disrupt and control an application or machine. Following a successful exploit, the attacker can slow down the target application (resulting in a denial-of-service state).

In IPS, we detect statistical anomaly samples of network traffic at random and compares them to a pre-calculated baseline performance level. When the sample of network traffic activity is outside the parameters of baseline performance, the IPS takes action to handle the situation. It sends these to the next part of our system honeypot which will analyze and find if it is false positive .

B. Distributed Honeypot in CDN

Content Delivery Network is the largest distributed server for smooth content sharing. It is distributed geographically along with proxy server and backup database to provide end-users high availability and high performance. Distributed servers are kept for distributing contents e.g. a University campus where students access to their personal data and class material resources through it. Therefore, assuming that there is a central or primary server that hosts the primary Honeypot and the network traffic attributes database, all other servers will host secondary honeypots as the initial target of attackers. As a result the edge servers will act as the monitoring agent that continuously monitor the data and provide feedback to the central control system through secure channel.

Following is the system architecture of our distributed server in **Figure 3**.

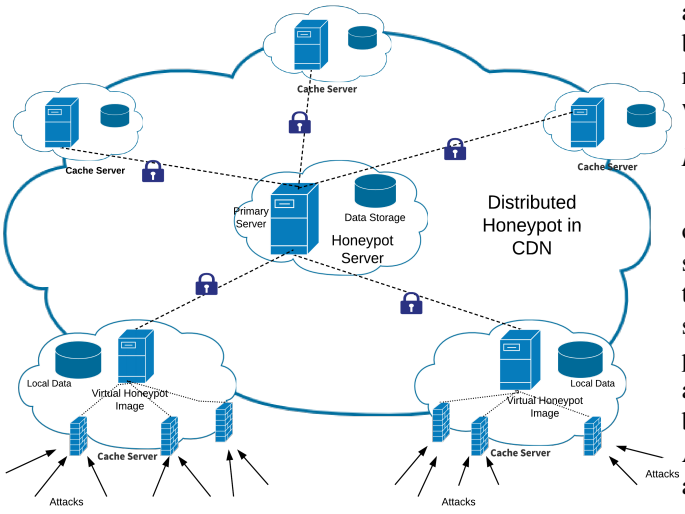


Fig. 3. System Architecture of distributed virtual Honeypot in CDN

Distributed honeypot based systems have two layer architecture in general as: Control agent and Monitoring agent. Details regarding the agents and secure communication management are provided in the following:

a) Control agent: The primary server acts as the control agent in CDN keeping the primary log file synced with other local log files residing in edge servers. The honeypot hosted by the primary server works in the second layer keeping the monitoring agents in the first layer of malicious attacks.

b) Monitoring agents: All the honeypots residing in the edge servers act like the monitoring agents. As users will be initially granted access to resources in the edge caching servers, the malicious user have the same access mechanism. Therefore, monitoring agents are the first layer to face any sort of intrusion or attacks.

c) Secure communication between agents: The communication between the control agent and the monitoring agents is secured using secure socket layer (SSL). If one monitoring agent reports anomalies, then it will be synced with the primary honeypot hosted by control agent. Later the same report will be synced with all other edge servers for further rule changing purpose in firewalls.

C. Virtual Honeypot using image

A virtual honeypot is a software that imitates an endangered system to draw in intruders and concentrate their behaviors. Honeypot proprietors can assemble data about attackers and their activities that can help recognize organize vulnerabilities and take activities to secure feeble focuses. The data can likewise be utilized to figure out what IP locations and requests ought to be blocked. Virtual honeypots are hosted in virtual machines (VMs) which may exist in numerous arrangements on a solitary PC or apparatus to copy different frameworks and vulnerabilities. As honeypot is comparatively costly and in a distributed framework like CDN it will increase the total cost for maintenance of the system, virtual honeypots seem to be a better solution and easier to implement. Virtual honeypots are less expensive to convey and more secure than equipment based systems. As Content delivery Network comprises of numerous intermediary topographically disseminated server so we utilized virtual honeypot so that attack can be halted.

D. Cache Servers

Cache server or proxy server is the distributed server of Content delivery Network because of its smooth content sharing. Due to dispersed nature of cache server it is vulnerable to various attacks. Actually it is the copy of the primary server which resides in edge servers. Therefore, if any attack is performed in any location then its regional users will be greatly affected. In our architecture, every cache server is protected by Virtual honeypots hosted in virtual machine containers. Also these honeypots will act as monitoring agent which will analyze the traffic and make decision whether or not a request will gain proper access to cached contents. As a result the actual user will get instant and reliable service and malicious intruder will be detected at once.

Now an overall in-detailed data flow diagram of our architecture is shown in **Figure 4**.

E. Algorithms

Initially user request will come at the monitoring agent which is the edge server. Then they will generate local log if it

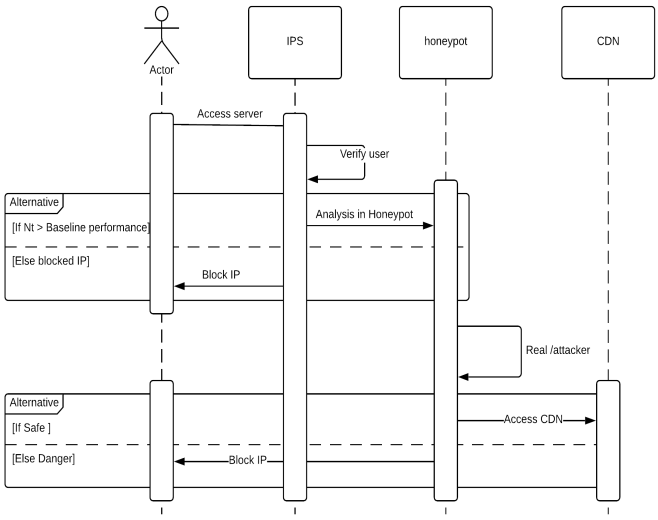


Fig. 4. Data Flow Diagram

is authentic user or attacker. Afterwards this information will be synced with primary server [Control Agent]. The complete process is presented in **Algorithm 1**.

Algorithm 1: Algorithm for attack mitigation using Honeypot

```
System Initialization;
while TRUE do
    Real user or attackers send request to access CDN;
    /* AT Edge Server                                  */
    if IP in Blocked list of Local Log then
        Reject Access;
    end
    if Network traffic > Baseline Performance[depends
     on the system] then
        Send Anomaly request to Monitoring
         Agent[Honeypot] for Analysis;
    end
    /* AT Monitoring Agent                             */
    if Analysis found safe then
        Grant access to use CDN;
    end
    else
        IP stored to the blocked list in Local Log;
        Updated List is synced with The Control Agent;
    end
end
```

IV. Implementation and Evaluation

A. DDoS Attack Simulation

In this test we will conduct trials using Ubuntu operating system and Python script for deploying attacks on our server, this attack testing will conduct a direct attack to the server so that in the presence of this test will go into the trap honeypot. The following **Figure 5** is demonstrate the script is running and it will flood our bandwidth with unnecessary bad traffic.

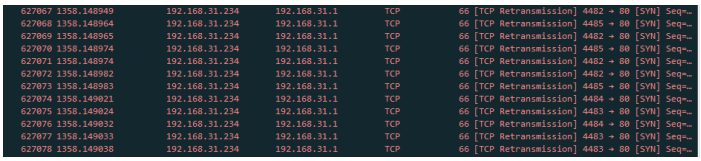


Fig. 5. Python Script for DoS attack

The **TABLE I** shows that we will make a real server that will act as a CDN server with port 80 and another IP address will be an intruders address from where we will perform attack and will make the system unreachable to the users.

TABLE I
Attacker IP and Server IP

IP Address	User Computer	Operating sytem	PORT
192.168.31.1	Real Server	Ubuntu	80
192.168.31.234	Client Intruder	Ubuntu	80

1) Normal Traffic: Network traffic or data traffic is the amount of data moving across a network at a given point of time. At first we measured the normal traffic movements **Figure 6** in our system and analyzed different parameter values.

TABLE II
Statistics of Normal Traffic

Measurement	**Captured**	**Displayed**
Packets	15196	15196 (100.0%)
Time span, s	58.391	58.391
Average pps	260.2	260.2
Average packet size, B	821	821
Bytes	12472428	12472428 (100.0%)
Average bytes/s	213 k	213 k
Average bits/s	1,708 k	1,708 k

We have captured the traffic for almost 60 seconds and found the statistics given at **TABLE II**. Here the average bytes/s was 213k and average packet/s was 260.20.

2) DDoS Attack Traffic: Not at all like other cyber criminal intrusions, DoS and DDoS attack don't attempt to penetrate a system rather they are regularly part of a bigger attack. They try to create force so that bandwidth allocated to the real user are used up by fake request **Figure 7** and create panic towards the users. Here is a statistics of the DDoS attack traffic performed on our system in **TABLE III**.

From the above table we found that here the average packet/sec was 3223 and the average byte 1154k which is far more than normal traffic.

The comparison between two traffic given in the following **Figure 8**.

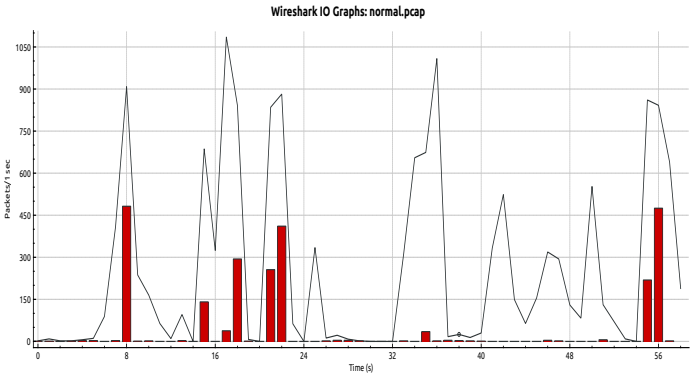


Fig. 6. Normal Traffic in server

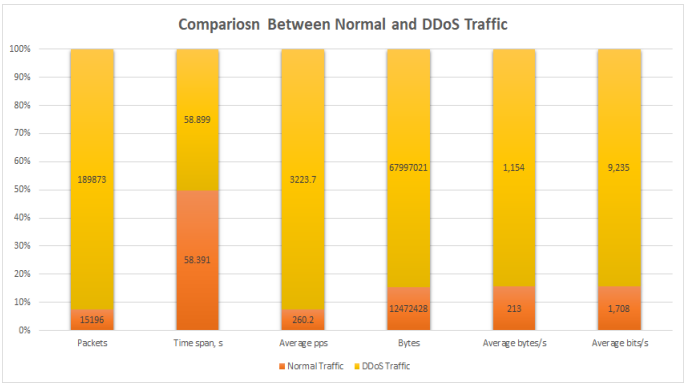


Fig. 8. Comparision between DDoS and Normal Traffic

TABLE III
STATISTICS OF DDOS TRAFFIC

Measurement	Captured	Displayed
Packets	189873	189873 (100.0%)
Time span, s	58.899	58.899
Average pps	3223.7	3223.7
Average packet size, B	358	358
Bytes	67997021	67997021 (100.0%)
Average bytes/s	1,154 k	1,154 k
Average bits/s	9,235 k	9,235 k

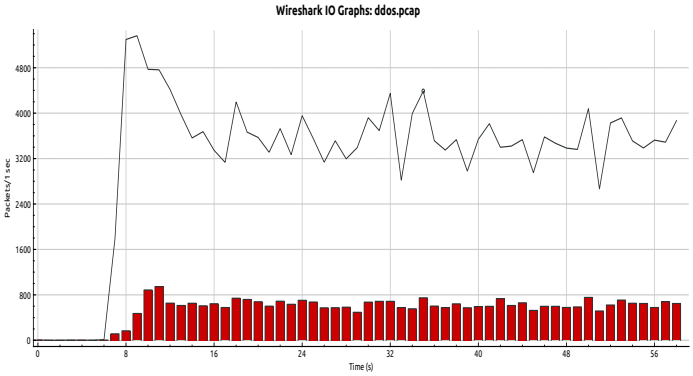


Fig. 7. DDoS Traffic in server

Now we will add honeypot and make a port open so that intruder try to find that way and want to perform their task. When any malicious attacker will find that open port and want to do something then our honeypot will show his IP address and we can block that IP address.

B. Intrusion Detection by Honeypot

In our Distributed Virtual honeypot model every edge server contains monitoring agent(Honeypot) that analyzes traffic if it is malicious and sent by the IPS then after analyzing, it will be added to the blocked IP list or sent to the server for their service .

This is the interface **Figure 9** of honeypot being deployed to our server. It is a Ruby script from where we can select various tool.



Fig. 9. Honeypot tool

Now we will go to the network tool by selecting second option. We will see that different network tools are on the list. From these we will choose honeypot . After choosing honeypot we will find two option for auto configuration and for manual configuration . We will go for manual configuration and give our preferred options like :

- Port number
- Any message want to show
- Log file generated by intrusion
- Alert option for intrusion

After enabling honeypot, we will try to DoS attack attempt on our server. We will use python script that creates many requests, that will try to flood the server.

Next for each intrusion detected by our honeypot we will get intruders information **Figure 10**. We will save the IP address in the blocked list of monitoring agent and later on this will be synced with the control agent of our main server.

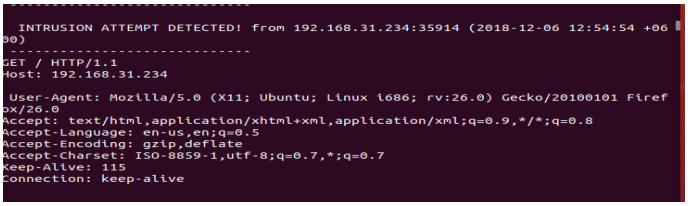


Fig. 10. Attacker IP address Detected by Honeypot

This blocked list will be helpful for detecting the blacklisted IP at the first place by the IPS and it will smooth our CDN services.

V. Discussion

A. Significance and Application

Presently, the most difficult issue is to keep up the unwavering quality and speed of a system. Content Delivery Network(CDN) is worked for quick service and smooth conveyance of assets to its clients. Sometimes its asset conveyance is being hampered by different Denial of Service attacks which spent the data bandwidth of the clients and make them sit tight for their services. As a result we implemented a distributed measure that will mitigate these DDoS attacks and serve the user in their best way. We have used two levels of security for better intrusion detection. And virtual honeypot adds numerous advantages, mostly they can essentially enhance our security pose along with IPS in CDN.

Unlike IPS, IDS can not block attacks rather it can give warning to the administration. Therefore, this changes the situation when we find the blocked IP wanted to access the contents then it will be rejected at the first attempt. If somehow any false positive IP were found then it will be analyzed by honeypot and it will take necessary action. Reduced cost is another advantage while implementing the model. As the edge servers are hosted in virtual machine containers, it will reduce the total cost tremendously which is much less than implementing traditional honeypot system in CDN.

B. Scope and Limitations

In our architecture we have taken a baseline performance that determines the traffic characteristics and then we give access to our main content server. The system may find limitations as it needs to study different baseline performance according to various networks.

C. Future Work

Future work will be maximizing more security by making our honeypot adaptive that is able classify the intrusion and anomaly patterns and modify firewall rules accordingly. Adaptive systems provide quicker response to malicious attacks and severe risks including zero day attack.

VI. Conclusion

The consistent advancement of DDoS attacks have rendered them one of the best weight to the present Internet transmission capacity. In spite of the fact that variety of DDoS attack prevention, mitigation, and traceback techniques methods have been proposed by analysts in the most recent decades, their relative take-up has been negligible at mammoth. In this paper, a mitigation strategy is exhibited for counteractive action of DDoS attacks on a CDN server by Distributed Virtual Honeypot while the proposed system does not influence the activity from genuine users, it effectively blocks traffic from the attack sources by watching their network traffic and sends those for more assurance to the honeypot where it analyze the packet if they are false positive or not. Thus our technique will maintain the real users bandwidth and blocks the false users. Moreover, the utilization of virtual honeypot could be reached out for different kinds of attacks as a research based honeypot to gather points of interest.

Acknowledgement

The research work was partially supported by Green University of Bangladesh research fund.

References

[1] C. Douligeris and A. Mitrokotsa, "DDoS attacks and defense mechanisms: A classification," *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2003*, vol. 44, pp. 190–193, 2003.

[2] V. V. Das, "Honeypot scheme for distributed denial-of-service attack," *Proceedings - International Conference on Advanced Computer Control, ICACC 2009*, pp. 497–501, 2009.

[3] I. Sembiring, "Implementation of honeypot to detect and prevent distributed denial of service attack," in *2016 3rd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*. Semarang, Indonesia: IEEE, 2016, pp. 345–350.

[4] R. Tiwari, "Design and Analysis of Distributed Honeypot System," *International Journal of Computer Aplications (0975-8887)*, vol. 55, no. 13, pp. 20–23, 2012.

[5] A. Zakari, A. A. Lawan, and G. Bekaroo, "Towards improving the security of low-interaction honeypots: Insights from a comparative analysis," in *Emerging Trends in Electrical, Electronic and Communications Engineering*, P. Fleming, N. Vyas, S. Sanei, and K. Deb, Eds. Cham: Springer International Publishing, 2017, pp. 314–321.

[6] Y. Yang, H. Yang, and J. Mi, "Design of distributed honeypot system based on intrusion tracking," *2011 IEEE 3rd International Conference on Communication Software and Networks, ICCSN 2011*, pp. 196–198, 2011.

[7] Hrishikesh Arun Deshpande, "HoneyMesh: Preventing Distributed Denial of Service Attacks using Virtualized Honeypots," *International Journal of Engineering Research and*, vol. V4, no. 08, pp. 263–267, 2015. [Online]. Available: http://www.ijert.org/view-pdf/13877/honeymesh-preventing-distributed-denial-of-service-attacks-using-virtualized-honeypots

[8] N. I. Mowla, I. Doh, and K. Chae, "Multi-defense Mechanism against DDoS in SDN Based CDNi," in *2014 Eighth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*. Birmingham, United Kingdom: IEEE, Jul. 2014, pp. 447–451.

[9] A. Resmi and R. M. Chezian, "An extension of intrusion prevention, detection and response system for secure content delivery networks," in *2016 IEEE International Conference on Advances in Computer Applications (ICACA)*. Coimbatore, India: IEEE, Oct. 2016, pp. 144–149.

[10] M. Anirudh, S. A. Thileeban, and D. J. Nallathambi, "Use of honeypots for mitigating DoS attacks targeted on IoT networks," in *2017 International Conference on Computer, Communication and Signal Processing (ICCCSP)*. Chennai, India: IEEE, Jan. 2017, pp. 1–4.

[11] S. Khattab, C. Sangpachatanaruk, D. Mosse, R. Melhem, and T. Znati, "Roaming honeypots for mitigating service-level denial-of-service attacks," in *24th International Conference on Distributed Computing Systems, 2004. Proceedings.* Tokyo, Japan: IEEE, 2004, pp. 328–337.

[12] Q. D. La, T. Q. Quek, J. Lee, S. Jin, and H. Zhu, "Deceptive Attack and Defense Game in Honeypot-Enabled Networks for the Internet of Things," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1025–1035, 2016.

[13] N. Garg and D. Grosu, "Deception in honeynets: A game-theoretic analysis," *Proceedings of the 2007 IEEE Workshop on Information Assurance, IAW*, no. June, pp. 107–113, 2007.

[14] T. Kaur, V. Malhotra, and D. D. Singh, "Comparison of network security tools- Firewall, Intrusion Detection System and Honeypot," *International Journal of Enhanced Research in Science Technology & Engineering*, vol. 3, no. 2, pp. 200–204, Feb. 2014.

Merging the communication protocols *6LoWPAN-CoAP and RPL-CoAP*

Simulation and performance analysis using Cooja simulator

Arif Mahmud
Dept. of CSE
Daffodil International University
Dhaka,Bangladesh
arif.cse@diu.edu.bd

Faria Hossain
Dept. of CSE
Daffodil International University
Dhaka,Bangladesh
faria.cse@diu.edu.bd

Faija Juhin
Dept. of CSE
Daffodil International University
Dhaka,Bangladesh
faija15-238@diu.edu.bd

Tasnim Ara Choity
Dept. of CSE
Daffodil International University
Dhaka,Bangladesh
tasnim15-245@diu.edu.bd

***Abstract*— IoT can prolong the internet communication to heterogeneous types of objects those are used within embedded technology to join with the surrounding through the active support of internet technology. The exploration and application of the Internet of things technology is certain to quicken the industrial promotion and revolution, at the same time, to securely promote the growth of the national economy, to constantly improve the comprehensive national power. A number of research works is going on these IoT nodes which consume limited resources and the development of communication protocols can be considered as the vital one. This paper is concentrated on finding out the best protocol for communication between IOT nodes after merging these three protocols RPL, CoAP and 6LoWPAN depending on different criteria and, therefore, the study can be useful in personal and home application, healthcare, utilities and services, enterprise application and industrial automation. Based on the results, it can be stated that the combination of RPL and CoAP performs better than the combination of 6LoWPAN and CoAP in terms of received packets for varying number of nodes, simulation time and communication range.**

Keywords— CoAP, Cooja, IoT, RPL, 6LoWPAN

I. INTRODUCTION

The amount of connected entities has already gone beyond the total amount of human being few years ago and it is predicted to reach 50 billion by the end of 2025 [1]. As more and more devices are being introduced in our regular life, has created immense pressure on internet connectivity in term of storage and routing. These devices are constantly communicating with each other and also with the server through the internet. Therefore, the necessities of application specific communication protocols have risen up and will continue too in future days to come.

Communication protocols are similar to traffic directors that maintain organized communication with each other. Protocol has the responsibilities to develop, maintain communication and without protocols, the nodes of various networks might get connected but won't be capable to interact with each other. Therefore, if one machine wants to receive or send data to another machine, both of them should use similar protocol for completing the task [2]. As a result, choosing an appropriate communication protocol is one of the vital concerns now a day.

This paper is concentrated on finding out the best protocol for communication between IOT nodes after merging these three protocols RPL, CoAP and 6LoWPAN depending on different criteria and, therefore, the study can be useful in personal and home application, healthcare, utilities and

services, enterprise application and industrial automation. Hence, the paper is schematized as the following methods:Section II describes IoT and Communication protocols;Section III illustrates Literature review; Section IV defines the sample communication scenario using these merged protocols; Section V explains results based on comparison among these protocols; and the paper is concluded in section VI.

II. IOT AND COMMUNICATION PROTOCOLS

Internet of Things (IoT) denotes the dynamic connectivity of physical devices with limited resources with the support of internet infrastructure. It helps to develop interaction between these entities and also every internet supported objects and networks [3]. IoT can prolong internet communication to heterogenous types of objects those are used within embedded technology to join with the surrounding through the active support of internet technology [4]

Based on the necessities and regular application IoT has already been included in the national five-year development plan outline of China [5]. The exploration and application of the Internet of things technology is certain to quicken the industrial promotion and revolution, at the same time, to securely promote the growth of the national economy, to constantly improve the comprehensive national power [6].

A number of research works is going on these IoT nodes which consume limited resources. Development of communication protocols can be considered as the vital one. At present, three communication protocols have become popular and therefore, we have chosen them to simulate using cooja simulator. The basic features of protocols are explained below in brief:

The 6LoWPAN concept originated from the idea that the Internet Protocol could be applied to the smallest devices, [8] and those low-power devices with limited processing capabilities. The standard has the liberty of choosing frequency band and the elasticity to complete over several communications stages, counting Ethernet, Wi-Fi and 6LoWPAN protocol. This protocol plays a most imperative role in IoT wireless communication as it stands for IPv6. It is demonstrated by auxiliary addresses with diverse lengths, low bandwidth, star and mesh topologies, battery supplied devices, low cost, large number of devices, unknown node positions, high unreliability, and long idle periods during the communications [7].

RPL is one of the distance vector routing protocols for network layer and it is developed to function with little energy and lossy systems using IPv6. Besides, message confidentiality and reliable message delivery are cared by RPL. It has been pre-arranged in such a technique that link layer appliances can be castoff when available and appropriate; yet, in their absence, RPL can custom its own mechanisms[8]. RPL is popular for many reasons, for example, it can develop its own route very fast, share the routing information to other nodes and have the capability to get adapted dynamically with the network topology [9].

CoAP is considered as one of the most recent application layer protocols founded by IETF to be utilized for intelligent objects. In addition, it is a trivial protocol for low resource consumed smart devices and can be embedded into buildings, vehicles. It is also considered as the substitution of HTTP protocol to be used in application layer of IoT devices [10]. As a result, the data reception from the sensor has become more convenient. In addition, this protocol has the multicast feature, a little overhead though and very effective in M2M interaction. CoAP nodes are divided into two types namely server and client from the architecture viewpoint. Server nodes are installed in sensors while client nodes are embedded into controller [11].

III. LITERATURE REVIEW

Currently, a number of research works is going on these communication protocols. Different types of simulators are being used for different applications. The papers [12] studied the performance of COAP protocols for different platforms like Contiki, TinyOS, POSIX, lwIP etc. C, JAVA, JAVA script and python were implemented as programming languages and several library functions were used, for example, ibcoap, FreeCoAP, smcp, CoAPthon, Californium etc. [13] used the COAP protocol for healthcare monitoring application. The paper [14] made the comparison between REST and SOAP architecture using the previously described three protocols to be used for building management system. Besides, the benefits of BMS were explained along with the methodologies that are needed, were analyzed. [15] also worked with the development of protocol stack.

This paper [16] described the strong points and limitations of RPL protocol. It also pointed out the vulnerability and proposed an intrusion detection technique in support of it to reduce the attack. The paper [17] found out the importance of RPL, COAP and 6Lowpan as the communication protocol for IOT to be used to develop smart city. The paper [18] proposed a new model for cross layer network infrastructure on NS-3 platform. It also addressed the congestion problem and solved with back pressure

algorithm. The paper [19] studied the three protocols on the development of sensor networks and their viability was evaluated for industrial purpose. However, no much study has been attempted on cooja simulators or on contiki nodes. Most importantly, these protocols were not merged, simulated and our study can be considered as the newest one.

IV. COMMUNICATION SCENARIO

In this chapter, the communication between the IOT nodes will be illustrated using RPL, CoAP and 6LoWPAN protocols. Besides, these protocols will be merged and their network architecture will be briefly explained.

IPv6 Low power Wireless Personal Area Network (6LoWPAN) is functional on adaptation layer that allows the consumption of IPv6 over low power wireless and User Datagram Protocol (UDP) header compression. As seen in Fig. 1, the 6LoWPAN stack is both highly scalable and reliable. The modular structure of the design makes possible for ARM to accommodate most of the requirements. The combination of 6LoWPAN stack and 6LoWPAN border router Access Point (AP) enables developers to use the extremely flexible and multipurpose mesh communication for most applications.

As in Fig.2, DIO carries information that allows a node to discover an RPL Instance, learn its configuration parameters and select DODAG parents. A DIS solicits a DODAG Information Object from an RPL node. Destination Advertisement Object (DAO) propagates destination information upwards along the DODAG.

As shown in figure 5, CoAP cooperative model is related to HTTP's client/server model and CoAP employs a two layers structure. The bottom layer is Message layer that has been designed to deal with UDP and asynchronous switching. The request/response layer concerns communication method and deal with request-response message [10].

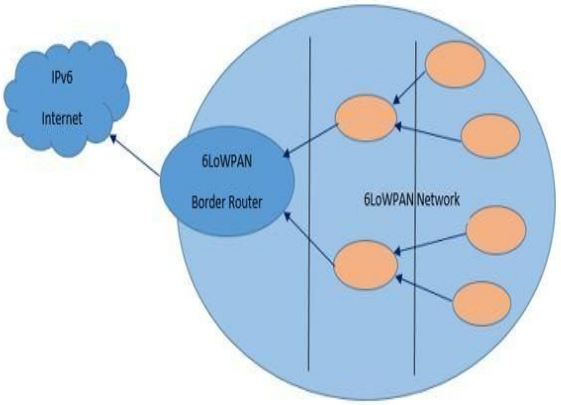


Fig. 1. Network architecture of 6LoWPAN [20]

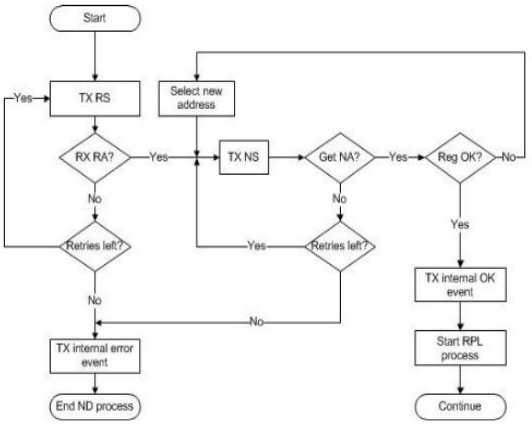


Fig. 2. Functional systen of RPL [10]



Fig. 3. Network Architecture of RPL [10]

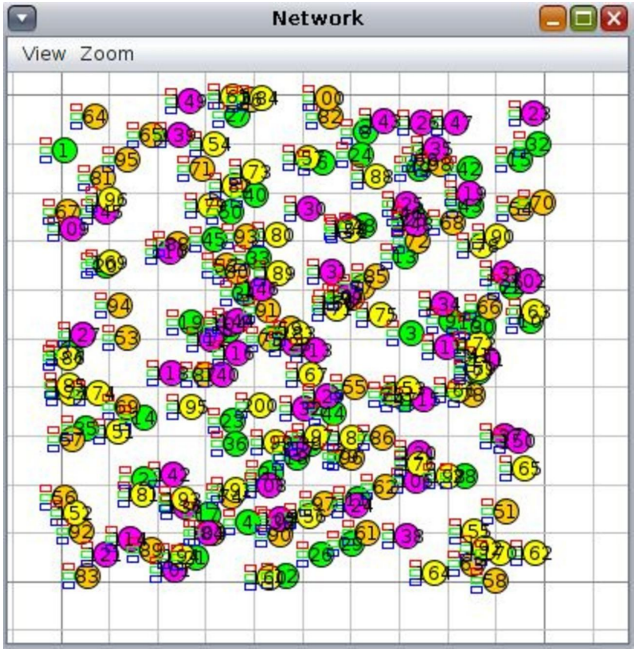


Fig. 4. 6LoWPAN and CoAP are communicating on cooja platform

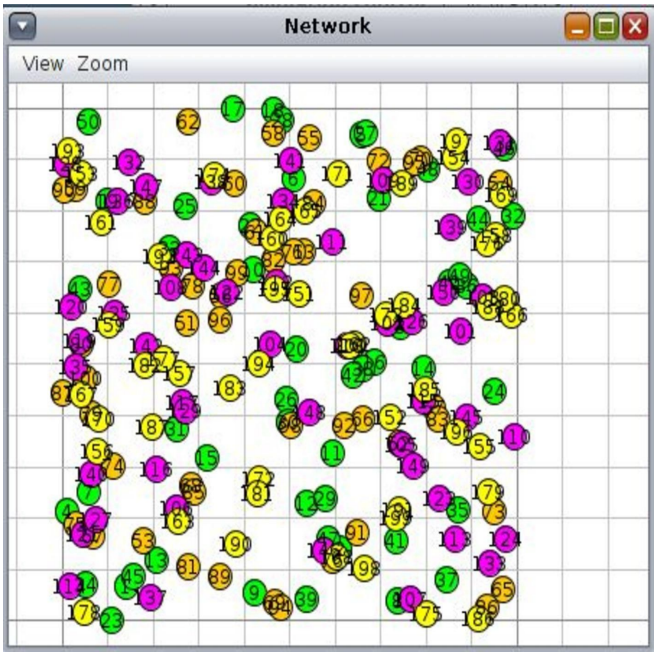


Fig. 5. RPL and CoAP are communicating on cooja platform

As can be seen from Fig. 4,the server and the client nodes are communicating with each other using 6LowPAN and CoAP protocols together.Here,Purple and yellow nodes are UDP client and server under 6LoWPAN protocol and green,orange nodes are UDP client and server under CoAP protocol.

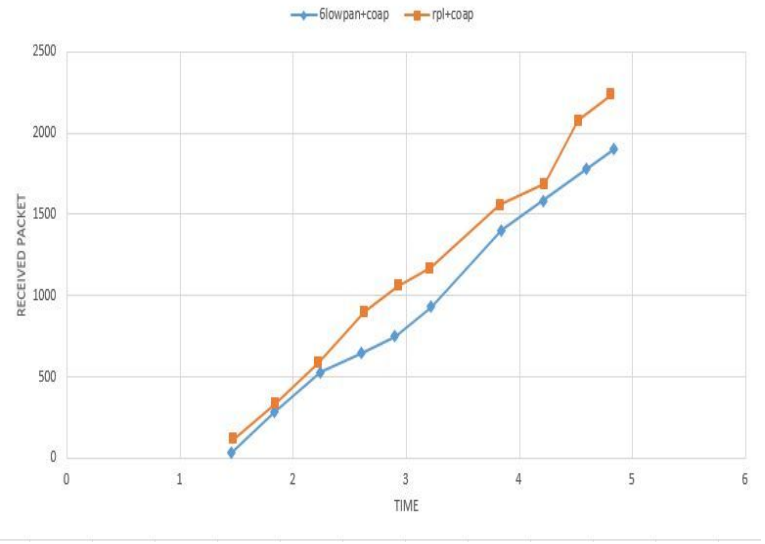


Fig. 6. Comparison of received packet for 50 nodes

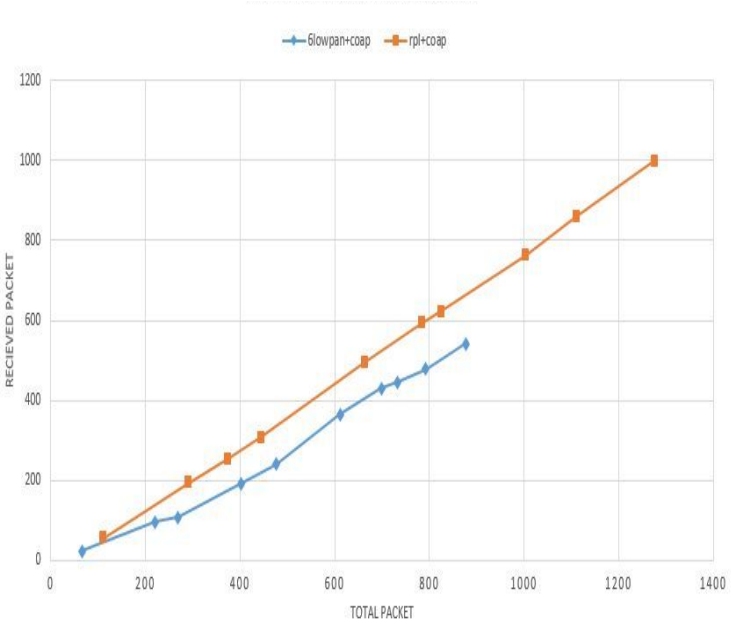


Fig. 7. Comparison of received packet for 70 nodes

As shown in Fig. 5, the server and the client nodes are communicating with each other using RPL and CoAP protocols together.Here,Purple and yellow nodes are UDP client and server under RPL protocol and green,orange nodes are UDP client and server under RPL protocol.

V. IMPLEMENTATION AND RESULT

The performance of merged (6LoWPAN and CoAP, RPL and CoAP) protocols will be compared depending on different criteria like received packet, number of nodes, time simulation and number of IOT nodes on cooja platform. Cooja is chosen to simulate contiki nodes which consume low resources like energy, bandwidth to communicate with each other on Ubuntu operating system. At present, this simulator is very popular among the network researchers since accurate behavior of the system can be examined and simulations of large networks are possible before embedded them into devices [21].

As seen in Fig. 5 ,the X-axis indicates the transmission time and Y-axis indicates number of received packets.The number of nodes has been taken as 50 for this simulation.It can be observed that 6LoWPAN and CoAP combination are bit slower than RPL and CoAP combination.For example, in 4.8 minutes 6LoWPAN and CoAP combination received 1900 packets out of 1977.However,at the same time RPL and CoAP combination received 2236 packets out of 2326.

As in Fig. 7, the X-axis indicates the total packets and Y-axis indicates number of received packets the number of nodes has been taken 20. 6LoWPAN and CoAP combination received 564 packets out of 877 and on the contrary RPL and CoAP combination receives 998 packets out of 1224.However 6LoWPAN and CoAP combination are slower than RPL and CoAP combination.

In Fig. 8, the X-axis indicates the transmission time and Y-axis indicates number of received packets.RPL-CoAP can receive more packets than 6LoWPAN-CoAP combination. In certain periods of time(12.51 minutes for both case) 6LoWPAN and CoAP combination received 641 packets on the other hand RPL and CoAP combination receives 1046 packet.

As seen in Fig. 9, Transmission range has been expanded from 10 meter to 60 meters.It can be seen from Fig. 6 that transmission rate of 6LoWPAN and CoAP is very slow. Noticed that, in 12.31 minute 6LoWPAN and CoAP combination received 841 packets.On the contrary RPL and CoAP combination received 3,545 packets for the same amount of time.

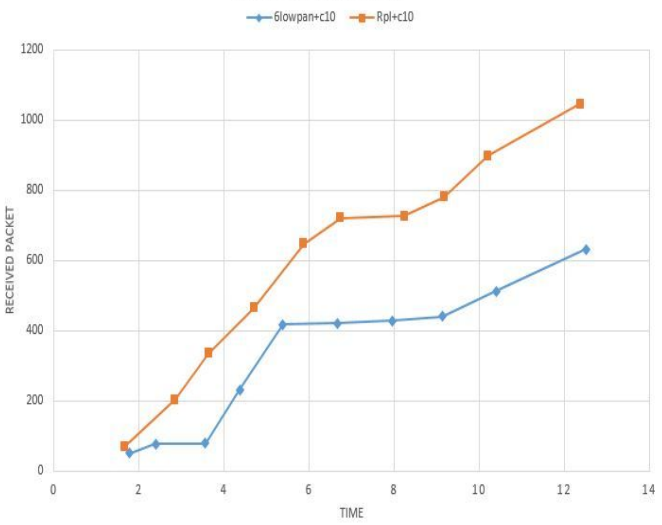


Fig. 8. Comparison of received packet for 10 meter range

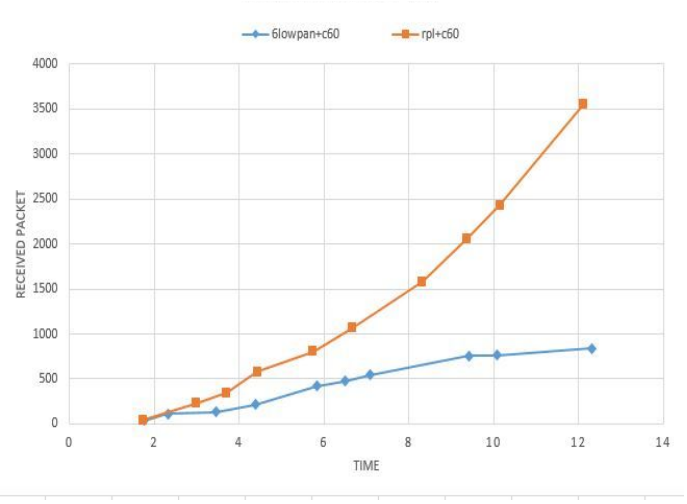


Fig. 9. Comparison of received packet for 60 meter range

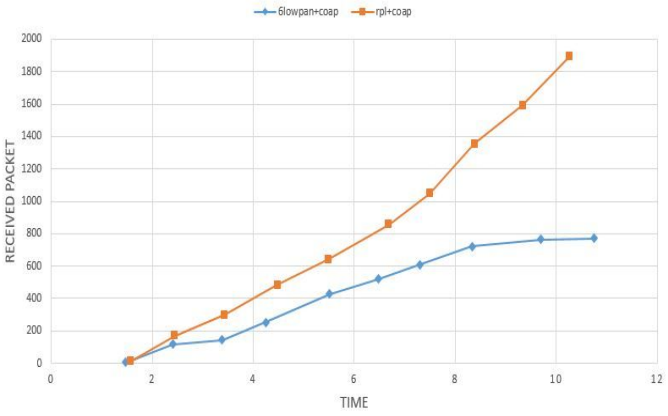


Fig. 10. Comparison of received packet for 100 meter range

Communication range has been changed in Fig. 10 for 100 meters.It can be find that in 10.74 minutes, 6LoWPAN and CoAP combination receives 771 packets but at the same time RPL and CoAP combination receives 1894 packets.

VI. Discussion

IoT is the concept in which the virtual world of information technology mixes flawlessly with the real world of things, it is one of the important topics in the current networking research.Since small devices which are embedded into IoT, are connected via the internet, consumption of power, memory, energy, routing, bandwidth, etc. are the vital issues.In this paper, mainly focused on investigating a better solution of finding out the best communication among these merged protocols 6LoWPAN and CoAP, RPL and CoAP.

In previous paper [], IoT protocol stack along with different protocols used at different layers have been studied for the efficient communication among IoT nodes through the Internet and focused on investigating a solution of finding out the best communication among three protocols 6LoWPAN, RPL and CoAP individually.But in this paper, find out the best communication after merging these protocols 6LoWPAN and CoAP, RPL and CoAP.

In this paper, the performance of merged protocols were examined depending on different criteria like different transmission range, different transmission time with number of data packets. After the analysis, it can be stated that packet transfer rate of 6LoWPAN and CoAP combination is slow and the combination of RPL and CoAP is more efficient in terms of packet reception. We are also hopeful to implement this protocol in real testbed and not so much deviations are expected though.

References

[1] Plamen Nedeltchev, “The Internet of Everything is the New Economy”,COSCO,2015.

[2] Elis Panaseyko, “Internet Technologies And Communication Protocols”,2018

[3] P.P.Ray,”A survey on Internet of Things architectures”,Journal of King Saud University, 2018, pp. 291-319

[4] Forrest Stroud, “Iot-Internet of things. Lecture Notes in Iot”, webopedia, 2018.

[5] Saber Talari , Miadreza Shafie-khah, Pierluigi Siano , Vincenzo Loia, Aurelio Tommasett and João P. S. Catalão:,”A Review of Smart Cities Based on the Internet of Things Concept”. MDPI,open access journals, 2017.

[6] Jean-Marc Frangos,”The Internet of Things will power the Fourth Industrial Revolution Annual Meeting of the New Champions”,world economic forum, 2017.

[7] Ziyad Khalaf Farej,Ali Maher Abdul-Hameed,”Performance Comparison among (Star, Tree and Mesh) Topologies for Large Scale WSN based IEEE 802.15.4 Standard”,International Journal of Computer Applications, 2015.

[8] JP Vasseur,Navneet Agarwal,Jonathan Hui,Zach Shelby,Paul Bertrand,Cedric Chauvenet, “RPL: The IP routing protocol designed for low power and lossy networks,Internet Protocol for Smart Objects (IPSO) “, Alliance, 2011.

[9] Muhmmad Rashid Khan, “Performance and route stability analysis of RPL protocol”, Masters' Degree Project Stockholm, Sweden, 2012.

[10] Xi Chen, “Constrained Application Protocol for Internet of Things”, Research paper on computer science,Washington University, 2012.

[11] Girum Ketema Teklemariam,Jeroen Hoebeke,Ingrid Moerman,,Piet Demeester, “Facilitating the creation of IoT applications through conditional observations in CoAP”, EURASIP Journal on Wireless Communications and Networking, 2013.

[12] Markel Iglesias-Urkia, Adrian Orive, Aitor Urbieta,”Analysis of CoAP Implementations for Industrial Internet of Things: A Survey”,The 8th International Conference on Ambient Systems, Networks and Technologies, 2017.

[13] Dejana Ugrenovic, Gordana Gardasevic,”CoAP protocol for Web-based monitoring in IoT healthcare applications”,23rd Telecommunications forum TELFOR 2015 Serbia, Belgrade, November 24-26, 2015.

[14] Ngoc-Thanh Dinh and Younghan Kim,”RESTful Architecture of Wireless Sensor Network for Building Management System”,KSII transaction on internet and information systems vol. 6, no. 1, Jan 2012

[15] Thomas Watteyne, Xavier Vilajosana, Branko Kerkez, Fabien Chraim1, Kevin Weekly, Qin Wang, Steven Glaser and Kris Pister1,”OpenWSN: a standards-based low-power wireless development environment”,Transaction of emerging telecommunications technoligies,Trans. Emerging Tel. Tech. 2012; 23:480–493

[16] Linus Wallgren,1 Shahid Raza,1 and Thiemo Voigt1,2,”Research Article Routing Attacks and Countermeasures in the RPL-Based Internet of Things”,Hindawi Publishing Corporation International Journal of Distributed Sensor Networks Volume 2013

[17] Andrea Zanella, Nicola Bui, Angelo Castellani, Lorenzo Vangelista,Michele Zorzi, “Internet of Things for Smart Cities”,22 IEEE Internet of things Journals, vol. 1, no. 1, Februrary 2014

[18] Angelo P. Castellani , Michele Rossi, Michele Zorzi, “Back pressure congestion control for CoAP/6LoWPAN networks”.

[19] C.P. Kruger and G.P. Hancke,”Implementing the Internet of Things Vision in Industrial Wireless Sensor Networks”.

[20] Geoff Mulligan, The 6LoWPAN architecture, 4th workshop on Embedded networked sensors, 2007.

[21] Saurabh Mehta,Najnin Sultana,Kyung Sup Kwak,”Network and System Simulation Tools for Next Generation Networks” a Case Study,Book of Modelling, Simulation and Identification, 2010.

A Low Cost Dual-Band Patch Antenna Array for Future 5G Communication Applications

Md Eshtiuk Hossain
Dept. of Electronics & Telecommunication Engineering, Rajshahi University of Engineering & Technology.
Rajshahi-6204, Bangladesh.
eshtiuk.ruet.ete13@gmail.com

Prosen Kirtonia
Dept. of Electronics & Telecommunication Engineering, Rajshahi University of Engineering & Technology.
Rajshahi-6204, Bangladesh.
Prosen.ete.ruet@gmail.com

Mst Fateha Samad
Dept. of Electronics & Telecommunication Engineering, Rajshahi University of Engineering & Technology.
Rajshahi-6204, Bangladesh.
fatehaeteruet@gmail.com

Abstract— A dual-band, 4×2 microstrip patch antenna (MPA) array for future 5G communication applications is proposed in this paper. The single element antenna of the proposed array configuration is a dual band patch antenna with dimension of 8×8 mm². This proposed single element antenna has resonance frequencies of 29 GHz and 40.5 GHz, and achieves 5.69 dBi and 4.54 dBi gain at the center frequencies respectively. The attained -10dB impedance bandwidth of the single element antenna is 10.34% and 12.34% at two operating frequency bands. Using corporate fed and series fed, the proposed 8-element MPA array is designed from this single element. The dimension of the proposed array antenna is 38 × 15 mm^2 and low-cost FR-4 material is used as the substrate. The proposed array antenna resonates at 29.5 GHz and 38 GHz, and provides 8.4 dBi and 9.3 dBi total gain at the resonance frequencies respectively. Thus, improved bandwidth and better radiation performance over required frequency bands, high gain and use of low-cost material makes the proposed antenna perfect for 5G communication/wireless communication applications.

Keywords- Antenna array, corporate fed, fifth generation (5G), millimeter-wave, dual-band.

I. INTRODUCTION

There are various generation of Mobile communication technology from 0*G* to 4*G* is introduced with the improvement of wireless communication system. The current generation of cellular communication system is 4G which is ineffective to solve the problems of poor coverage, crowded channel, poor quality, dropped connection and flexibility that unable to fulfill the consumer's requirements [1]. To fulfill these tradeoff, investigation and development of Fifth Generation (5*G)* of wireless communication have been started and it's expected to commercialize in 2020 [2].

5*G* is capable to provide a data rates of at least one gigabit per second for tens of thousands of users simultaneously, improved security features, large bandwidth, increase the energy efficiency and higher resolution for cells. So, 5G will provide support for interactive multimedia such as, virtual reality, autonomous driving etc. [3][4].

Modern/5G communication system requires the antennas with high standard of data transfer, small size, low cost of fabrication, more bandwidth which increases the demand of microstrip patch antenna. The limitation of low impedance bandwidth, low gain, and low efficiency of patch antenna can be overcome by using ARRAY configuration of a single element patch antenna. Also, the dimensions of the mobile phone rapidly decreasing that lead to the evolution of compact antenna structures. Multi-band characteristics, compact structure has emerged the patch as a promising candidate for handheld devices [5].

5G will use millimeter bands and probably require the generalization of the implementation of small-cell radio transmitter though no standard will be available until 2019 [6]. Fifth Generation (5G), here millimeter wave radio frequencies can provide the basic ground. Millimeter wave have unexploited spectrum (3GHz-300GHz) to fulfill the new generation needs. The range of spectrum for 5G application is 20-90GHz and different frequency bands will be allocated for different countries and different applications [7]. Recently, the 28 GHz and 38 GHz bands have been approved for mm-wave 5G operation in the USA by FCC [8].

Microstrip patch antenna is used in wireless and mobile communication because of its small size, low cost of fabrication, capability of supporting multiple frequency bands, and array configuration for the better gain. Each and every element of the array configuration has to be fed properly to make the array a perfect radiator. Single or multiple line feeding can be used in array. Series feeding network is single line feeding and multiple line feeding is called corporate feeding network [9].

In this work, a slotted microstrip patch antenna is designed which resonates at 29 GHz and 40.5GHz with total gain of 5.69 dBi and 4.54 dBi and impedance bandwidth of 3 GHz and 5 GHz at the resonance frequencies. FR-4 material has been used as substrate and thickness of the substrate is 0.8 mm. Then, for enhancing the gain 2×1, 4×1, and 4×2 patch antenna arrays is designed using corporate fed and series fed connection. Finally, an eight-element array antenna is developed which resonates at dual frequency bands dedicated for 5G communication and provides better radiation performance.

II. ANTENNA CONFIGURATION AND DESIGN

A. Single Antenna Element Design.

The schematic representation of the dual band single element patch antenna is shown in Fig.1. The dimension of the antenna is 8×8 mm², and designed on low cost FR-4 substrate with a dielectric constant (ε_r) of 4.4, loss tangent (tan δ) of 0.02 and thickness of 0.8 mm. It consists of a rectangular patch of dimension 3.25 ×2.3 mm² and microstrip feed line of 1 ×2 mm² having characteristic impedance equal to 50Ω. An E-shaped slot is truncated from the conventional rectangular patch to achieve dual-band characteristics and increase current flow path. Then the size of the slot is optimized to adjust the frequencies and achieve desired gain.

The detailed dimensions of optimized design of single element patch antenna are given in TABLE.I.

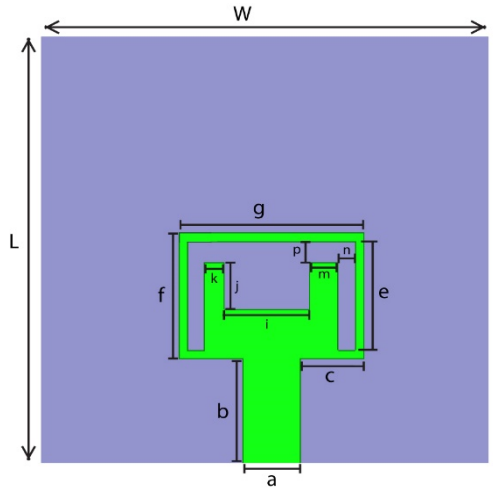


Fig.1 Single Element Patch Antenna.

TABLE.I. DETAILED DIMENSION OF THE SINGLE ELEMENT DESIGN

Parameter	Value(mm)	Parameter	Value(mm)
W	8	g	3.25
L	8	k	0.35
a	1	i	1.5
b	2	j	0.86
c	1.125	p	0.47
e	2	m	0.5
f	2.3	n	0.3

B. Array Antenna Design.

To ameliorate the total gain value array configurations are used. Here 4×1, four elements and 4×2, eight elements array antennas are designed. The schematic representation of 4×1, four element array antenna is shown in Fig.2.

The elements are connected using corporate fed in this configuration. Multiple fed lines are used to connect the elements, and match the impedance of the feeding line and antenna. The dimension of the 4×1, four element antenna array is 38 × 16 mm². The detailed dimensions of optimized fed line and distance between elements are given in TABLE.II.

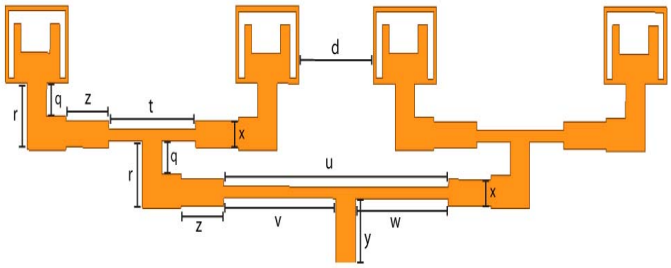


Fig.2 4×1 Array Configuration.

TABLE.II. DETAILED DIMENSION OF THE 4×1 ARRAY DESIGN

Parameter	Value (mm)	Parameter	Value (mm)
u	11.6	t	4.5
v	5.8	r	2
w	4.8	q	1
x	0.8	d	3.85
y	2.05	z	2.2

Finally, the proposed 4×2, eight elements antenna array is found by adding single element using corporate fed and series fed, which is shown in Fig.3. The size of proposed 4×2, eight elements antenna array is 38 × 15 mm². In the proposed 4×2 antenna array design, the dimensions of the fed line and single element are same as TABLE.I and TABLE.II . Here, series fed connections are used to add four antenna elements with previously described 4×1 antenna array to achieve dual-band characteristics and improved gain. The proposed antenna array is designed and simulated using Ansys HFSS software.

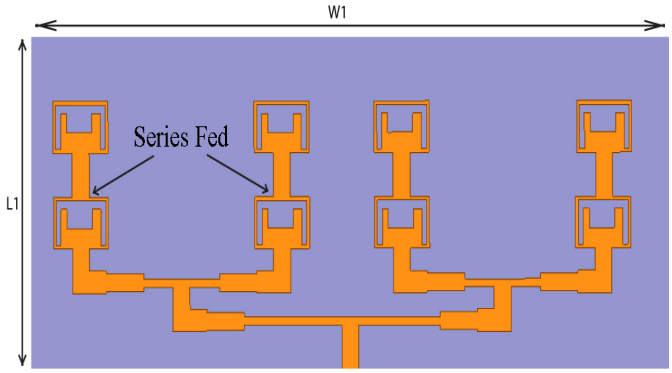


Fig.3 Proposed 4×2 Array Antenna

III. SIMULATION AND RESULTS

A. Simulation of Single Antenna Element.

Reflection coefficient shows the return loss due to mismatch in impedance of fed line and antenna. So, the lower value of reflection coefficient is preferred and -10 dB is the standard value for calculating impedance bandwidth. The proposed single element design has a return loss of -25.03 dB and -29.7 dB with VSWR of 1.11 and 1.06 at the resonating frequencies respectively which is shown is Fig.4. The achieved -10dB impedance bandwidth of this single element antenna is 10.34% and 12.34% at two resonating frequency bands which covers two 5G communication band.

The single element design provides total gain of 5.69 dBi and 4.54 dBi at two resonating frequencies respectively which are illustrated in Fig.6 using radiation pattern in E-plane and H-plane. The simulated 3d polar plots of gain at resonance frequencies are delineated in Fig.7 and Fig.8. This design provides directive radiation pattern and maximum gain at a specific direction and angle. In the

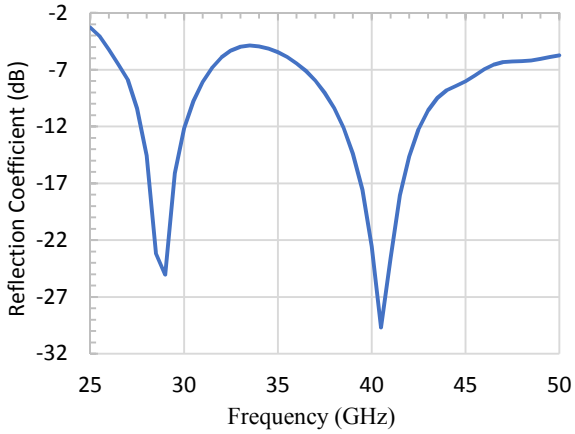


Fig.4 Reflection Coefficient of Single Element Antenna.

simulation, the radiation efficiency of single element antenna is found 79.26% and 87.13% at 29 GHz and 40.5 GHz respectively.

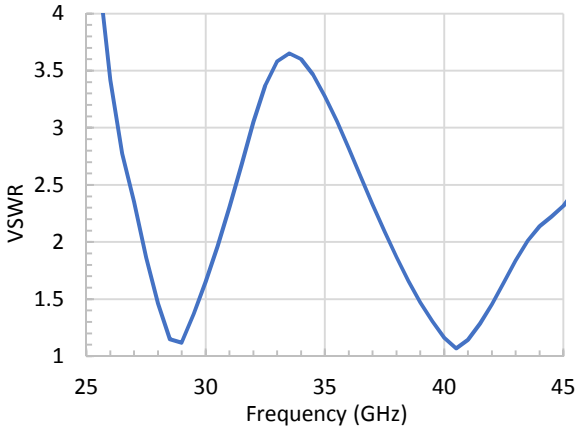


Fig.5 VSWR of Single Element Antenna

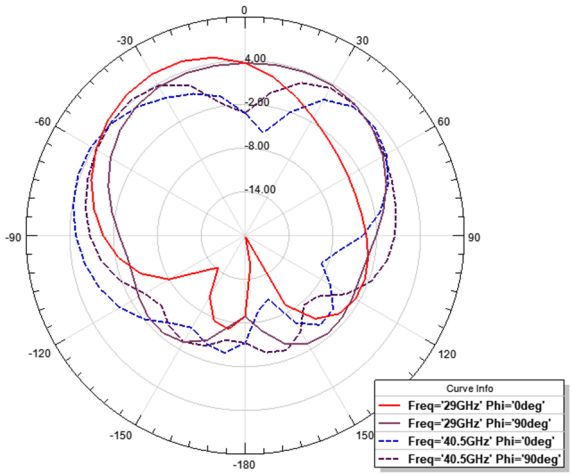


Fig.6 Gain of Single Element Antenna in Radiation Pattern

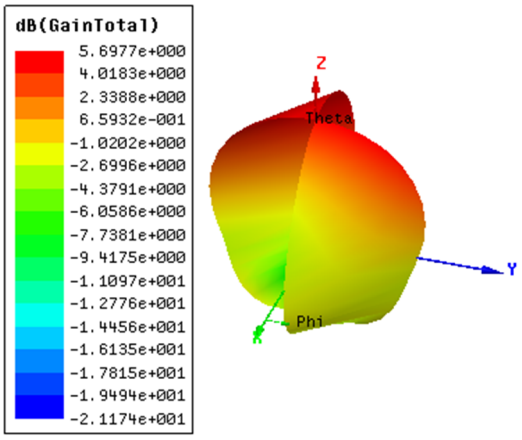


Fig.7 Simulated Gain of Single Element Antenna in 3d Plot at 29 GHz.

B. Simulation of Proposed Array Antenna.

Here, Fig.9 shows the simulated return loss of 4×1, four elements and 4×2, eight elements array antenna. Both are evident that they operate over two desired frequency bands allocated for 5G communications. The designed 4×1, four element array antenna resonates at 30 GHz, 32 GHz, 34.5 GHz, 38 GHz and provides return loss of -17 dB, -20.5 dB, -15.5 dB, and -18.7dB respectively. Thus, the simulated resonance frequency of proposed 4×2, eight elements array antenna is 29.5 GHz, 32 GHz, 35 GHz, and 38 GHz, and provides return loss of -18 dB, -17.9 dB, -16.1 dB, and -17.6 dB at the resonating frequencies respectively. Among the operating frequency bands, two bands are extensively used for 5G communication research.

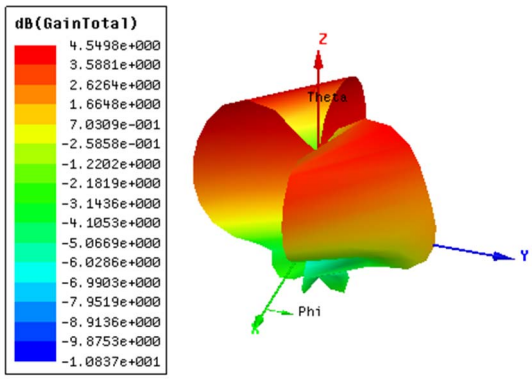


Fig.8 simulated Gain of Single Element Antenna in 3d Plot at 40.5 GHz.

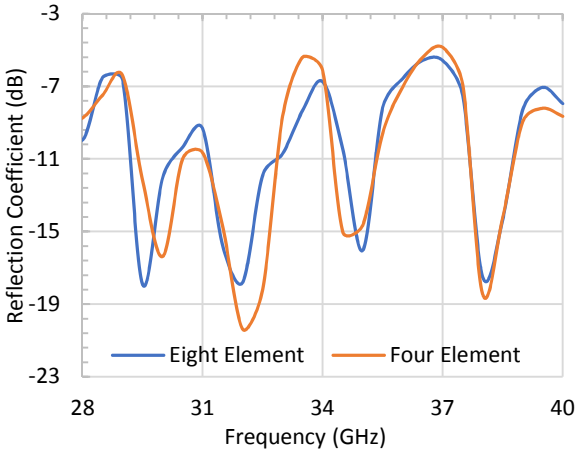


Fig.9 Reflection Coefficients of 4-Element and 8-Elements Array Antenna.

The achieved -10dB impedance bandwidth of 4×1 antenna array is 4.1% and 3.6% at resonance frequencies of 30 GHz and 38 GHz respectively. Here, the proposed 4×2 antenna array attains -10dB impedance bandwidth of 4.5% and 3.6% at the resonance frequencies of 29.5 GHz and 38 GHz respectively. The radiation pattern of the total gain in E-plane and H-plane of 4×1 antenna array at the resonating frequencies is shown in Fig.10. The total gain of 4 element array antenna is found 8.1 dBi and 8.8 dBi at the operating frequencies respectively.

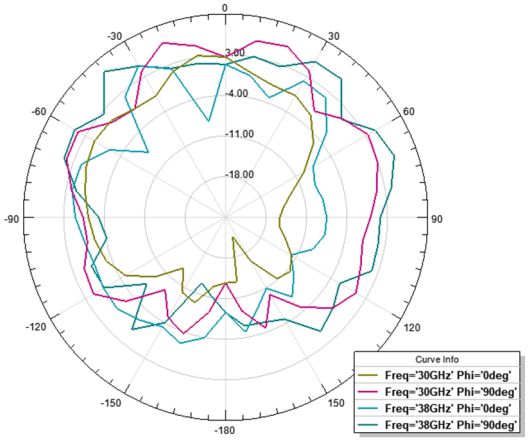


Fig.10 Gain of 4×1, 4-Element Array Antenna.

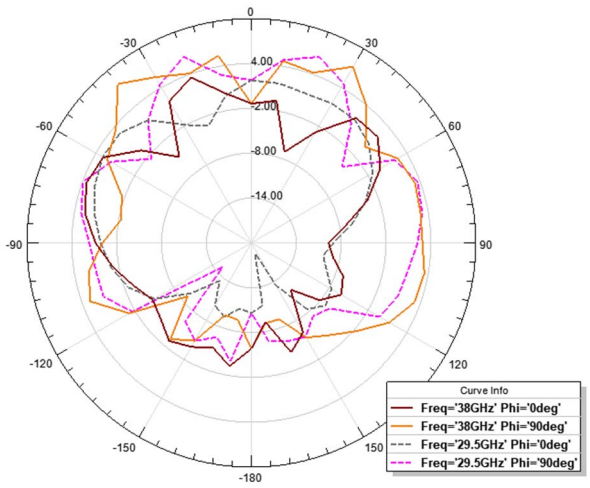


Fig.11 Gain of proposed 4×2, 8-Element Array Antenna.

The gain of the proposed 4×2, eight element array antenna is illustrated in Fig.11 using radiation pattern. Total gain of this proposed 4×2, eight element array antenna is achieved 8.4 dBi and 9.3 dBi at resonance frequencies respectively. The radiation efficiency of the proposed array configuration is 98.3% and 98% at the center frequencies respectively. Here, Table.III and Table.IV contain the performance comparison of single element design and proposed array design with similar existing works to limn the improvement in outcomes of the proposed antennas.

TABLE. III. PERFORMANCE COMPARISON BETWEEN PROPOSED SINGLE ELEMENT ANTENNA AND SIMILAR EXISTING WORKS

No	Dimension (mm²)	Resonance Frequency (GHz)		Gain (dBi)		Bandwidth (GHz)	
		F1	*F2*	*F1*	*F2*	*F1*	*F2*
[7]	7 × 3	28	38	3.75	5.06	3.34	1.395
[10]	19 × 19	10.15	28.1	5.51	7.54	.278	1.085
Proposed	8 × 8	29	40.5	5.69	4.54	3	5

TABLE. IV. PERFORMANCE COMPARISON BETWEEN PROPOSED ARRAY ANTENNA AND SIMILAR EXISTING WORKS

No.	Dimension (mm²)	Resonance Frequency (GHz)		Gain (dBi)	
		F1	*F2*	*F1*	*F2*
[11]	(64-elements)	-	38	-	6.5
[12]	130 × 74	29	32	5.39	4.92
Proposed	38 × 15	29.5	38	8.4	9.3

IV. CONCLUSION

A compact array antenna operates at two frequency bands which are dedicated for 5G communication is proposed in this work. A dual-band single element patch antenna is designed to make the desired MPA array configuration. Here, different array configurations are used using different feeding technique to achieve the desired radiation performance. The material used in this design is low cost FR-4. The proposed MPA antenna encloses two 5G communication frequency bands (28GHz and 39.5GHz) entirely and provides improved gain of 8.4dBi and 9.3dBi. Low cost, operation over two bands, wide bandwidth and high gain with compact structure ensures that the proposed antenna has great potential for 5G mobile communication applications.

V. REFERENCES

[1] Meenal G. Kachhavay and Ajay P.Thakare; "5G Technology-Evolution and Revolution: Overview on Generations of Network: 1G,2G,3G,4G,5G", *International Journal of Computer Science and Mobile Computing*, Vol.3, Issue. 3, March 2014.

[2] Palak Sharma, Megha Verma, Neha Sundriyal and Jyoti Chauhan, **"**5G Mobile Wireless Technology**"**, *International Journal of Research*, Vol-1,Issue-9, October 2014.

[3] Abhishek Gupta, Dr.Anupam Gupta and Sarthak Gupta; "5G : The Future Mobile Wireless Technology by 2020", *International Journal of Engineering Research & Technology*, Vol.2 Issue 9, September 2013.

[4] Shivangi Verma, Leena Mahajan, Rajesh Kumar, Hardeep Singh Saini and Naveen Kumar, "A Small Microstrip Patch Antenna for Future 5G Applications",*5th International Conf. on Reliability, Infocom Technologies and Optimization*, AIIT, India Sep. 7-9, 2016.

[5] Ilhem Gharbi, Rim Barrak, Mourad Menif and Hedi Ragad, "Design of patch array antennas for future 5G applications",*18th international conference on Sciences and Techniques of Automatic control computer engineering* - STA'2017, Monastir, Tunisia, December 21-23, 2017.

[6] Naser Al-Falajy and Omar Y.K. Alani, **"**Design considerations of ultra-dense 5G network in millimeter wave band",*2017 Ninth international Conference on Ubiquitous and Future Networks (ICUFN)*, pp.141-146, 2017.

[7] Waleed Ahmad and Wasif Tanveer Khan, "Small form factor dual band (28/38 GHz) PIFA antenna for 5G applications."*In 2017IEEE MTT-S International Conference on microwaves for Intelligent Mobility (ICMIM)*, pp.21-24, 2017.

[8] Mohammad S.Sharawi & Mohammad Ikram, "Slot Based Connected Antenna Array for 5G Mobile Terminal.", *International Workshop on Antenna Technology (iWAT),* Nanjing, China , pp: 1-3, 2018.

[9] Dilipsingh Rawat, Gaurav singh and R P singh, "Design of a corporate fed 1*2 Microstrip Array Antenna for X Band Application.*", International Journal of computer Application* (0975-8887), Volume 95- No.5, June 2014.

[10] Yassine JANDI, Fatima GHARNATI & Oulad said ahmed, "Design of a compact Dual bands patch antenna for 5G Applications", *International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS),* Fez, Morocco, 2017.

[11] Cheng-Nan Hu , Kevin Peng , Chung-Hang Yu , Tsai-Wen Hsaio & Der-Phone Lin,"Design of a mm-wave microstrip antenna array.",*International Workshop on Electromagnetics: Applications and Student Innovation Competition (iWEM),* Taiwan , 2015.

[12] Manoj Stanley, Yi Huang, Tian Loh, Qian Xu3, Hanyang Wang & Hai Zhou, "A High Gain Steerable Millimeter-Wave Antenna Array for 5G Smartphone Applications*.", 11th European Conference on Antennas and Propagation (EUCAP)*, 2017.

Differentially Fed Compact Triple-band Antenna for Implantable Medical Devices

Prosen Kirtonia
Electronics & Telecommunication Engineering.
Rajshahi University of Engineering & Technology.
Rajshahi-6204, Bangladesh.
prosen.ete.ruet@gmail.com

Mst. Fateha Samad
Electronics & Telecommunication Engineering.
Rajshahi University of Engineering & Technology.
Rajshahi-6204, Bangladesh.
fatehaeteruet@gmail.com

Md Kamal Hosain
Electronics & Telecommunication Engineering.
Rajshahi University of Engineering & Technology.
Rajshahi-6204, Bangladesh.
kamaleteruet@gmail.com

Md Eshtiuk Hossain
Electronics & Telecommunication Engineering.
Rajshahi University of Engineering & Technology.
Rajshahi-6204, Bangladesh.
hridoy6699@gmail.com

***Abstract*— A compact triple-band antenna for biomedical telemetry applications is presented in this paper. The proposed antenna operates at MedRadio band (402-406 MHz) and ISM bands (902-928 MHz & 2.40-2.48 GHz) whose resonance frequencies are 402 MHz, 914 MHz, and 2.43 GHz. A meandering slotted patch over the silicon substrate of standard 0.275 mm thickness and a sorting pin is used to achieve the desired results and size. The performances of the proposed antenna are evaluated and studied in free space, in a skin phantom cube and inside the fat of a six-layer phantom model using Ansoft HFSS software. Differential feeding is utilized in the design and the volume of the presented antenna is only 46.475 cubic millimeter excluding superstrate. The simulated bandwidth of the proposed antenna is 11.7%, 0.77%, and 2.88% at the resonant frequencies respectively. Finally, the specific absorption rate is determined at the center frequencies in the phantom model for the assessment of patient safety. Compact structure, bio-compatibility, improved realized gain and operation over three different biomedical frequency bands enables the antenna to be used with implantable medical devices with great possibilities.**

Keywords*— Biomedical Telemetry, Implantable antenna, Low profile, Miniaturization, Triple-band.

I. INTRODUCTION

The use of biotelemetry system in the implantable medical devices (IMD) like pacemaker, defibrillator, glucose monitor etc. are increasing gradually because of thriving demand in continuous monitoring of physiological parameters, diagnosing diseases of patients, and treating some diseases [1]. In a biotelemetry system, implantable antenna is the major element for transmitting and receiving physiological signals. In this perspective, an implantable antenna must be stable in terms of performance and biocompatible for assuring patient safety. Moreover, it has to be compact in structure as the IMDs are very small in size [2].

Various types of antenna such as PIFA [3], slot antenna, patch antenna [4], spiral antenna [5] etc. have been proposed in last few years for biotelemetry applications. These implantable antennas are operated with radio-frequency coupling strategy which is a great alternative to traditional inductive coupling coils. There are few radio frequency bands including MedRadio bands (402-406 MHz, 413-419 MHz, 426-432 MHz, 451-457 MHz) [6], ISM bands (433-434.8 MHz, 902-928 MHz, 2.4-2.48 GHz, and 5.725-5.875 GHz) which have been dedicated to biomedical telemetry applications by the Federal Communications Commission (FCC). Antennas for biotelemetry applications can operate at single or the multiple numbers of frequency bands from the bands mentioned above.

In literature, some work has conducted on multiband antennas. For instance, a differentially fed dual band implantable antenna operating at MICS band (402-405 MHz) and ISM band (2.4-2.48 GHz) has been proposed in [7]. The simulated -10dB bandwidth of that antenna was 12.6% at MICS and 5.7% at ISM band respectively. However, the volume of this antenna is 691.515 mm^3 which is comparatively large. Another differentially fed dual band implantable antenna is reported in [4]. This antenna also operates at MICS and ISM band mentioned above. After using the shorting pin and meandering slots the obtained volume of that antenna was 642.62 mm^3.

The important requirements while designing an antenna for biomedical application are patient safety and stable radiation performance. For assuring patient safety, the assessment of biocompatibility usually done through specific absorption rate (SAR) calculation at the resonant frequencies. The SAR of the antenna at the operating frequencies must follow the maximum allowable standard. The maximum value of the average SAR should be equal to 1.6 W/kg for 1-g tissue model as IEEE C95.1-1999 standard [8]. According to IEEE C95.1-2005 standard the maximum average SAR should be equal to 2 W/kg [9]. Moreover, after implantation, the surrounded environment (e.g. human/animal tissues) influences the performance of the antenna. So, the performance of the antenna must be optimized by considering tissues in the surrounding environment. Thus, the antenna performance needs to be evaluated in free space and in different tissue models to figure out the change in antenna performance in different conditions.

In this work, a compact triple-band antenna is proposed. The antenna is simulated in free space, in a skin phantom, and in six-layer phantom model to find the resonance frequencies and realized gain. Finally, SAR of the antenna is determined for assuring patients safety and biocompatibility of the proposed implantable antenna.

II. ANTENNA CONFIGURATION

In this paper, a patch antenna is proposed because of low cost, easy fabrication, compact size, and flexibility of design. The dimension of the proposed design is 13×13×0.275 mm^3 excluding superstrate. The schematic representation of the proposed antenna is shown in Fig.1. A silicon layer of 0.275mm thickness with the dielectric constant of $\in_r$=11.9, loss tangent of tanδ = 0.0012 is used as the substrate layer. In this design, a meandering slot, shorting pin, and rectangular slots are applied to minimize the dimension of the antenna and find the desired output.

Firstly, a meandering slot is truncated on the center of the patch to increase the current path between the feed and the radiating patch hence to reduce the antenna size significantly. Secondly, two rectangular slots are mangled to adjust the resonating frequencies and antenna size by increasing current patch or altering patch current distribution. Thirdly, a sorting pin of 0.4 mm is used for altering the reactance of the antenna which is required to further miniaturization of the antenna and reduce the resonant frequencies. For achieving the lower value of the axial ratio, two slots are made on the ground which can be adjusted to obtain circular polarization.

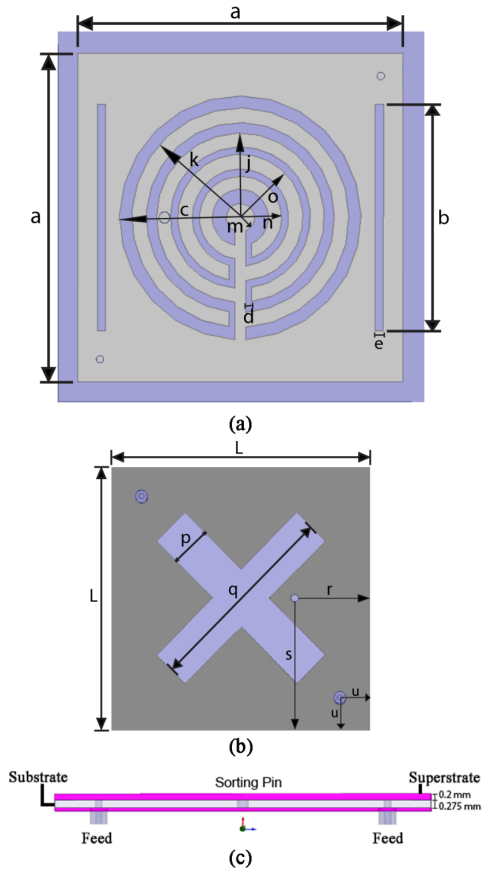


Fig.1. Schematic Representation of Proposed Antenna a) Patch, b) Ground, and c) Side View.

For avoiding the short circuit between metallic ground/patch and tissue after implantation, a separation layer is used which is called superstrate layer. A thin layer of silicon-di-oxide with the dielectric constant of $\in_r$= 4.0, loss tangent of tanδ = 0.006 is used as a superstrate. The thickness of the superstrate is 0.1 mm. Superstrates are attached all over the metallic patch/ground to reduce the losses by human/animal tissues and improve the radiation performance [7].

The size of the slots, the location of sorting pin is optimized by iterative simulation to obtain desired resonance frequencies and radiation performance. The feeding ports are connected to 50 Ω coaxial cable. Iterative simulation of the antenna with different feeding position is done to obtain impedance matching of antenna with the feed line. Proper impedance matching is needed to reduce return loss and increase the efficiency of the antenna. After the optimization and adjustment of size, the optimal geometric parameters are obtained which are listed in TABLE.I.

TABLE I. DETAILED DIMENSION OF THE PROPOSED ANTENNA

Parameters	Value (mm)	Parameters	Value (mm)
L	13	e	0.3
p	2	d	0.2
q	10	c	4.30
r	3.8	m	0.5
s	6.5	n	1.46
u	1.5	o	2.23
a	11.6	j	2.98
b	8	k	3.86

III. SIMULATION AND RESULTS

A. *Simulation in free space.*

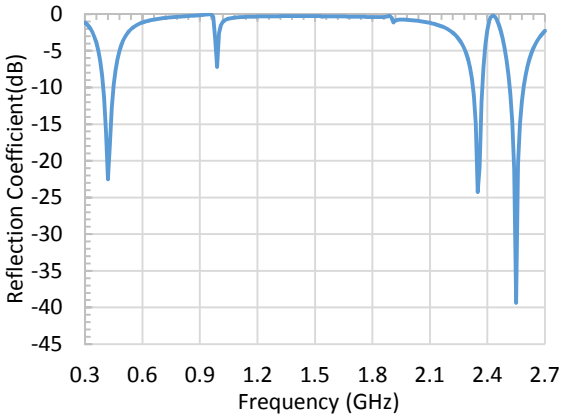


Fig.2. Reflection Coefficient in Free Space.

The return loss response of the proposed antenna in free space is illustrated in Fig.2. In free space, the antenna resonates at 0.42 GHz, 0.99 GHz, 2.35 GHz and 2.55 GHz which are shown in Fig.2. The proposed antenna is optimized following the phantom model simulation result because of purpose of implantation. It is simulated in the free space to observe the antenna performance and effect of phantom model on the antenna only that's why the antenna is not tuned in desired frequency bands. The simulated radiation efficiency of the antenna is found 94.84%, 98.49%, and 95.91% at 0.42 GHz, 2.35 GHz, and 2.55 GHz respectively in free space.

B. *Design of Phantom Models.*

The performance of an implantable antenna should be determined inside the phantom model for estimating the antenna performance after implantation. Antenna performance depends on surrounded or radiating environment. Moreover, for analyzing the biocompatibility of the designed implantable antenna, it is required to simulate the antenna in phantom models also. An antenna generally shows different characteristics in the different environments like air, fat, skin, tissue, multi-layer model etc.

Depending on requirements, the antenna must be simulated inside single/multi-layer skin model or human

body model. Here, a skin phantom cube and a six-layer model is designed to simulate the proposed antenna. Although the body model of an animal or human is very complex, here, a simple six-layer model is considered for simulation of the design. Human/animal body tissues show different dielectric characteristics in different frequency and an average value is used in the design of the models as the proposed antenna operates over multiple frequency bands. Geometric representation and detailed dimension of skin model and six-layer phantom model are shown in the Fig.3 and dielectric properties of the tissues of the models are given on Table.II.

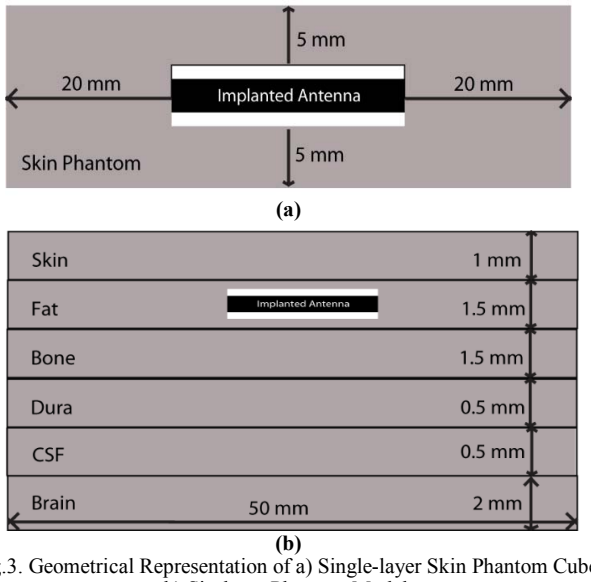


Fig.3. Geometrical Representation of a) Single-layer Skin Phantom Cube, b) Six-layer Phantom Model.

TABLE II. THE DIELECTRIC PROPERTIES OF DIFEERENT TISSUES [10]

Tissues	Permittivity	Mass Density	Conductivity(S/m)
Skin	40.2	1080 kg/m^3	0.8
Fat	15	920 kg/m^3	0.35
Bone	4.9	1180 kg/m^3	0.15
Dura mater	44.39	1030 kg/m^3	0.966
CSF	68.61	1110 kg/m^3	2.419
Brain	50	1050 kg/m^3	1

C. Antenna Performance in Phantom Models.

The proposed antenna is optimized in a single layer skin phantom cube, and the optimized design is simulated in free space and six-layer rat head model later. The resonance frequencies of the proposed design are found 402 MHz, 914 MHZ, and 2.43 GHz in the simulation inside skin phantom cube. Value of S_{11} parameter at the resonant frequencies (402 MHz, 914 MHZ, and 2.43 GHz) are -24.74 dB, -16.27 dB, and -14.52 dB respectively as shown in Fig.4. The VSWR at resonance frequencies are 1.12, 1.37, and 1.50 respectively. The achieved -10dB impedance bandwidths of the antenna are 11.7%, 0.77%, and 2.88% at 402 MHz, 914 MHZ, and 2.43 GHz frequency respectively. Simulated radiation efficiency of the proposed antenna in the skin phantom cube is attained 80.32%, 60.10%, and 88.18% at 2.43 GHz, 914 MHz, and 402 MHz respectively.

In the simulation inside the six-layer model, the optimized design resonates at 360 MHz, 2.13 GHz, and 2.3 GHz. The simulated data of the reflection coefficient are plotted and shown in Fig.4 with the comparison of single layer model results.

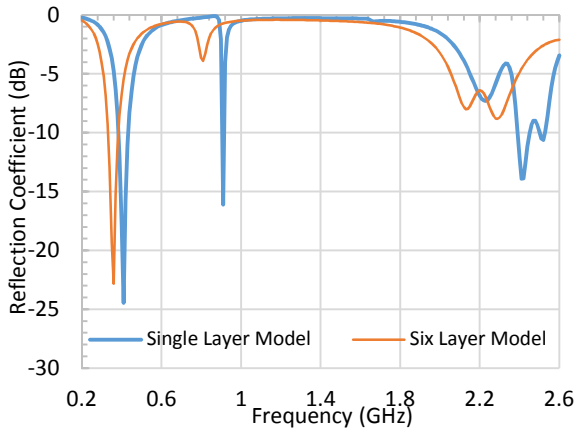


Fig.4. Reflection Coefficient of the Proposed Antenna in Skin Models.

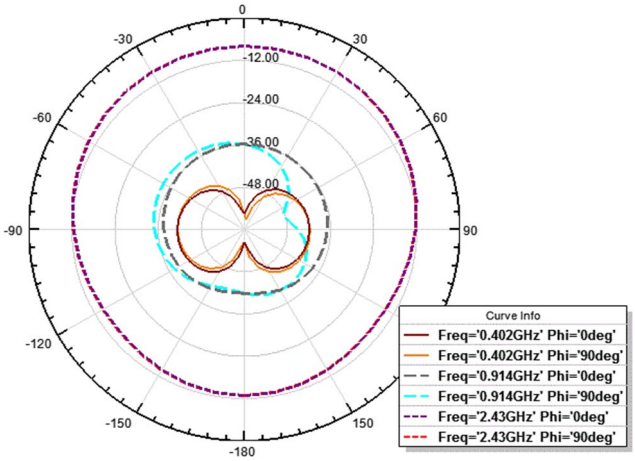


Fig.5. Realized Gain of the Proposed Antenna in Skin Cube

The peak realized gains of the proposed antenna inside skin phantom cube are attained of -41.30 dBi, -33.35 dBi, and -7.90 dBi at 402 MHz, 914 MHZ, and 2.43 GHz respectively which are illustrated in Fig.5. The angles at which the peak realized gains are found are -90-degree, -50 degree, and 0 degree. In the simulation inside the fat of six-layer model, the peak realized gain of the antenna are -16.01 dBi, -11.52 dBi and -42.9dB at resonance frequencies found in the simulation in six-layer model respectively.

The maximum value of current distribution over the patch of the proposed antenna is found at 2.45 GHz. Current distribution of the antenna is represented at Fig.6. High value of current on the patch can cause heat there which must be avoided. Here, the value of Jsurf over the patch is sustainable, and superstrate layers are applied over the patch and ground.

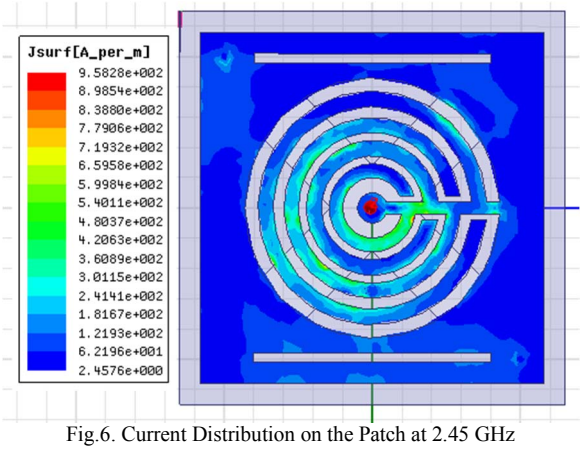


Fig.6. Current Distribution on the Patch at 2.45 GHz

A performance comparison between the proposed design and similar existing works are delineated in Table. III to limn the performance improvement of the proposed antenna.

TABLE III. COMPARISON BETWEEN PROPOSED AND SIMILAR EXISTING ANTENNA

No.	Vol. (mm^3)	Implant Tissue	Resonating Frequency (MHz)	-10 dB impedance bandwidth	Gain (dBi)
[7]	691.515	Skin	400 2400	12.6% 5.7%	-30 -
[4]	642.62	Skin	402 2400	7.4% 6.6%	-36.7 -27.1
This work	46.475	Skin Fat	402 914 2430	11.7% 0.22% 2.88%	-41.30 -33.35 -7.90

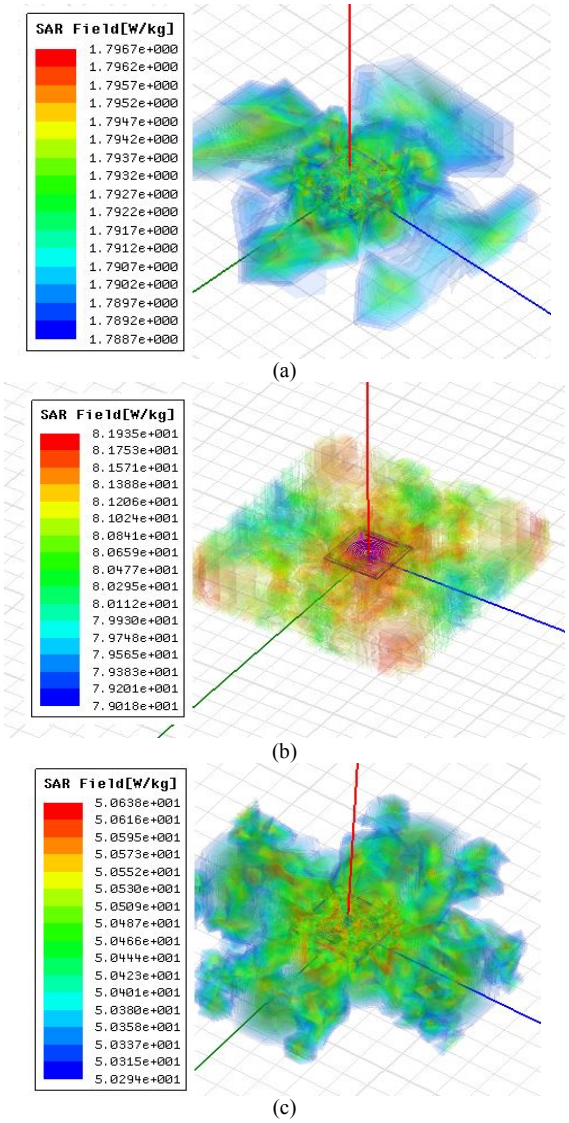


Fig.7. The Value of Specific Absorption Rate at a) 402 MHz, b) 914 MHz, and c) 2.43 GHz.

D. Biocompatibility Analysis.

Specific absorption rate analysis is the factor to evaluate the radiation effect of the antenna on the human/animal body to determine the safety level. The proposed patch antenna transmits and receives the electromagnetic energy wirelessly while working with an implantable device. A portion of the electromagnetic energy which is transmitted/received interacts with the tissue around the implantable antenna and assessment of that absorbed energy is called biocompatibility analysis. Maximum allowable specific absorption rate and conditions discussed in previous section [9][10]. The specific absorption of the proposed antenna is simulated at the resonating frequencies, assuming 1W power is radiated by the antenna and mass of the skin phantom cube is 1g. After the SAR analysis of the proposed antenna at the resonance frequencies, the maximum specific absorption rate is found 81.93 W/Kg at 914 MHz. So, the radiated power of the antenna must be less than 19.53 mW to satisfy the maximum allowable specific absorption rate of 1.6 W/Kg for 1g tissue. The results of the specific absorption rate analysis at the resonance frequencies are shown in Fig.7 respectively.

IV. CONCLUSION

A differential fed compact in size triple-band antenna is designed in this paper. The volume of the proposed antenna is 46.475 mm^3 without superstrate and the antenna resonates at three frequency bands which are dedicated for biomedical applications (MedRadio and ISM band). The -10dB impedance bandwidth of the antenna fully encloses one MedRadio (401-405 MHz) and one ISM band (2.40-2.48 GHz), and partially covers another ISM band (910-917 MHz). Due to implantation purpose, the performance of the designed antenna is assessed in free space and inside phantom models. The SAR is calculated for ensuring biocompatibility and patient safety. Before concluding about the antenna performance and patient safety, an extensive experimental study is needed as the results of this paper is totally simulation-based.

V. REFERENCES

[1] F. Nebeker, "Golden accomplishments in biomedical engineering," *IEEE Eng. Med. Biol. Mag.*, vol. 21, no. 3, pp. 17–47, May/June 2002.

[2] C. Liu, Y. D. Zhang, X. G. Liu, "Circularly Polarized Implantable Antenna for 915 MHz ISM-Band Far-field Wireless Power Transmission," *IEEE Antennas Wireless Propag. lett.,* vol. 17, No. 3, March 2018

[3] A. Kiourti and K. S. Nikita, "Miniature scalp-implantable antennas for telemetry in the MICS and ISM bands: design, safety considerations and link budget analysis," *IEEE Trans. Antennas Propag*., vol. 60, no. 8, pp.3568-3575, Aug. 2012.

[4] Y. Liu, Y. Chen, H. Lin, and F. H. Juwono, "A Novel Differentially Fed Compact Dual-Band Implantable Antenna for Biotelemetry Applications," *IEEE Antennas and Wireless Propag. letters*, vol.15, pp. 1791–1794, 2016.

[5] Jae-Ho Lee, Dong-Wook Seo, and Hyung Soo Lee "Design of Implantable Rectangular Spiral Antenna for Wireless Biotelemetry in the MICS Band," *ETRI Journal*, Volume 37, Number 2, April 2015.

[6] FCC MedRadio Specification [Online]. Available: https://www.fcc.gov/wireless/bureau-divisions/broadband-division/medical-device-radiocommunications-service-medradio.

[7] C. Liu, Y. X. Guo, and S. Xiao, "Compact Dual-Band Antenna for Implantable Devices," *IEEE Antennas and Wireless Propagation letters*, vol. 11, no. 8, pp. 1508–1511, Aug. 2012.

[8] IEEE. Standard for Safety Levels with Respect to Human Exposure to Radiofrequency Electromagnetic Fields, 3 kHz to 300GHz. IEEE Standard C95.1. New York, NY, USA: IEEE, 1999.

[9] IEEE Standard for Safety Levels with Respect to Human Exposure to Radiofrequency Electromagnetic Fields, 3 kHz to 300 GHz, IEEE Standard C95.1, 2005.

[10] Md Kamal Hosain, Abbas Z. Kouzani, Susannah J. Tye, Osama A. Abulseoud, Andrew Amiet, Amir Galehdar, Akif Kaynak, and Michael Berk,"Development of a Compact Rectenna for Wireless Powering of a Head-Mountable Deep Brain Stimulation Device", *IEEE Journal of Transitional Eng. in Health and Medi.,* Vol.2, 2014.

Design & Analysis of a Micro-strip Patch Antenna for RFID, WiMAX and X-band Applications

Mst. Dilshad Jahan [1]
Department of Electrical & Electronic Engineering
Bangladesh University of Business & Technology
Dhaka, Bangladesh
dilshadjahan011@gmail.com

Chinmoy Das [2]
Department of Electrical & Electronic Engineering
Bangladesh University of Business & Technology
Dhaka, Bangladesh
chinmoyd299@gmail.com

Nayan Sarker [3]
Department of Electrical & Electronic Engineering
Bangladesh University of Business & Technology
Dhaka, Bangladesh
nayan_ece09@yahoo.com

***Abstract*— This work deals with the design and simulation of patch antenna that provides resonance at frequency 9.968 GHz and applicable for RFID, WiMAX, and X-band applications. The antenna has a -10 dB impedance bandwidth of 300 MHz ranging from 9.8 GHz to 10.1 GHz. The prospect of this design is to obtain a small size, light weight and low-cost miniaturized antenna with good antenna characteristics and ease of integration using feed-networks. The basic theory and design are analyzed, and the simulation results are obtained using computer simulation technology (CST) Microwave Studio. The proposed antenna provides excellent VSWR of 1.208 and a gain of 5.960 dB. The simulated directivity, return loss (S11) and radiation efficiency are 6.015 dBi, -20.49 dB and 98.14% respectively. The simulation and PCB design of the proposed antenna are carried out using FR4 (lossy) substrate. The experimental radiation pattern is investigated using wave and antenna training system (WATS 2002) module. Finally, the simulated radiation pattern is compared with the experimental one.**

Keywords*— Patch Antenna, Microstrip, RFID, X-band, VSWR.

I. INTRODUCTION

Now-a-days, technology is rapidly switching and make many things easier day by day. Since the innovative thinking of a person is increasing day-by-day, new methods for wireless communication have been evolved. As the lower frequency band are almost occupied and due to spectral crunch wireless technology are switched to the higher frequency band. Higher frequency band means there is a greater chance to achieve higher bandwidth as a result multi-gigabit data speed can be achieved according to the user demand [1]. Since the performance of an antenna directly affects the overall communication system performance, so an accurate antenna design, especially at a higher frequency band, is a very essential task. Huge research has been done in the area of antenna design at the industrial, scientific and medical band (ISM), high frequency (HF) band, ultra-high frequency (UHF) band and centimeter band [1]. It should be great concern that the designed antenna should be low profile antenna that means the antenna have light weight, lower substrate thickness, and lower dimensional parameters with acceptable antenna performance metrics. At present low profile antennas with good performance characteristics are a suitable candidate for any wireless technology especially at wireless-fidelity (Wi-Fi), worldwide interoperability for microwave access (WiMAX), and X-band applications like satellite communications [2,4]. On the other hand, if we consider both commercial and military applications wideband, multiband and low profile antennas are great demand for a modern wireless communication system that influenced the research on antennas in many directions. The researcher is exploiting various types of antennas like as L-shaped, F shaped, inverted F-shaped, and E-shaped microstrip patch antenna for different applications like WLAN, GPS, GSM, mobile radio, wireless communications [2,3]. The microstrip patch antenna is a very good candidate and can perform very well in WiMAX and X-band satellite communication due to its ease of fabrication, low profile and acceptable performance [2,4]. The patch of a microstrip patch antenna is generally made of metal like copper or gold. But, the surface of the dielectric substrate feed line and the radiating patch are basically photo etched. Just for its fringing fields, the patch is electrically bigger than its physical size. The rectangular microstrip patch antenna is a very simple design considering its geometric structure but the electromagnetic field is complex. Calculating all the parameters need thorough analysis and accurate calculation. The main advantage of a microstrip patch antenna is its light weight and simple design. But the drawbacks are the narrow frequency bandwidth, deceptive feed radiation, poor polarization purity, limited power capacity, tolerance problem and the excitation of surface waves [5]. To design microstrip patch antenna substrate plays an important role to achieve good antenna performance. Because higher the substrate thickness lowers the antenna resonance frequency but higher the antenna gain, directivity, and bandwidth. In addition, if the dielectric constant (permittivity) of the antenna substrate is decreased then antenna resonance frequency and bandwidth are increased. So it is necessary to adjust antenna dielectric substrate thickness as well as substrate permittivity according to the technology requirements [8].

II. LITERATURE REVIEW

We have reviewed some of the previous antennae works as well as we have discussed their antenna size, shape, operating frequency band, gain, return loss etc. In Section V we have compared our resultant values with some of the reviewed works. The literature review table is shown in "TABLE.1

TABLE I:LITERATURE REVIEW TABLE

References	Antenna Size in mm^2	Antenna Shape	Operating Freq. Band	f_r in GHz	BW in GHz	Gain in dB	Return Loss	Substrate Material	η_{radeff} (%)
[12]	40 × 60	Tri-band fractal antenna	Centimeter band	3.85, 5.80, 8.12	3.869 ,4.034	2.6 dBi	−10 dB	FR4 epoxy substrate	91.2
[2]	46.4×14.8	Planar monopole antenna	Centimeter band	2.01	2–12.5	5.3 ,7.4	−10..2dB	Metamate-rial substrate	84.6
[4]	24×37	Multiband microstrip antenna	Centimeter band	4.30.5.51 GHz	5.8,4.1	4.49,0.44 dB	-18.85,-14.72dB	FR4 epoxy substrate	98
[3]	50×40	Double U-slot	Centimeter band	3.2 to 3.8	0.48, 0.3	-	-18.5, -14.5 dB	Dielectric substrate	-
[8]	40.29×35.12	Novel multiband patch antennas	Centimeter band	7.30, 9.50	7.66-7.84	7.628dB	-10 dB	Rogers RT5880	85, 87.14
[14]	36 × 34	Dual-band dipole antenna	Centimeter band	2.39, 5.86 GHz	2.50-2.42	2.45, 5.8dB	<=-10 dB	FR4 epoxy substrate	-
[10]	83 × 83	Multiservice RFID reader antenna	Centimeter band	0.91	0.86	7.55 dBi	-	FR4 substrate	85, 89%
[11]	98 × 98	Multiband Coil antenna	Centimeter band	2.45, 5.8	2.42-2.49,5.59-5.93	4.1dBi	-25dB	FR4 substrate	-
[7]	55 × 45	U-shaped micro strip	Centimeter band	4-4.5 GHz	7.83, 8.99	8.99dB	-43dB	RT/DUROID 5880	-
[9]	19× 25	F shaped	Centimeter band	2.40, 5.20, 5.80	0.76, 0.964, 0.80	1.5, 1.7, 3.05	$\leq$-20, $\leq$-42, $\leq$-30	FR4 substrate	$\geq$87
[15]	38 × 33.5	Planer end fire CP antenna	Centimeter band	5.8	0.11	4.9	-20	Printed FR4 substrate	63
[13]	178.04×178.04	Elliptical ring patch	UHF	0.8425, 0.92	0.025, 0.025,	4.5, 5.5	-19, -30	FR4 substrate	60

III. OBJECTIVES AND ORGANIZATION OF THIS RESEARCH WORK

The main contributions of this paper are mentioned below:

A. A microstrip patch antenna is proposed having resonance frequencies of 9.968 GHz, and 12.3 GHz to 13 GHz for wireless applications such as X-band, WiMAX, and RFID applications. The simulation of the proposed antenna is carried out using CST Microwave Studio and the simulated Gain, S11 (Return loss), VSWR, and the radiation pattern is presented.

B. The proposed antenna is implemented in hardware and the experimental radiation pattern is measured using WATS 2002 module. Finally, a comparison between simulated and experimental radiation pattern is reported in this research.

The rest of the paper is organized as follows. In Section IV, the design of microstrip patch antenna is designed. Simulation results on the return loss (S_{11}), voltage standing wave ratio (VSWR), directivity, radiation efficiency for the proposed antenna are presented in Section V. Finally, Section VI provides concluding remarks.

IV. ANTENNA DESIGN METHODOLOGY

Our research work is primarily focused on designing a microstrip patch antenna. The proposed antenna is formulated based on the empirical formula that helps to obtain desired geometry. This equation based designed results in the larger size of the antenna. As well as in most cases they may not provide excellent results. Hence, in order to achieve better results, further adjustments are required. After designing an antenna, we emphasized on antenna design with irregular geometry.

A. Antenna Design

Computer Simulation Technology (CST) Microwave Studio is used for antenna simulation and optimization. Commercially available FR4 substrate of thickness 2.5 mm is used on which the total structure is mounted. Dielectric constant (ϵ) is chosen 4.3 while loss tangent is 0.025. The opposite side of the substrate contains the copper plate that acts as a ground plane. This enhanced the directivity of the antenna to the desired level. Copper metal is also used as the radiating element. The thicknesses of both the radiator and ground plane are 0.313 mm. Plan view of the centimeter wave microstrip patch antenna is shown in "Fig. 1". Here, the proposed dual-band antenna is designed and optimized with first resonance frequency band at f_{r1} = 9.868 GHz and second resonance frequency band at f_{r2} = 12.3-13 GHz on

FR4 substrate. The overall ground plane dimension of the proposed antenna is 40 × 39 mm^2.

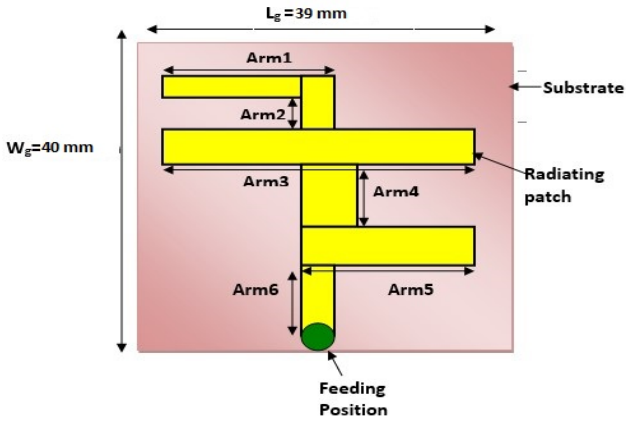


Fig. 1. Schematic diagram of the proposed antenna.

The antennas are designed based on the "(1)" – "(4)". In order to obtain the antenna's higher order Transverse Electromagnetic Modes (TEM) whose attributes are very closely matched with the fundamental mode, a technique is introduced to calculate the proposed antenna's length and width. This technique is commonly known as size extension method [10]. According to the size extension method, the extended patch antenna width 'W' and length 'L' can be expressed as below:

$$W = \frac{(2N+1)}{\sqrt{(2\varepsilon_r+1)/2}} \times (\frac{\lambda}{2}) \quad (1)$$

$$L = \frac{(2N+1)}{\sqrt{\varepsilon_{eff}}} \times (\frac{\lambda}{2}) - 2\Delta L \quad (2)$$

where λ is the proposed antenna's operating wavelength, ε_r is the relative permittivity (dielectric constant) and N is a positive valued integer number (in this antenna design we assume N =1). The physical dimensions of the microstrip patch antenna would look electrically wider for the fringing field effect. The extended length of the patch ΔL on each side is a function of antenna width to substrate height ratio (W/h) and the effective dielectric constant ε_{eff} [8]. So, ΔL and ε_{eff} are obtained by the following equations

$$\Delta L = 0.412h \frac{(\varepsilon_{eff}+0.3)(0.264+W/h)}{(\varepsilon_{eff}-0.3)(0.8+W/h)} \quad (3)$$

Where, effective dielectric constant,

$$\varepsilon_{eff} = \frac{\varepsilon_r+1}{2} + (\frac{\varepsilon_r-1}{2})(1+12\frac{h}{W})^{-\frac{1}{2}} \quad (4)$$

The initial length and width of the two antennas are obtained by taking 10 GHz resonance frequency. The optimized dimensional parameter values of the proposed dual-band antenna are shown in "TABLE. II".

TABLE II: GEOMETRICAL PARAMETERS OF ANTENNA STRUCTURE

Arm name	Length(mm)
Arm1	20.645
Arm2	5
Arm3	33.29
Arm4	8
Arm5	19.145
Arm6 (Feed line)	4
Patch thickness	0.313
Lg (Substrate Length)	39
Wg (Substrate Width)	40
h (substrate thickness)	2.5

"Fig. 2" and "Fig. 3" shows the front and back view of the hardware implementation of the proposed antenna. The hardware design of the antenna is done using toner transfer process.



Fig. 3. Back view of the proposed antenna hardware.

B. Antenna Performance Metrics

The performance of an antenna depends on several performance metrics. These metrics are S-parameter, Voltage Standing Wave ratio (VSWR), Half Power Beam Width (HPBW), antenna bandwidth, gain, directivity, and radiation efficiency. In this research work, the antenna designs are evaluated by the following metrics:

- *Scattering parameter (S-parameter):* S-parameters represent the amount of power reflected back from the antenna. It is also called return loss (RL). RL is related to reflection coefficient (Γ). "Equation (5)" and "(6)" are used to represent reflection coefficient and return loss of an antenna respectively.

$$\Gamma = \frac{Z_L - Z_{in}}{Z_L + Z_{in}} \quad (5)$$

$$RL = -20\text{Log}(\Gamma) \quad (6)$$

where Z_L is the load impedance (antenna impedance) and Z_{in} is the input impedance (or source impedance Z_S).

- *VSWR:* VSWR indicates impedance mismatch and higher the impedance mismatch, higher the value of VSWR. For effective radiation, the ideal value of VSWR is 1. But a VSWR value less than 2 is acceptable.

$$VSWR = \frac{1+|\Gamma|}{1-|\Gamma|} \quad (7)$$

- *Directivity (D):* The antenna directivity is defined as the ratio of the radiation intensity in a given direction from the antenna to the radiation intensity averaged over all directions.

$$D = \frac{4\pi U(\theta,\phi)}{P_{rad}} \quad (8)$$

Where D is the directivity (dimensionless), P_{rad} is the total radiated power in Watt, $U(\theta,\phi)$ radiation intensity (W/sr).

- *Gain (G):* Gain is closely related to the antenna directivity and can be defined as the ratio of the radiation intensity in a given direction to the radiation intensity that would be obtained if the power accepted by the antenna were radiated isotopically.

$$G = \frac{4\pi U(\theta,\phi)}{P_{in}} \quad (9)$$

- *Radiation Efficiency (η_{rad}):* Antenna radiation efficiency is related to the antenna gain and directivity.

$$\eta_{rad} = \frac{G}{D} \quad (10)$$

V. Simulated and Experimental Results of the Proposed Antenna

The designed antenna provides resonance at frequencies at = 9.868 GHz and second resonance frequency band at = 12.3-13 GHz. The excitation signal to the antenna is supplied by a 50Ω waveguide port using CST simulation software. From the "Fig. 4" it can be seen that the return loss at first resonance frequency is -20.49 dB and at the second resonance frequency is -10 dB. The proposed antenna provides 3 dB impedance bandwidth at first resonance frequency is 300 MHz ranging from 9.8 GHz to 10.1 GHz and -8 dB bandwidth at the second resonance frequency is 1 GHz ranging from 12.2 GHz to 13.2 GHz. The simulated VSWR curve of the proposed antenna is shown in "Fig. 5".

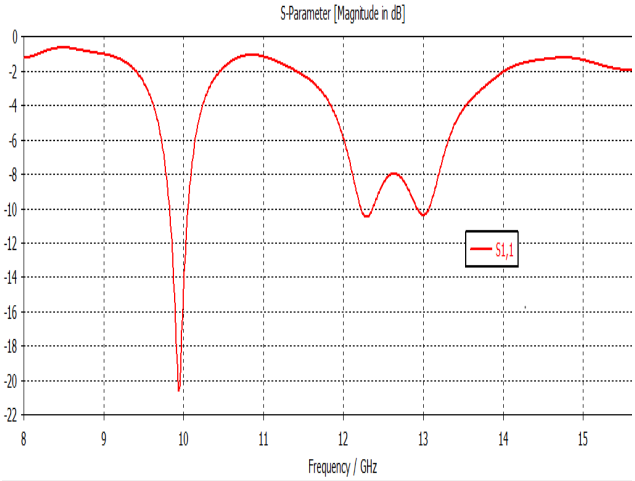


Fig. 4. Simulated return loss of the proposed rectangular patch antenna.

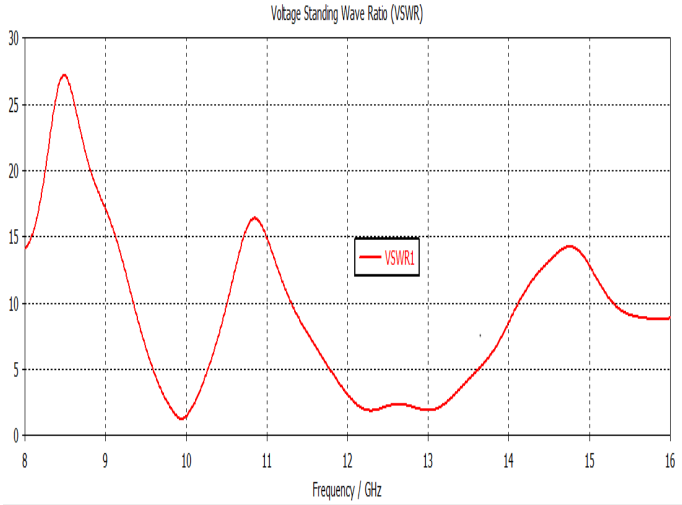


Fig. 5. Simulated VSWR of the proposed rectangular patch antenna.

From "Fig. 5" it is evident that at both the resonance frequencies the numeric value of VSWR is less than 1.5 that is effective for any wireless communication.

The gain and directivity of the proposed antenna at resonance frequency 9.968 GHz is 5.960 dB and 6.015 dBi respectively which is suitable for WiMAX, RFID, and X-band wireless applications. The 3D gain and directivity of the proposed antenna are shown in "Fig. 6" and "Fig. 7".

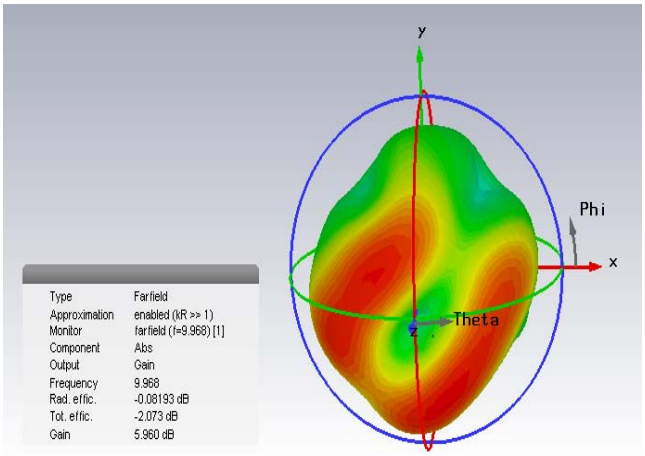


Fig. 6. 3D radiation pattern (Gain) of the proposed antenna.

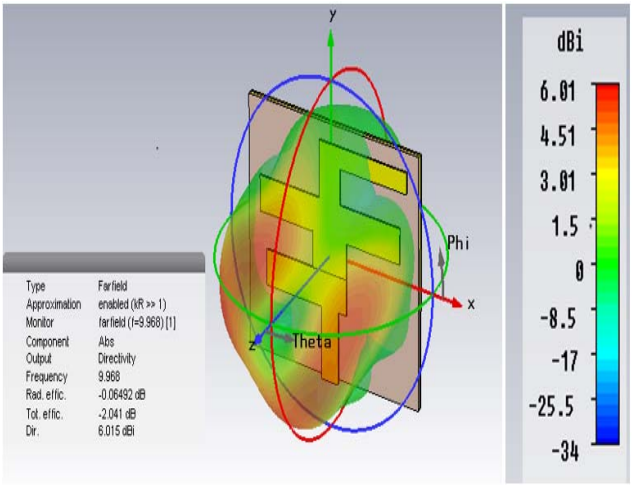


Fig. 7. 3D radiation pattern (Directivity) of the proposed antenna.

In a practical scenario, the VSWR is 1.1 to 2.5 and S-parameter range is less than -10 dB respectively although the ideal value of VSWR is unity. In our designed antenna the VSWR is less than 1.5 and S-parameter is less than -10dB. "Fig. 8" & "Fig. 9" shows the polar plot of radiation pattern (both gain and directivity) of the proposed antenna which indicates that the side lobe level is -7.6 dB, main lobe magnitude is 5.8dB and 3dB angular width is 40 degree at resonance frequency 9.968 GHz.

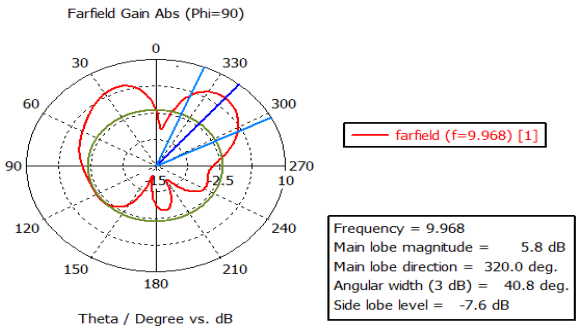


Fig. 8. The polar radiation pattern of Gain of the proposed antenna.

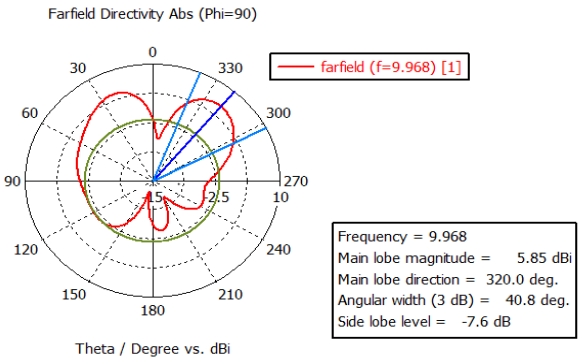


Fig. 9. The polar radiation pattern of directivity of the proposed antenna.

The antenna radiation efficiency is related to the gain and directivity and it can be written as $\eta_{rad} = G(dB) - D(dB)$. So, the antenna radiation efficiency at 9.968 GHz is almost 98 %. "Fig. 10" shows that experimental radiation pattern of the proposed antenna using WATS 2002 Module. To compare with the simulation radiation pattern shown in "Fig. 8" & "Fig. 9" is slightly different with the experimental result shown in "Fig. 10".

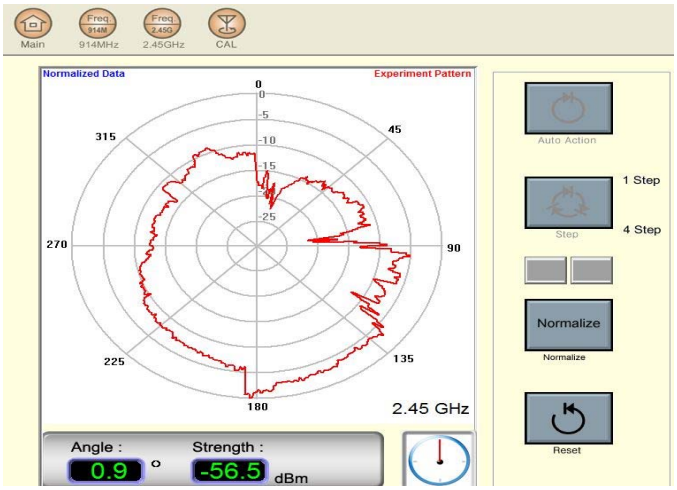


Fig. 10. Experimental radiation pattern of the proposed antenna using WATS 2002 Module.

The variation of the resonance frequency, VSWR, gain and return loss due to the variation of antenna arm length (Arm3) are shown in "Fig. 11".

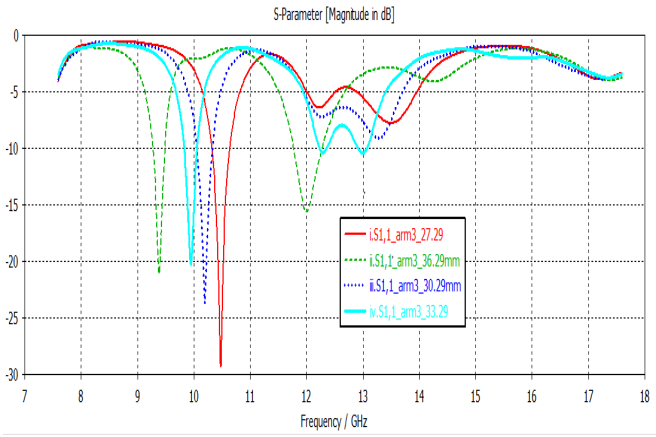


Fig. 11. Effect of length variation of Arm3 on the Resonance frequency and S-parameter.

TABLE III: EFFECT OF LENGTH ADJUSTMENT OF ARM 3

Length of Arm3(mm)	Resonant frequency (GHz)	Return loss (dB)	Gain(dB)	VSWR
33.29	9.97	-20.49	5.96	1.21
30.29	10.21	-23.33	6.18	1.15
27.29	10.49	-26.99	6.86	3.02
36.29	9.39	-21.05	6.61	1.19
	12	-15.56		1.40

From the "TABLE. III", it is seen that when the length of Arm3 decreases then resonance frequency as well as gain increases. It is noted that when Arm3 decreases from 30.29 mm to 27.29 mm the resonance frequency is sifted 10.21 GHz to 10.49 GHz and gain increases from 6.18 dB to 6.86dB. Also, we observed that when the length of Arm3 is increased to 36.29 mm then gain is slightly decreased but resonance frequency shifted to upper frequency than others. When the Amr3 is 36.29 mm, the proposed antenna provides two band; one at 12 GHz and other at 9.39 GHz. So, at this length of Arm3, the designed antenna acts as dual-band antenna. The S_{11} parameter, gain, VSWR at resonance frequency 12 GHz and 9.39 GHz becomes -21.05 dB, -15.56 dB, 6.61dB, and 1.40, 1.19 respectively.

TABLE IV: COMPARISON OF THE PROPOSED ANTENNA WITH THE LITERATURE

Reference Antennas	Size in mm^2	Operating Bands in GHz	Bandwidth in GHz	Gain (dB)
[12]	40×60	3.85,5.80,8.12	3.869,4.034	2.6
[8]	83 × 83	0.91	0.86	7.55
[11]	98 × 98	2.45, 5.8	2.42-2.49,5.59-5.93	4.1
[13]	178.04×178.04	0.8425, 0.92	0.025, 0.025	4.5, 5.5
[15]	38 × 33.5	5.8	0.11	4.9
Our Proposed antenna	**40 × 39**	**9.968, 12.3-13**	**0.30. 1.0**	**5.96**

The overall simulated results of our designed antenna are compared with the various reference antennas are presented in "TABLE IV". From the comparison table, we can see that our proposed design has better gain with less weight & dimension than some other existing antennas.

VI. CONCLUSION

In this contribution, we have analyzed and studied the characteristics of a rectangular microstrip patch antenna. The antenna is simulated on FR4 (lossy) substrate in CST simulation software and all the possible results are obtained using CST software. The proposed design is implemented in hardware and measure the radiation pattern using WATS 2002 module. The physical dimension of our designed antenna is 40×39 mm^2. The antenna provides high gain, high directivity, lower side lobe level and the standard amount of VSWR and S-parameter. Due to the variation of the length of Arm 3 provides the variation of resonance frequency with a variation of S-parameter, gain, VSWR and bandwidth. The proposed antenna provides all the required features within the simplest design possible. The results obtained from the designed microstrip antenna is useful in X-band, RFID and WiMAX applications. In the future, an array antenna will be designed using the proposed antenna to achieve higher gain and directivity.

REFERENCES

1. A. K. Evizal, T. A. Rahman, S. K. B. A. Rahim and M. B. Jamlos, 2011. "A MULTI BAND MINI PRINTED OMNI DIRECTIONAL ANTENNA WITH V-SHAPED FOR RFID APPLICA-TIONS," Progress In Electromagnetics Research B, 27, pp.385-399, 2011.
2. W. T. Li, Y. Q. Hei, W. Feng and X. W. Shi, "Planar Antenna for 3G/Bluetooth/WiMAX and UWB Applications With Dual Band-Notched Characteristics," in IEEE Antennas and Wireless Propagation Letters, vol. 11, pp. 61-64, 2012.
3. L. Dang, Z. Y. Lei, Y. J. Xie, G. L. Ning and J. Fan, "A Compact Microstrip Slot Triple-Band Antenna for WLAN/WiMAX Applications," in IEEE Antennas and Wireless Propagation Letters, vol. 9, pp. 1178-1181, 2010.
4. N. Prema, and A. Kumar "Design of multiband microstrip patch antenna for C and X band."optic, vol.127(20), pp. 8812-8818, Oct 1 2016.
5. Y. Liu, L.-M. Si, M. Wei, P. Yan, P. Yang, H. Lu, C. Zheng, Y. Yuan, J. Mou, X. Lv, and H. Sun, "Some Recent Developments of Microstrip Antenna," International Journal of Antennas and Propagation, vol. 2012, Article ID 428284, pp. 1-10 , 2012.
6. Y. Dong, H. Toyao and T. Itoh, "Design and Characterization of Miniaturized Patch Antennas Loaded With Complementary Split-Ring Resonators," in IEEE Transactions on Antennas and Propagation, vol. 60, no. 2, pp. 772-785, Feb. 2012.
7. K. N. Lal and A. K. Singh, "Modified design of microstrip patch antenna for WiMAX communication system," Proceedings of the 2014 IEEE Students' Technology Symposium, Kharagpur, pp. 386-389, 2014.
8. N. Sarker , M. A. Islam, and M. R. H. Mondal. "Two novel multiband centimetre-wave patch antennas for a novel OFDM based RFID system." Journal of Communications 13.6 (2018).
9. A. K. Gautam, L. Kumar, B. K. Kanaujia and K. Rambabu, "Design of Compact F-Shaped Slot Triple-Band Antenna for WLAN/WiMAX Applications," in IEEE Transactions on Antennas and Propagation, vol. 64, no. 3, pp. 1101-1105, March 2016.
10. J. K. Pakkathillam, M. Kanagasabai and M. G. N. Alsath, "Compact Multiservice UHF RFID Reader Antenna for Near-Field and Far-Field Operations," in IEEE Antennas and Wireless Propagation Letters, vol. 16, pp. 149-152, 2017.
11. A. Sharma, I. J. G. Zuazola and A. Perallos, "Multipurpose Near- and Far-Field Switched Multiband Coil Antenna for 915-MHz/2.45/5.8-GHz RFIDs," in IEEE Antennas and Wireless Propagation Letters, vol. 16, pp. 2562-2565, 2017.
12. C. Varadhan, J. K. Pakkathillam, M. Kanagasabai, R. Sivasamy, R. Natarajan and S. K. Palaniswamy, "Triband Antenna Structures for RFID Systems Deploying Fractal Geometry," in IEEE Antennas and Wireless Propagation Letters, vol. 12, pp. 437-440, 2013.
13. Z. Wang, R. She, J. Han, S. Fang and Y. Liu, "Dual-Band Dual-Sense Circularly Polarized Stacked Patch Antenna With a Small Frequency Ratio for UHF RFID Reader Applications," in IEEE Access, vol. 5, pp. 15260-15270, 2017.
14. Y. Yu, J. Ni, and Z. Xu, "Dual-band Dipole Antenna for 2.45 GHz and 5.8 GHz RFID Tag Application", AEM, vol. 4, no. 1, pp. 31-35, Jun. 2015.
15. W. Zhang, P. Cheong, W. Lu and K. Tam, "Planar Endfire Circularly Polarized Antenna for Low Profile Handheld RFID Reader," in IEEE Journal of Radio Frequency Identification, vol. 2, no. 1, pp. 15-22, March 2018.

Traffic Load Based Efficient Energy Management Technique for 5G Small Cell Networks

Syed Zahidur Rashid[1]
Dept. of Electronic and Telecommunication Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
szrashidcce@yahoo.com

Abdul Gafur[2]
Dept. of Electronic and Telecommunication Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
agafur_cox@yahoo.com

Atikur Rahaman[3]
Dept. of Electronic and Telecommunication Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
rahmanatik171993@gmail.com

Yaheya Solaiman[4]
Dept. of Electronic and Telecommunication Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
mdriad290@gmail.com

Erfanul Hoque Bahadur[5]
Dept. of Computer Science & Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
ehb1729@gmail.com

Abstract—There will be high energy wastage and CO_2 emissions will be increased in 5G network because of dense deployment of small cell base stations. This paper brings an efficient energy management technique for 5G green small cell network which switches on as such number of antennas from the array as needed. This is done by monitoring and counting the amount of data traffic within the area of the small cell. It is shown that the proposed scheme is capable to improve energy saving significantly. We have found that the scheme consumes up to 96.875% less energy than without energy management technique. Besides, the emission of CO_2 is reduced by 2.3166kg per day when the small cell is operated by a diesel power source and for all kinds of power sources on average CO_2 emission can be reduced by 1.612 kg per day.

Keywords— Energy efficiency; 5G Green network; CO_2 emissions; massive MIMO; Small cell network.

I. INTRODUCTION

It has been reported that [1], the mobile data traffic is globally increased nearly 11-fold within 2013 to 2018. Now it is reached to 15.9 exabytes per month approximately. The only reason for it, is the gradually increasing number of devices that are mobile connected. Currently it is about 1.4 per capita. Hence, the demand of mobile data is growing very rapidly. So to say, the one and only solution to solve such demand is deploying small cells. It is a matter of concern that, there is high data traffic density in 5G small cell network(SCN). To improve the capacity also, deploying dense small cell is mandatory. Actually it not only helps to improve the capacity, but also it helps to reuse the radio resources. Again, it helps to reduce the base station transmit power. Thus, it helps to fulfill the demand of next generation network by improving both spectral efficiency and energy efficiency.

In accordance with [2] and [3], one of the promising solutions of 5G networks to cope with the sustaining trend of ever growing mobile data traffic is to deploy dense small cell base stations. A small cell is a low-cost radio Base Station which operates at low-power and provides better cellular coverage. The cell radius of the small cell can range from tens of meters to a few hundred meters. Because of a short distance between the Base Station and the user equipment(UE) the transmit power is reduced significantly in small cell network.

Small cell can be deployed in places of high user density, i.e. stadiums, pedestrian malls, exhibition centers, airports and so on. But, at off-peak time like as, nights, the user density is decreased noticeably. So, we do not need to activate all the antennas of the array at that time.

To develop a green network, the energy consumption of base stations must be reduced. Because 56% power consumption of cellular network is from base stations [4]. This great amount of power consumption increases the amount of CO_2 in the environment. Although, the amount of CO_2 emitted from cellular network industry is relatively meager, it will be increased in the next generation mobile network. Because, there will be a very large number of base stations in the network. Therefore, developing a green network is dependent on decreasing the power radiation from the future small cell base stations which will be densely deployed.

For boosting the network throughput and to save energy at low traffic hours, it is highly required to develop efficient cell management technique and user access mechanisms for the dense deployments of small cell networks. Because, the active hardware not only consumes power, but also increases the amount of CO_2 in the environment.

Data traffic is varied because of the mobility of the user. Hence, the demand of the antenna changes continuously. So,

obviously there should be some techniques to manage the cell for saving energy.

II. LITERATURE REVIEW

Different techniques are brought to achieve energy efficiency. In [5], we have seen that, the coverage areas of some stations are adjusted and other base stations are turned off. They achieved 40% less power consumption in the sleeping base stations. In another study [6], the base stations are switched to different sleep modes step by step, where less power is consumed to deeper sleep modes. They could save 92% energy in empty network. In the study of [7], we have seen that, they did the load analysis. On that basis, they have chosen sleep mode technique and thus found energy efficiency. They achieved nearly 22% and 38% improvement in network throughput without and with delay respectively.

In [8], the authors divided the geographical area into different grids. To give satisfactory service to all the users at the peak hours, they selected maximum number of small cell base stations in every grid. On the other hand, required number of base stations are kept on and others are turned off in the idle periods. They saved up to 53% and 23% energy in the dense and uncrowded area respectively.

19% energy efficiency is achieved in [9] compared to a scheme where the Base Stations run with a set number of antennas. In this study the authors used a game theoretic framework. In another study [10], the authors proposed minimum 40% energy efficiency and QoS improvement. They used the combination of an adaptive discontinuous transmission and antenna optimization.

This paper deals with an unaccustomed traffic load based energy management technique where the base stations decide how much antennas will be activated according to the total number of user equipment. In every 3 seconds the base station will count the number of users in its monitoring area(M_A) and required number of antennas will be activated. So, this scheme is reducing the wastage of energy by small cell base station. As a result, the small cell base stations can form an adaptive, stable and energy efficient small cell network. The energy saving attributes and CO_2 reduction characteristics of the technique are then estimated and compared with a scheme with no energy management.

The remainder of this paper is organized as follows: The model of the system is defined in Section III. The energy management technique is described in Section IV. Section V presents the results and analysis. Comparison with similar research works is shown in Section VI. Finally, the work is concluded in Section VII.

III. MODEL OF SYSTEM

A. Network Architecture

Basically, there will be four kinds of small cell in future 5G network, namely, femto-cell, pico-cell, micro-cell and metro-cell. We have worked with pico-cells. Pico-cells can be deployed both indoors and outdoors, i.e., stadiums, pedestrian malls, exhibition centers, airports and so on. Here, each pico-cell contains 128 (16*8 array) massive MIMO antennas in its array. For applying beamforming technology, a dedicated antenna will be allocated for each user. So, the maximum number of users that can be supported by the cell is 128. The energy consumption of each antenna is approximately 1.5625Wh [11]. Hence, the energy consumption of all the antennas is 200Wh. The monitoring radius of the cell is 50 meters or less [12]. A pico-cell base station is located at the center of figure 1.The User Equipments (UE- the dots in the figure) are arbitrarily distributed within and out of the monitoring area. The monitoring area(M_A) is depicted as the circle shown in figure 1. The user equipments that are within the monitoring area will only be served by the base station.

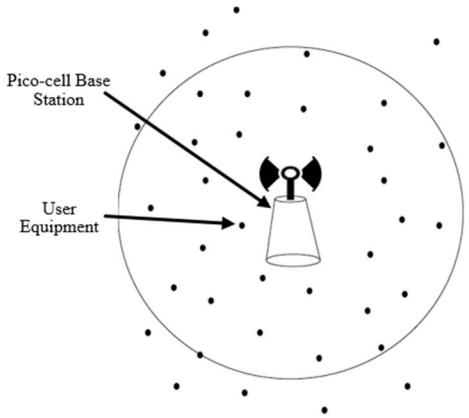


Fig. 1: Proposed Model

B. Model of Path Loss

By using WINNER II model, which is given in [13] the path loss between the small cell base station and the UE is calculated. Assuming the SCNs with high population density such as train stations or airports, we have chosen model B3 as shown below:

$$PL_{LOS} = 13.9\log_{10}(r) + 64.4 + 20\log_{10}(f_C/5) + SF_{LOS} \quad (1)$$

$$PL_{NLOS} = 37.8\log_{10}(r) + 36.5 + 23\log_{10}(f_C/5) + SF_{NLOS} \quad (2)$$

$$P_{LOS} = 1 \quad ; \text{if } r \leq 10 \quad (3)$$

$$P_{LOS} = \exp(-(r-10)/45) \quad ; \text{if } r > 10 \quad (4)$$

where path loss for line of sight (LOS) is PL_{LOS} (in dB) and the path loss for non-line of sight (NLOS) is PL_{NLOS} (in dB). Again the parameter r represents the distance between the small cell base station and the user equipment, f_C represents GHz level of carrier frequency. The shadow fading for LOS and NLOS are SF_{LOS} and SF_{NLOS} and valued them 3 dB and 4 dB respectively. Finally, the probability of having LOS propagation condition is represented as P_{LOS}.

Here, different path losses will be resulted for line of sight and non-line of sight communications. Obviously, the path loss for non-line of sight will be more, and in 5G, it will be a mandatorily considerable factor. Because, as it is known that penetration is increased with increasing frequency.

IV. ENERGY MANAGEMENT TECHNIQUE

The small cell has a circular monitoring area M_A with a radius of R. Here, the monitoring radius is 50 meters. The distance between the base station and the UE is $D_{c,x}$. The small cell base station will estimate the distance of a UE by measuring the power because main cause of power attenuation is path loss which is a function of distance as given in equation (1), (2), (3) and (4). The path loss can be shown by a function, f(D).

While transmitting data, a UE_x sends a reference signal, P_{ref} periodically. The small cell base station, receiving this signal will record the receive power $P^{(r)}_{c,x}$, which is similar to recording the $D_{c,x}$ due to the relationship given below [5]:

$$P^{(r)}_{c,x} = P_{ref} - f(D_{c,x}) \quad (5)$$

Where $P^{(r)}_{c,x}$ and P_{ref} are in dBm corresponding to $f(D_{c,x})$ in dB.

In this way, the small cell base station measures whether the UE_x is within its monitoring area. If the UE_x is within the monitoring area, it will forward the UE_x into its monitoring list. And finally, the base station counts the number of users in its cell area.

Here, the users that are within the cell area will only be supported by the base station. The period of measuring reference signal by the base station sent by the user equipment is considered to be 3 seconds. It is thought-out that the number of users will not be changed significantly within the cell area within this period.

A. Working Principle by Pseudo Code

By using the equation (5) the base station calculates the path loss and then by using equations (1), (2), (3) and (4) the base station will calculate the distances of the UEs. After that it will count the total number of user equipments in the cell area. And then, it will activate antennas according to the number of users. Because of keeping the blocking probability low, the number of active antennas are more than the number of users. So there are some extra active antennas for the upcoming possible users. It will continuously do the same thing in every 3 seconds.

The algorithm of the model's working principle is illustrated below in the pseudo code:

Here, estimated_time_duration denotes the period for which the pico-cell is deployed.

```
time1 := starting_time()
FOR time() time_i ε estimated_time_duration do
  time2 := time()
  IF time_difference(time2, time1) is equal to 3 seconds
    scan energy_database from buffer memory
    calculate distance and push to distance_database
    IF distance[i] is less than range
      push the user in user database
      calculate total permitted waiting user number
      IF user number is 0
        provide array (2*2)
      ELSE IF user number is 1 to 2
        provide array (2*4)
      ELSE IF user number is 3 to 5
        provide array (2*6)
      ELSE IF user number is 6 to 10
        provide array (2*8)
      ELSE IF user number is 11 to 19
        provide array (3*8)
      ELSE IF user number is 20 to 27
        provide array (4*8)
      ELSE IF user number is 28 to 35
        provide array (5*8)
      ELSE IF user number is 36 to 43
        provide array (6*8)
      ELSE IF user number is 44 to 51
        provide array (7*8)
      ELSE IF user number is 52 to 59
        provide array (8*8)
      ELSE IF user number is 60 to 67
        provide array (9*8)
      ELSE IF user number is 68 to 75
        provide array (10*8)
      ELSE IF user number is 76 to 83
        provide array (11*8)
      ELSE IF user number is 84 to 91
        provide array (12*8)
      ELSE IF user number is 92 to 99
        provide array (13*8)
      ELSE IF user number is 100 to 107
        provide array (14*8)
      ELSE IF user number is 108 to 115
        provide array (15*8)
      ELSE user number is 116 to 128
        provide array (16*8)
      END IF
    END IF
  ELSE difference (time2, time1) not equal 3 seconds
  time2 := time()
  END FOR
```

V. RESULTS AND ANALYSIS

For showing the effectiveness of our proposed scheme, we have compared it with the scheme without energy management technique. We have analyzed this technique with respect to a pico cell. We considered 128 (16*8 array) massive MIMO antennas in a pico cell. In a cell with no energy management technique, all the antennas are kept on at all the time. But our technique will activate as such number of antennas which are actually required by the surrounding user equipments. We have gained a good amount of energy saving and CO_2 emission is also minimized.

The number of active links during a day are approximated on rational basis. In figure 2, horizontal axis represents hours of a day and vertical axis represents number of active links. We can see from figure 2 that, the number of active links may become zero during off-peak hours. There is considerable variation in the number users during a day because of user mobility.

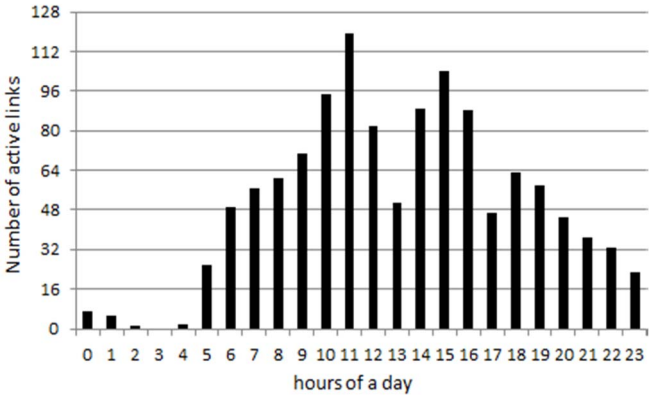


Fig. 2 Number of active links against hours of a day

In figure 3, horizontal axis represents hours of a day and vertical axis represents required number of antennas. Here, the comparison between the proposed scheme and the scheme without energy management is shown. In the fixed topology, 128 antennas are kept on continuously. On the other hand, the number of active antennas are adapted according to the number of users in the proposed scheme. Antennas will be activated according to the pseudo code described in Section III. We can see from figure 2 that in most of the time of a day, the number of user is less than its maximum value. That's why it ensures energy saving significantly.

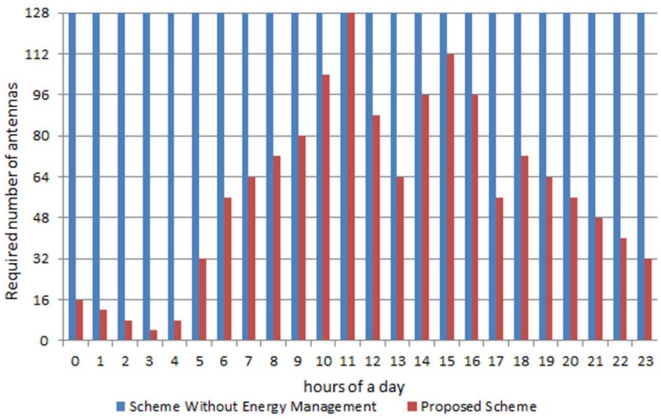


Fig. 3: Required number of antennas against hours of a day

In figure 4, horizontal axis represents hours of a day and vertical axis represents required energy in Watt-hour. Here, the energy consumption of the base station during a day is shown. The base station consumes 200Watt-hour of energy in every hour if there is not any energy management. But, the energy consumption is varied during a day with the variation of the number of users in the suggested scheme. A very small amount of energy is consumed during off peak time. A big amount of energy is saved and energy wastage is reduced by a great amount therefore. Thus maximum energy saving during off-peak time and average energy saving during a day are calculated by,

$$ES_{OPT} = (CE_{WNEM} - CE_{WEM}) / CE_{WNEM} \quad (6)$$

$$ES_{AVG} = (\sum CE_{WNEM} - \sum CE_{WEM}) / \sum CE_{WNEM} \quad (7)$$

Here, CE_{WNEM} and $\sum CE_{WNEM}$ are representing cell energy consumption at off-peak time and total cell energy consumption during a day respectively with no energy management. CE_{WEM} and $\sum CE_{WEM}$ are representing cell energy consumption at off-peak time and total cell energy consumption during a day respectively with energy management. We can notice from figure 4 that, our proposed scheme can achieve 96.875% energy saving during off-peak time and 54.167% energy saving on average during a day.

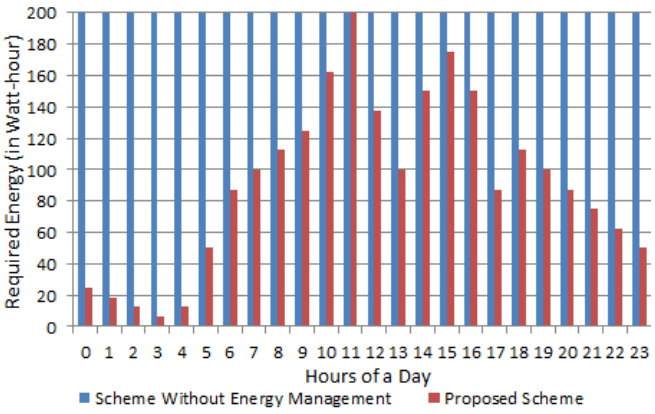


Fig. 4: Required Energy (in Watt-hour) against hours of a day

In figure 5, horizontal axis represents hours of a day and vertical axis represents the amount of CO_2 in kg. Here, amount of CO_2 emission in the proposed scheme is compared with the scheme without energy management.

Because of dense deployment of small cell base stations, the CO_2 emissions will be considerably increased. So, it will be a threat to the environment and to reduce the amount of CO_2 emitted from base station is a major challenge to develop a green network. In India the CO_2 emission from a base station is 0.000891kg for 1W of power radiation when a diesel engine is used as the power source. Maximum CO2 reduction during off-peak time and average CO_2 reduction during a day are calculated by,

$$CR_{OPT} = (CE_{WNEM} - CE_{WEM}) / CE_{WNEM} \quad (6)$$

$$CR_{AVG} = (\sum CE_{WNEM} - \sum CE_{WEM}) / \sum CE_{WNEM} \quad (7)$$

Here, CE_{WNEM} and $\sum CE_{WNEM}$ are representing CO_2 emission at off-peak time and total CO_2 emission during a day

respectively with no energy management. CE_{WEM} and $\sum CE_{WEM}$ are representing CO_2 emission at off-peak time and CO_2 emission during a day respectively with energy management. Without energy management technique, the amount of CO_2 emitted from the base station is 4.2768kg per day. But, only 1.9602kg CO_2 will be emitted by proposed scheme per day. Hence, this technique is offering a good amount of CO_2 reduction.

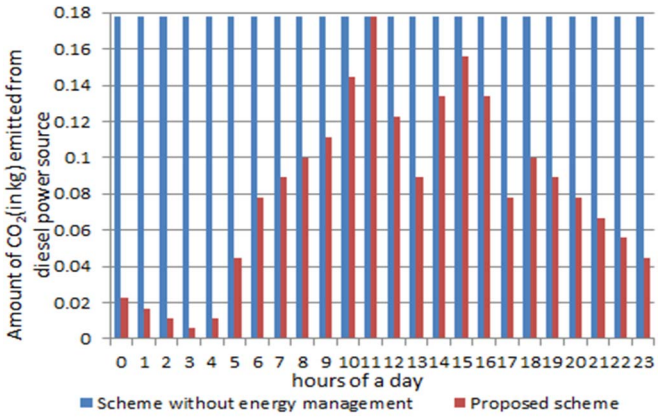


Fig. 5: Amount of CO_2 (in kg) emitted from diesel power source against hours of a day

In the United States of America, the average amount of CO_2 emitted per KWh is about 0.62 kg. Without energy management technique, the amount of CO_2 emitted from the base station is 2.976kg per day. But, only 1.364kg CO_2 will be emitted by proposed scheme per day. Thus this technique is helping to develop a green network.

VI. Comparison Table

Research Works	Approach used	Outcomes
Our Proposed work	Switching antennas according to user number by proper coding.	Up to 96.875% less energy consumption in empty network and 54.167% energy saving on average.
Ref [6]	Base stations are switched to different sleep modes step by step.	They could save 92% energy in empty network.
Ref [7]	They did the load analysis and have chosen sleep mode technique.	Achieved nearly 22% and 38% improvement in network throughput without and with delay respectively.
Ref [9]	They used a Game theoretic framework.	19% energy efficiency is achieved.
Ref [10]	They used the combination of an adaptive discontinuous transmission and antenna optimization.	Achieved minimum 40% energy efficiency.

VII. Conclusion

This paper introduced an efficient energy management technique for small cell networks based on traffic load. It is evident that it can save significant amount of energy. Here, as such number of antennas will be activated as needed. It offers up to 96.875% less energy consumption in empty network and 54.167% energy saving on average than with no energy management technique. Besides, our proposed scheme can reduce the emissions of CO_2 from small cell base stations. CO_2 emission is reduced from 4.2768kg to 2.9602kg in the suppositional scheme per day when the small cell is operated by a diesel power source. For all kinds of power sources in average CO_2 emission can be reduced by 1.612 kg per day. Moreover, our proposed algorithm can be utilized for any massive MIMO antenna system. So, this technique is ancillary to develop an efficient, stable and green network.

References

[1] C. V. N. I. Cisco, "Global Mobile Data Traffic Forecast Update, 2013-2018 (White Paper, 2014)," ed.

[2] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, et al., "Network densification: the dominant theme for wireless evolution into 5G," IEEE Communications Magazine, vol. 52, pp. 82-89, 2014.

[3] T. Q. Quek, G. de La Roche, I. Güvenç, and M. Kountouris, Small cell networks: Deployment, PHY techniques, and resource management: Cambridge University Press, 2013.

[4] C. Han, T. Harrold, I. Krikidis, I. Ku, T. A. Le, S. Videv, et al., "Green radio: radio techniques to enable energy efficient wireless networks," IEEE Communications Magazine Special Issue: Green Communications, vol. 49, pp. 46-54, 2011.

[5] Z. Li, D. Grace, and P. Mitchell, "Traffic perception based topology management for 5G green ultra-small cell networks," in 2014 1st International Workshop on Cognitive Cellular Systems (CCS), 2014, pp. 1-5.

[6] P. Lähdekorpi, M. Hronec, P. Jolma, and J. Moilanen, "Energy efficiency of 5G mobile networks with base station sleep modes," in 2017 IEEE Conference on Standards for Communications and Networking (CSCN), 2017, pp. 163-168.

[7] H. Celebi and İ. Güvenç, "Load analysis and sleep mode optimization for energy-efficient 5G small cell networks," in 2017 IEEE International Conference on Communications Workshops (ICC Workshops), 2017, pp. 1159-1164.

[8] C. Peng, S.-B. Lee, S. Lu, and H. Luo, "GreenBSN: Enabling energy-proportional cellular base station networks," IEEE Transactions on Mobile Computing, vol. 13, pp. 2537-2551, 2014.

[9] M. A. Hossain, C. Cavdar, E. Bjornson, and R. Jantti, "Energy-efficient load-adaptive massive MIMO," in 2015 IEEE Globecom Workshops (GC Wkshps), 2015, pp. 1-6.

[10] W. Pramudito, E. Alsusa, D. K. So, and K. A. Hamdi, "Load aware adaptive scheduling for energy efficient suburban Massive MIMO networks," in 2015 IEEE Global Communications Conference (GLOBECOM), 2015, pp. 1-6.

[11] K. Chen, C. Cui, Y. Huang, B. Huang, J. Wu, S. Rangan, et al., "C-RAN: A green RAN framework," in Green Communications: Theoretical Fundamentals, Algorithms and Applications, ed: CRC Press, 2013, pp. 279-304.

[12] T. Nakamura, S. Nagata, A. Benjebbour, Y. Kishiyama, T. Hai, S. Xiaodong, et al., "Trends in small cell enhancements in LTE advanced," IEEE Communications Magazine, vol. 51, pp. 98-105, 2013.

[13] Y. d. J. Bultitude and T. Rautiainen, "IST-4-027756 WINNER II D1. 1.2 V1. 2 WINNER II Channel Models," EBITG, TUI, UOULU, CU/CRC, NOKIA, Tech. Rep, 2007.

GDPR Compliance: Implementation Use Cases for User Data Privacy in News Media Industry

Ahmedur Rahman Shovon*, Shanto Roy†, Arnab Kumar Shil‡, and Tanjila Atik§
Institute of IT*†§, Department of CSE‡, Jahangirnagar University, BD
Department of CSE†,Green University of Bangladesh, BD
Department of CSE§, Daffodil International University, BD
Cefalo Bangladesh Limited*‡, BD
Email: shovon.sylhet*@gmail.com, shantoroy†@ieee.org, me‡@ruddra.com, tanjila.tanji15§@gmail.com

***Abstract*—The paper presents implementation use cases towards the consequences of maintaining user data privacy after the adoption of GDPR; specifically in the news media industry. General data privacy regulation (GDPR) is a European Union general data protection regulation adopted in 2016 subjected to protect personal data of the citizens in the EU. Besides, it implies to particular restrictions and obligations for handling user data by different companies or organizations. However, although the rule is applicable only if EU citizens are involved, all the companies started adopting the preparation and practice to maintain compliance with GDPR. In this paper, we identify and present the system design and implementation use cases for the news media industry that is compliant with the new regulations. The use cases indicate and explain significant transformation required in user data management process according to the GDPR.**

Keywords- General Data Protection Regulation, GDPR and Compliance, Privacy/Policy by design

I. Introduction

In order to restrict the use of user data by media or service industry, General Data Policy Regulation (GDPR) has been introduced. GDPR defines some particular data usage policies based on certain criteria that must be fulfilled by the developers so that a system preserves it's user rights. GDPR came in front with a view to solving the issue regarding user data misuse without the consent of unaware European Union citizens. However, as every system has worldwide user access including EU citizens, all companies had to change their user data access policies across the world to follow GDPR compliance.

News media industries generally partake various user information by selling subscriptions to a wide range of users. Therefore, it raises the question of how they are managing all these data and using further for advertisement or latter analytics purpose. Now, GDPR is compelling the industry to expose how they are managing all the user data so that the whole process of reusing data remains transparent to users. In addition to that, while mailing offers or showing adds, they have to act carefully in order to avoid user's right violation.

As the data collection and maintenance depend on the user consent, there should be a proper way to communicate between the industry and user's end through an easy to access user dashboard. Applications require to be developed keeping in mind that users should be aware of where, when and how their data is collected and used for further analytics purposes.

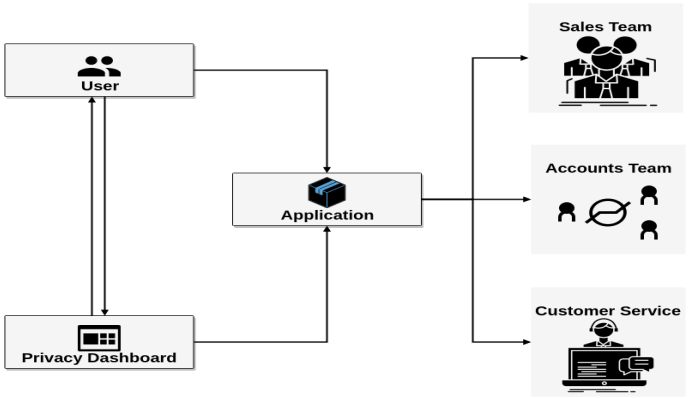


Fig. 1. Privacy by design and GDPR

Therefore, the industries end up enforcing transparent data usage policies that require user permission to use the data for future analysis. This also ensures that if the owner of data or user wants to delete his information, the industries will be responsible to erase that particular data entirely from their database. This is how GDPR is leading towards a policy/privacy by design solution that comes with options towards users for better accessibility, understandings, and management of their owned data. A simple policy/privacy by design with GDPR compliance is shown in **Figure 1**.

In this paper, we designed and developed a model framework that describes the initiation of a user dashboard. The dashboard is a better media of communication and understanding between the industry and the users. Users can read the policies provided by the company and therefore assign the data access rights through their consents. In addition to that, we provided the implementation use cases in the news media industry. We also identified the required transformation of the present system to meet the GDPR compliance.

The rest of the paper is as follows: **Section II** presents the issues and summary guidelines of GDPR compliance. **Section III** identifies and summarizes the implementation use cases of GDPR compliance. Then **Section IV** presents the design and workflow of the news media industry system followed by the consequence discussion in **Section V** and a few related work summaries in **Section VI**.

II. Background and rationale

A. Principle rules of GDPR

There are primarily six principles related to data subject rights that comprise breach notification (Article 33) [1], right to access (Article 15) [2], right to be forgotten (Article 17) [3], data portability (article 20) [4], privacy by design (Article 25) [5] and data protection officer (Article 39) [6]. The principles basically describes the rights a data subject should have on basis of new regulations. Therefore, if there is a data breach the owner of data must be notified as soon as possible. Data subject will need to have the right to know whether or not his particular portion of data is being used for any purpose; and somehow if he wants to erase his previous data, processor or service provider will not be able to store that information further. Another aspect is, the introducing of data portability where the data subject can move or transmit his data from one data controller to another. Privacy by design is not a jargon anymore, it is being applied to different sectors including big data analysis [7] to crisis management using social media [8]. Considering the regulations, controllers need to minimize the amount of data by deducting additional data as well as restricting the use of personal data during data processing (Article 23) [9]. In addition, controllers need to notify the data processing activities along with the local data protection act (DPA) in order to grab approval from the authority.

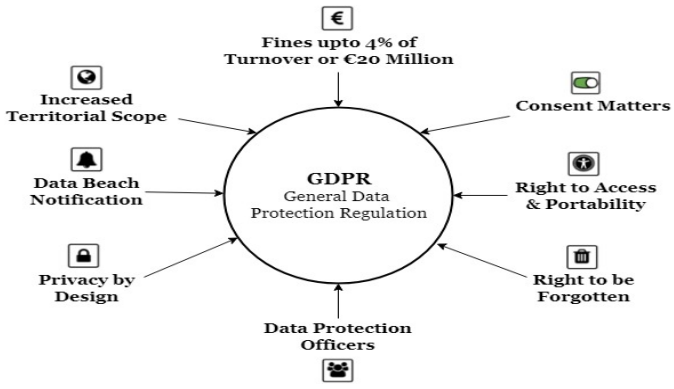


Fig. 2. General Overview of GDPR

Therefore, considering the rights of the data subject, controller or service providers have to perform minimum changes in their user policies and interfaces. Prior to the rules, in privacy program management (PPM) and privacy enterprise management (PEM), system engineering used to exclude consideration regarding user's data behavior. GDPR changed the view and directed it towards more data subject centric decisions to ensure privacy and protection for personal data.

B. Personal data protection

GDPR introduces the protection of data in personal level in the updated regulation policies applicable for only European Union citizens.

C. User consent centric Privacy by Design

Another principle introduced by GDPR is the privacy by design. It refers to the incorporation of the privacy practices from the initial design to the final specification of a system [10]. Privacy by default means that strict privacy settings will be automatically applied upon running or using the system for the first time. GDPR further emphasizes to provide enough protection mechanisms while the personal data of EU citizens is transferred to other countries and requires both a data controller and processor to take the liability.

D. Adoption to the new regulations

GDPR has far more influences over the policies and user experiences beyond regional boundaries. Although the rules are applied for ensuring the data protection across Europe, companies, and service providers have been taking initiatives to amend all the user policies as giant companies like Facebook and Google gave their commitment [11] [12]. But do every company need to take initiative for that? The answer is 'no'. The very first question that comes before deciding is whether or not the company sell products to individuals. Another question is whether or not the company keeps or collects personal data. If the answer to any between these two questions is- yes, the company need to measure whether their clients are from the EU or not. If the potential client is from the EU, the company must follow the regulations to avoid punishments. Since people across the world access to the news pages throughout the day, they register to access news articles and pay for subscription along with providing personal sensitive data. Therefore, news media must adopt the GDPR to establish synchronized user policies all over the world. While adopting, there are several things to keep in mind that include educating employees about the rules, revising user policies, reviewing personal data consent requests, managing and securing data, implementing privacy by design for data, performing data protection by hiring an officer, analyzing third-party risks, etc.

According to the new GDPR, users must have been informed about all the data practices to be done with their personal data. Also, users must have the opportunity to hold their right to give and at the same time to withheld their consent if they disagree to any of the usages of their personal data [13]. However, most of the time the privacy policies are written in an unclear and complicated way [14] and because of the policies being too long, users find it difficult and are unwilling to read the whole content [15].

GDPR aims to improve transparency [16] in terms of the information collected from the data subjects and encourages the use of visualization along with the requirement for using clear and simple language (Art 58, GDPR). To get compliant to this legislation, lately, the Usable Privacy Policy Project [17] in the US has given a try to represent the policy notices in a machine-readable way to display the relevant content in a more user-friendly manner.

Another important issue is to assess the users whether they have understood everything they read on the privacy notice

or not. In this paper distinct concepts of GDPR are discussed and which information must be given to the users while getting consent are extracted. Furthermore, a comparison was made with the old data protection directive 95/46/EC [18].

Some approaches are suggested here related to the notice:

- Firstly, all the data practices and their implementation procedures must be provided.
- A demo approach on how to read the document step by step can be included.
- Different colors and fonts can be used for different topics to keep the users attention.
- For critical legal notations, hyperlinks and pop-up windows can be provided containing illustration and examples related to those notations. These can be accomplished by using Akoma Ntoso mark-up.
- For assessing the proper understanding of the policies by the users, a consent form can be designed where clickable icons can be used to gather explicit consent.
- After taking the consent, same icon can be presented to the users and asked to connect different given icons according to its meaning.

This approach can seem to be annoying thus, needs to be refined further and tested on different categories of users.

III. Implementation use cases

According to the statement of [19], privacy is breached when personal information is available outdoor its supposed context. [20] prescribed some approaches to maintain privacy by utilizing data minimization which may limit the data controllers task of processing and collecting personal data. Giving users control over their own personal data to decide who they want to share those data and in what condition, can strengthen their privacy protection.

The natural persons whose personal data(PD) is processed by a processor or controller are identified as data subject(DS). GDPR compliant system should ensure the rights of Data Subjects(DS) which includes:

- **Right to be informed**: Provide the information listed in Art. 13 [13] if the personal data is supplied by data user. Otherwise, provide information listed in Art. 14 [13].
- **Access right**: Provide access to personal data and information listed in Article 15 [2].
- **Rectification right**: Allow rectification of incorrect personal data and the provision of supplementary data.
- **Right to be forgotten**: Allow erasing personal data when requested. Once deleted it should not be retained.
- **Right to restriction of processing**: Restrain processing of personal data under the situation stated in Art. 18 [21].
- **Notification obligation**: Notify any rectification or erasure or restriction of processing to each recipient stated in Art. 19 [22].
- **Right to data portability**: Allow data subject to transfer personal data as required to another data controller providing that Article 20(1) [4] is applicable.
- **Right to object**: Provide the option to object the processing if the conditions in Article 21 [23] apply. Respond and demonstrate legitimate grounds at earliest convenience.
- **Automated decision making**: Avoid taking decision merely on automated processing including profiling which produces legal effects Article 22 [24].

[25] show an effective methodology for analyzing privacy threats essentially aims on the basic privacy concepts such as non-repudiation, information disclosure, policy content awareness etc. It gives the engineers an initial knowledge about how an intruder might attack a system and helps them setting up effective safeguards while implementing the system.

The primary implementation approaches are as follows:

1) Appoint a data protection officer
2) Stick to a code of conduct
3) Acquire consent for data
4) Ensure 3rd party data processing is compliant to user consent
5) Provide set of policies to end users
6) Develop a dashboard where user can update the consent according to each policies
7) Categorize sources of personal data
8) Protect data subject right to erasure of personal data
9) Establish anonymized processes
10) Establish governance programs
11) Access restriction to approved processes

Our proposed implementation provides GDPR compliance in the following areas:

- Systems and data auditing
- Data maintenance
- Work-flows and data processing
- Policy dashboard configuration
- User consent preference
- User subscription management
- User personal data capture

IV. System Design and Development

User interface (UI) and user experience (UX) design require a user dashboard through which a user can easily gain access to the privacy related options and gives consent for further use of his data. The following paragraphs discusses the user dashboard requirements and functionalities that are required for privacy by design approach. Therefore, the whole management workflow of the news media industry is elaborated followed by the database design issue and miscellaneous.

A. System design for News Media Industry

In order to design a sustainable system with GDPR compliance, we needed to follow the primary implementation approaches mentioned in **Section III**. Therefore at first we identified the required interactions between the user and the system through the user dashboard and added contents for user consent according to that. A web media system basically contains advertisements, data analytics module, cookies, and the site contents- which is generally accessible to users after

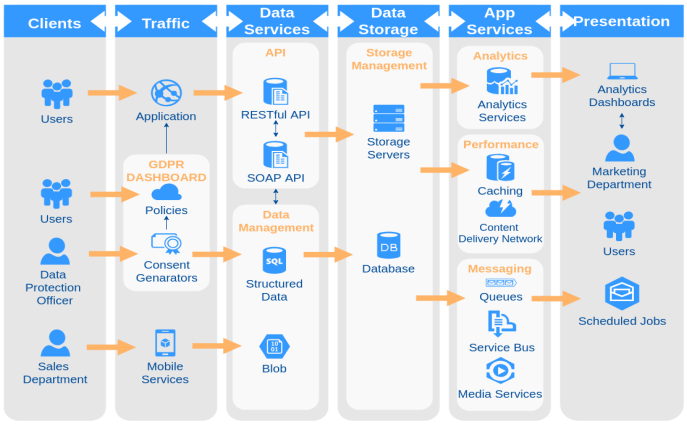


Fig. 3. System Architecture

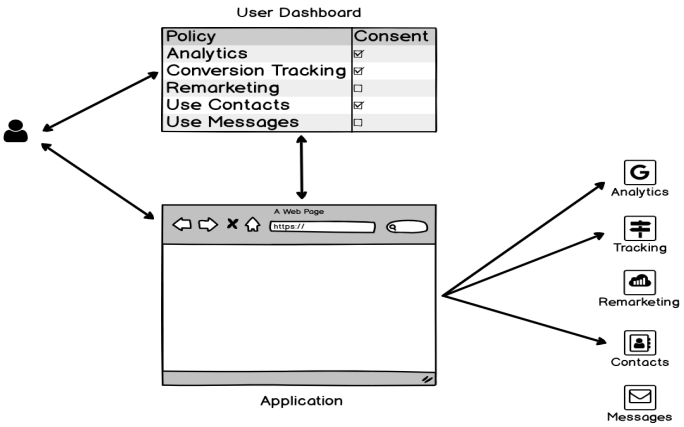


Fig. 4. Dashboard Functions

certain authentication process for users. Except those, the news media website also contains free site contents, newsletter, and other payment oriented monetary modules. The additional part here indicates the policy dashboard for users where user can put his preferences from which the system learns how to handle user data. As the owner of the data, it is completely up to the user whether or not the system can use his data for further analytics or business processes. According to the user consent input, the system stores the privacy data in an access control system through which other analytics module have permission or not to use user data in later cases. The design issues and processes are presented in **Figure 3**.

B. Dashboard Requirement and Functionalities

The user dashboard is the primary concern for the industry as it is the communication and understanding media for them with the users. The dashboard should include the user data policies along with the form that presents user consent options. The application checks initially the user preferences and then acts accordingly. For example, if the user gives his consent to use his data for only 2 options among 5 options, then the application will collect only those data and pass to other analytical machines for further use or analysis. In **Figure 4**, there are five policies for a user including analytics, conversation tracking, remarketing, use contacts, and use messages. A general user has given his consent for only analytics, conversation tracking, and contacts by ticking on the checkmark. Now the application collects data from the user accordingly and lets the data pass to other engines accordingly. Point to be noted that the system will also be responsible to convey the message towards the user on how his data is being used.

The user dashboard is required to have the following functionalities in general:

- Add or delete policies as needed for the application
- User can Check/Uncheck per policy (Give Consents)
- User can See and Update Consents in a dashboard

C. News media management workflow

Data management workflow for a news media includes from the data collection through application according to the user consents to the data transferring gateway to other applications. The application is always responsible for tracking any change that user dashboard provides and collecting the user data. Then the data is stored privately in secure storage lied in their own data centers according to various categories. Later the stored data is passed to other applications for further processing through privacy dashboard. **Figure 5** represents the data flow in a particular system.

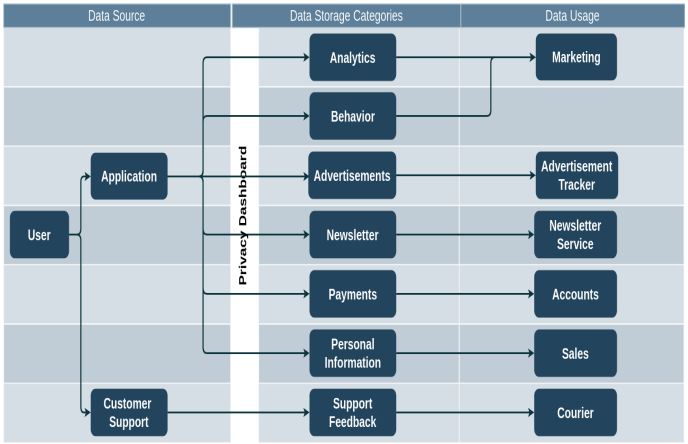


Fig. 5. News Media Corporation Data flow

D. Sample User Dashboard

A standard user dashboard contains the data usability policies and related options to users requesting for their permission in re-usability of user created contents. **Figure 6** represents a sample parallel user dashboard that is easily accessible to users. After reading the company policies, users become able to make the changes in remarketing and recommendation use cases that enforces credibility towards the reuse of user data.

The dashboard expresses whether or not a user is interested in receiving offers from the company through mail or phone. Also as the owner or creator of data, user is able to recommend whether the company will be able to use user data for future customer experience or for enabling personalized advertising

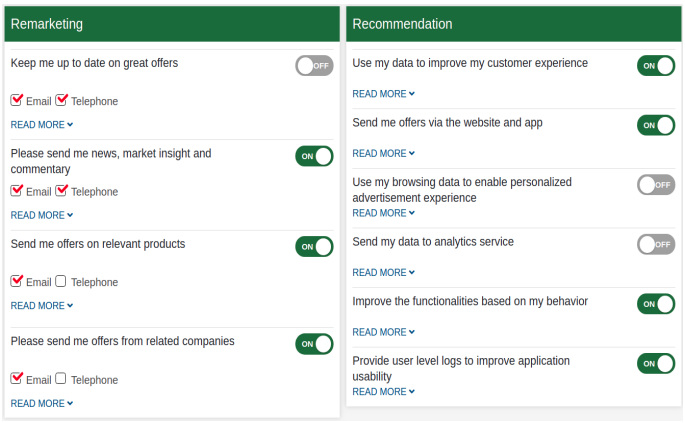


Fig. 6. Sample User Dashboard

experiences. Whether other third parties will be able to use data for behavior analysis or other analytic services- options are present within the dashboard for user permission as well.

V. Discussion

A. Penalty

[26] analyzes how the EUs GDPR will affect the IoT firms in terms of legal and preventive cost all over the world which deals with huge amount of personal information. According to the reference given by Ponemon Institute in 2016 [27], four personal data breach cases out of around twenty four thousands were analyzed to compare the estimated cost damage before and after GDPR. Assuming the average per capita cost of a data breach over the last four years as $150 and for each violation of GDPR costs 2% of a firms global turnover, the results of the analysis shows that in some cases a firms cost may be increased 3-18 times than before GDPR is applied based on the annual turnover of any firm for the past year. Vizio, an electronic product development company was charged with a $2.2 million fine after using content-aware apps to track user data without permission. The data of the users was collected intending to analyze personal TV habits info. If it would happen in current time with GDPR enforced, it would have been $292 million, more than hundred times larger penalty. In addition to that, users have full right to demand compensation for damages if experienced any material or non-material damage due to a violation of GDPR rules.

B. Liabilities of both service provider and consumer

Maintaining proper data privacy requires both parties to be aware of the data usage policies and regulations. Liability goes to both parties as users are responsible for getting familiar with the policies and put ticks on their preferable choices of data usage. On the other hand, service providers must follow user's preference and use data for further purpose thereby.

C. Issues

GDPR has several issues though. [28] enlightens some problems that can be caused when the new GDPR will be in force. The information flow within the digital market can be scattered because of the uneven national provisions. It will be infeasible for a cloud provider to have prior written consent from all of its data subjects before starting any processing task. Furthermore, to ensure highest level of security in every processing the CPs must carry out security assessments more frequently increasing the cost in a considerable amount and eventually the customers will have to bear that cost.

Also, cloud providers must optimize operational cost by adopting efficient privacy compliant practices while handling the data [29] as any information which can identify an individual including name, location, ID number, genetic, and biometric data etc. are considered as personal data.

VI. Related Works

In [28], the major changes in the new GDPR in cloud environment relating the field of healthcare are discussed. Much significance is given on the role of data controller and data processors in the GDPR and also their roles have been refined in some extent. The data subjects are now given more privileges which in turn increases the responsibilities of data controllers. A data controller must find a trustful processor that will maintain the compliance with the new principles such as the right for data access, data portability and also to be forgotten. To maintain appropriate security, both the controller and the data processor have to evaluate the risks associated with the processing of personal data and take feasible measures to eradicate those risks.

[30] focused to solve these problems by introducing the use of some widely used icon sets while preparing the privacy policy to facilitate the users task of decision making. Having the ability to present the legal information in a much clear, fast and effective way, graphical elements such as images icons etc. can improve the readers understanding capability.

The primary focus in [31] is to create a supportive environment for health and social care sectors and an analysis of the impact and related changes to the Information Governance Toolkit (IGTK) in the primary care sector in England considering the presence of new GDPR which will replace the UK data protection act,1998. The paper also describes the impact of introducing cloud computing in e-health, different risk factors and the maintenance of the security in the cloud. 8 principles of the data protection act,1998 are summarized and core 5 elements of information governance are discussed. [32] identified three major threats to security, that is tampering with data (Integrity), loss of data(availability), and unauthorized access to data (confidentiality). With these three under consideration, the risk of data security will be minimized.

The new GDPR that is in force from 25 May,2018; has imposed more obligations to regulate the privacy issue while collecting, storing and processing the personal data of EU citizens [16]. This paper aims at capturing the privacy requirements introduced by GDPR and providing general guidance to the engineers for developing GDPR-standard software. This paper further gave an overview on the roles and responsibilities of the liable authorities while developing a GDPR compliant

system via UML use-case diagrams. While implementing privacy policies within an organization several factors are too be considered such as directives, inspections, fine amount, and other profit and loss side effects [33].

VII. Conclusion

The primary concern of GDPR is to protect user personal data by enforcing the law. These days, all companies including Google, Facebook, and Amazon etc are following the rules and regulations. The news media industry is still adopting the new framework towards more user aware data usage policies. Processing of data is now more transparent towards users end. As the owner of data, users can change their preference at any time and become more aware of the policies on where and how the data is being used by other applications. This leads to privacy or policy by design that is more concerned about the security of personal data. Our paper discussed the design and developmental adoption to the new rules and regulations. Furthermore, we discussed the user dashboard in details so that other companies can follow the adoption with framework transformations that comply with the new regulations.

Acknowledgement

The work has been partially funded by Green University Research Fund.

References

[1] "Article 33 eu general data protection regulation (eu-gdpr). privacy/privazy according to plan." http://www.privacy-regulation.eu/en/article-33-notification-of-a-personal-data-breach-to-the-supervisory-authority-GDPR.htm, (Accessed on 07/29/2018).

[2] "Article 15 eu general data protection regulation (eu-gdpr). privacy/privazy according to plan." http://www.privacy-regulation.eu/en/article-15-right-of-access-by-the-data-subject-GDPR.htm, (Accessed on 07/29/2018).

[3] "Article 17 eu general data protection regulation (eu-gdpr). privacy/privazy according to plan." http://www.privacy-regulation.eu/en/article-17-right-to-erasure-'right-to-be-forgotten'-GDPR.htm, (Accessed on 07/29/2018).

[4] "Article 20 eu general data protection regulation (eu-gdpr). privacy/privazy according to plan." http://www.privacy-regulation.eu/en/article-20-right-to-data-portability-GDPR.htm, (Accessed on 07/29/2018).

[5] "Article 25 eu general data protection regulation (eu-gdpr). privacy/privazy according to plan." http://www.privacy-regulation.eu/en/article-25-data-protection-by-design-and-by-default-GDPR.htm, (Accessed on 07/29/2018).

[6] "Article 39 eu general data protection regulation (eu-gdpr). privacy/privazy according to plan." http://www.privacy-regulation.eu/en/article-39-tasks-of-the-data-protection-officer-GDPR.htm, (Accessed on 07/29/2018).

[7] G. D'Acquisto, J. Domingo-Ferrer, P. Kikiras, V. Torra, Y.-A. de Montjoye, and A. Bourka, "Privacy by design in big data: an overview of privacy enhancing technologies in the era of big data analytics," *arXiv preprint arXiv:1512.06000*, 2015.

[8] K. Boersma, P. Meier, and M. Imran, "The use of social media for crisis management: a privacy by design approach," in *Big Data, Surveillance and Crisis Management*. Routledge, 2017, pp. 33–51.

[9] "Article 23 eu general data protection regulation (eu-gdpr). privacy/privazy according to plan." http://www.privacy-regulation.eu/en/article-23-restrictions-GDPR.htm, (Accessed on 07/29/2018).

[10] S. Gürses, C. Troncoso, and C. Diaz, "Engineering privacy by design," 2011.

[11] "Facebook's commitment to data protection and privacy in compliance with the gdpr — facebook business," https://www.facebook.com/business/news/facebooks-commitment-to-data-protection-and-privacy-in-compliance-with-the-gdpr, (Accessed on 07/29/2018).

[12] "General data protection regulation (gdpr) — google cloud," https://cloud.google.com/security/gdpr/, (Accessed on 07/29/2018).

[13] P. O. P. DATA, "Article 29 data protection working party," 2012.

[14] E. Commission, "Special eurobarometer 431: Data protection," 2015.

[15] F. Schaub, R. Balebako, A. L. Durity, and L. F. Cranor, "A design space for effective privacy notices," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, 2015, pp. 1–17.

[16] G. D. P. Regulation, "Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46," *Official Journal of the European Union (OJ)*, vol. 59, no. 1-88, p. 294, 2016.

[17] N. Sadeh, A. Acquisti, T. D. Breaux, L. F. Cranor, A. M. McDonald, J. R. Reidenberg, N. A. Smith, F. Liu, N. C. Russell, F. Schaub *et al.*, "The usable privacy policy project," Technical report, Technical Report, CMU-ISR-13-119, Carnegie Mellon University, Tech. Rep., 2013.

[18] E. Directive, "95/46/ec of the european parliament and of the council of 24 october 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data," *Official Journal of the EC*, vol. 23, no. 6, 1995.

[19] H. Nissenbaum, "Privacy as contextual integrity," *Wash. L. Rev.*, vol. 79, p. 119, 2004.

[20] K. Borcea-Pfitzmann, A. Pfitzmann, and M. Berg, "Privacy 3.0:= data minimization+ user control+ contextual integrity," *IT-Information Technology Methoden und Innovative Anwendungen der Informatik und Informationstechnik*, vol. 53, no. 1, pp. 34–40, 2011.

[21] "18. right to restriction of processing - easygdpr," https://easygdpr.eu/gdpr-article/18/, (Accessed on 02/05/2019).

[22] "19. notification obligation regarding rectification or erasure of personal data or restriction of processing - easygdpr," https://easygdpr.eu/gdpr-article/19/, (Accessed on 02/05/2019).

[23] "21. right to object - easygdpr," https://easygdpr.eu/gdpr-article/21/, (Accessed on 02/05/2019).

[24] "22. automated individual decision-making, including profiling - easygdpr," https://easygdpr.eu/gdpr-article/22/, (Accessed on 02/05/2019).

[25] K. Wuyts, R. Scandariato, and W. Joosen, "Empirical evaluation of a privacy-focused threat modeling methodology," *Journal of Systems and Software*, vol. 96, pp. 122–138, oct 2014. [Online]. Available: https://doi.org/10.1016/j.jss.2014.05.075

[26] J. Seo, K. Kim, M. Park, M. Park, and K. Lee, "An analysis of economic impact on iot under gdpr," in *Information and Communication Technology Convergence (ICTC), 2017 International Conference on*. IEEE, 2017, pp. 879–881.

[27] "2017 cost of data breach study: United states," https://www.ponemon.org/blog/2017-cost-of-data-breach-study-united-states, (Accessed on 08/25/2018).

[28] R. Ducato, "Cloud computing for s-health and the data protection challenge: Getting ready for the general data protection regulation," in *Smart Cities Conference (ISC2), 2016 IEEE International*. IEEE, 2016, pp. 1–4.

[29] P. Blume, "It is time for tomorrow:: Eu data protection reform and the internet," *Journal of Internet law*, vol. 18, no. 8, pp. 3–13, 2015.

[30] A. Rossi and M. Palmirani, "A visualization approach for adaptive consent in the european data protection framework," in *E-Democracy and Open Government (CeDEM), 2017 Conference for*. IEEE, 2017, pp. 159–170.

[31] I. N. Shu and H. Jahankhani, "The impact of the new european general data protection regulation (gdpr) on the information governance toolkit in health and social care with special reference to primary care in england," in *2017 Cybersecurity and Cyberforensics Conference (CCC)*. IEEE, 2017, pp. 31–37.

[32] I. de la Torre Díez, K. Saleem, S. G. Alonso, M. S. Khalil, S. Hamrioui, J. J. Rodrigues, and M. L. Coronado, "Requisites of security, reliability and usability in 2 mhealth apps: Systematic analysis and proposed 3 architecture 4," 2018.

[33] M. J. Culnan and C. C. Williams, "How ethics can enhance organizational privacy: lessons from the choicepoint and tjx data breaches," *Mis Quarterly*, pp. 673–687, 2009.

Applying Model-Free Reinforcement Learning Algorithm in Network Slicing for 5G

Mostafa Zaman Chowdhury[1,2], Md. Tanvir Hossan[1], and Yeong Min Jang[1]
[1]Dept. of Electronics Engineering, Kookmin University, South Korea
[2]Dept. of EEE, Khulna University of Engineering & Technology, Bangladesh
Email: mzceee@ieee.org, mthossan@ieee.org, and yjang@kookmin.ac.kr

Abstract– A network slicing (NS) represents a form of isolated logical or virtual network architecture for a fixed physical network of 5G to support different radio access networks and adequately provide a specific service platform for the customers. The possible application of NS will find in the varying infrastructure of a network. Moreover, among several research challenges in NS, e.g., trouble to isolate the slices, mobility management, cooperation with other 5G technologies, and management of NS; the most challenging tasks are allocating resources of the slice efficiently and ensuring security. Reinforcement learning (RL) approach in NS for 5G helps encounter these critical challenges. It automatically reduces the latency and block suspected threat profile to ensure security in that slice. Additionally, the possible uses of allocated resources in SN improved sufficiently by applying RL technique. This paper will appropriately address to develop a policy-gradient based model-free RL approach to achieve optimal policy for NS with current values and action variables and accomplish optimal solutions for security improvement and detecting malicious nodes within the NS for 5G.

Keywords–Network slicing, reinforcement learning algorithm, policy-gradient approach, 5G.

I. INTRODUCTION

The term network slicing (NS) implies several logical or virtual networks of a core network and/or radio access networks (RAN) [1], whereas each slice of the network dedicates to serve a specific functionality (e.g., security, mobility support) and performance (e.g., latency, reliability, throughput) depending on the user requirement. The key objective to implicate NS to provide a platform for coexisted architecture in a fixed and compressed physical infrastructure to allow a flexible virtual network for the 5G mobile network alliance. It is required due to the considerable increase of RAN spectral efficiency in a significant number as well as the mobile networks need significantly scaling up and improve efficiency. NS take care of various services on mobile networks and makes it more flexible and programmable fabric. Additionally, it offers a control and management platform for the network infrastructure as well as allocated resources for optimal uses. Figure 1 describes the general architecture of sliced networks for 5G.

Meaningful discussion on NS is stronger than before due to the drawbacks of the existing networks, where typical networks deliver support various services from the same network region. The strategy of this network architecture naturally fails to make available a proper platform for internet-of-things devices, low latency, high throughput, high data rate, etc. Meanwhile, NS will provide valued support to 5G networks to deal with some unpredictable network applications, like haptic communications. Although it also enhances the potential scopes and benefits in automotive sectors. These support is required not only for radio frequency (RF) based application in 5G but also for several other technologies, which wants to merge with it. One of the technologies is based on optical wireless communications [2], which opens different futuristic scenarios, e.g., human bond communications [3]. Despite advantages and importance, NS is experiencing a few challenges (i.e., mobility management, cooperation with other 5G technologies, slices isolation and management) [1], which has caught the attention of the other researchers. Most challenging tasks are radio resource optimization and provide better solutions at the point of security of NS [4].

The radio resource slicing is typically defined as the NS in the wireless domain. The quality of service (QoS) of entire radio resource measures with slices. Resource slice isolation is a fundamental feature that ensures traffic from one slice never interfere to the other slice. Optimization of a radio resource slice decides the reconfiguration resolution of it. This optimization strategy can carry better QoS and ensures highest utilization of these potential resources. It enables the terminal between a network and terminal resource manager [4]. On the other hand, a cryptography protection algorithm is typically required for each slice of the networks from being hacked by anonymous nodes/networks and make these slices more accessible to multiple users [5].

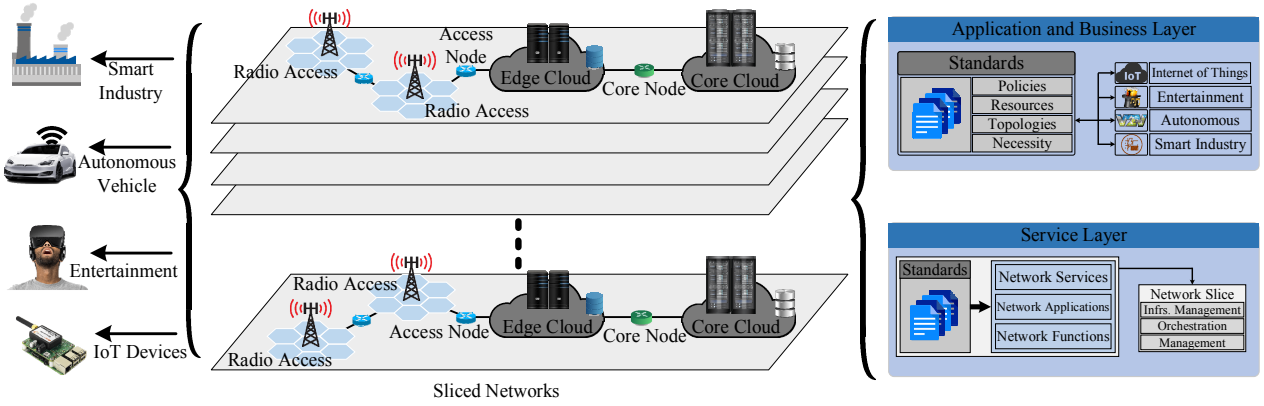


Figure 1. Overall architecture of sliced networks for 5G.

In this paper, we are deploying reinforcement learning (RL) approach, which can ensure proper resource slice optimization and security in the same slice. It is an artificial intelligence and unsupervised approach to carry out decisions about a critical environment (e.g., sliced networks) using the information from agents. Recent progress for artificial intelligence based network slicing has discussed in section II. Section III, explains briefly the innovation in our research approach. In the following section (section IV) we are discussing our proposed idea along with the simulation result. The way to reach the research destination has discussed in section V. The following two sections (i.e., section VI and section VII) illuminate the potential challenges and approaches for those challenges, respectively, before it ends up with conclusions in section VIII.

II. Related Works

NS in 5G is recent research concern. In [8], authors are introduced a possible solution for exchanging necessary data between vehicle-to-everything by slicing the core network and the radio access network resources. The article also includes the key contributions and configurations in the isolated treatment and guaranteed performance for same application scenarios of critical 5G cases. A complete integration of software defined networks and network function virtualization architecture for 5G sliced network, which is presented in [5]. For visualizing such cases, the authors also bring a practical example of the integration technologies to address its importance for the sliced network. At the end of their article, open research issues are enlisted. The overall characteristics of a distributed cloud infrastructure for NS are carefully analyzed along with the illustration of the optimization enhancement of the same network for 5G, has shown by the authors in [9]. However, possible deployment of a sliced network in specific application, e.g., ultra-dense cellular network and mobile edge computing based 5G wearable network architecture is discussed in [10]. In this article, the authors introduce the data driven network slicing management technique for proper resource sharing and energy-efficient application.

The extensive involvement of artificial intelligence in sliced networks is reasonably demanding to face existing challenges, especially in resource management and adequate security. However, the research progress is unsatisfactory enough in these regards and few of those discussions RL for a comprehensive solution. In [11], the authors present RL based slice admission policy to enhance the considerable profit of an infrastructure. Their proposed RL based policy shows up to 54.5% improved performance than the standards. In addition, authors in [12], agree that deep-RL is an essential part for resource management of each sliced. It helps those NS to provide better performance and services hasty. Here, deep-RL deals with several key issues, such as typically providing quality of experience and reduces the computational cost within the network resources. Eventually, deep-RL solves the optimal control policies in problems of RL with raw visual input. Conversely, the authors never discuss in their article about the measures like ensuring security by RL.

III. Innovation in Research

The possible and efficient utilization of the network resources can be solved by optimization approaches. The application of model-free RL structure can solve the optimization problems by analyzing the sliced networks into a finite set of discrete values [7], [13]. Moreover, the decision maker (or agent) helps predict new malicious attacks and achieve optimal or near-optimal performance of the sliced networks. RL can identify the dynamically changing strategies of malicious nodes in time progress. However, adding or removing flexibility of the sliced network puts the core network in a security risk. RL reliably detects the malicious node and convenient store this information until the formation or removal of sliced networks [14].

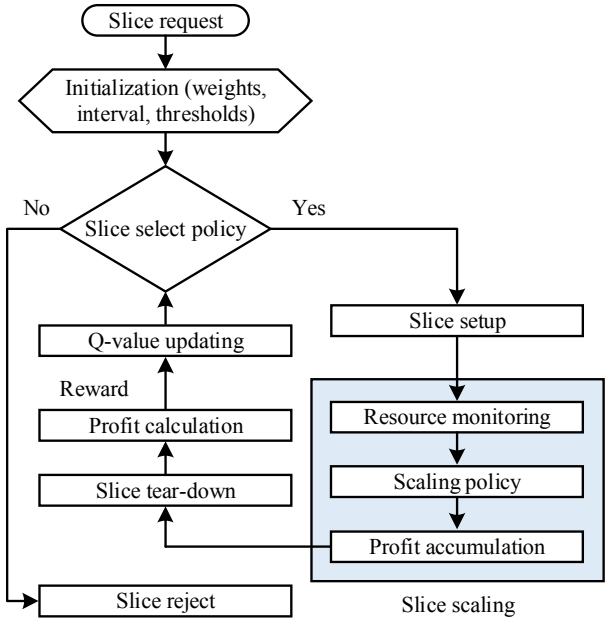


Figure 2. Model-free RL approach for 5G NS.

IV. Proposed Idea and Simulation Result

Applying a model-free RL approach to optimize available resources efficiently for the network slice as shown in Figure 2. The agent of this model typically depends on trial and error based feature for action selection. Without the extensive knowledge of transition and reinforcement function, this model can adequately provide necessary decision to evaluate the complex network scenarios like sliced network. The trial-and error learning experience helps to optimize the network by accurately mapping the available resources of the sliced network. However, with the several iterations of actions, the availability of the malicious node accurately identifies within the NS and adequately provides possible solution for security improvement with this identification.

Furthermore, the policy-gradient-based reinforcement learning algorithm [15] select an action from most significant reward to achieve precisely the optimal policy for NS with current values and action variables. This algorithm includes three distinctive methods, i.e., policy-based, value-based, and actor-critic methods.

The proposed algorithm has compared with two different algorithms, such as random algorithm and greedy-based algorithm. Under certain condition, random algorithm randomly allocates the sliced network resources. Figure 3 and Figure 4, show the total revenue and cost of energy with respect to the number of sliced network, respectively. In all cases, proposed algorithm almost always shows better performance than others algorithms.

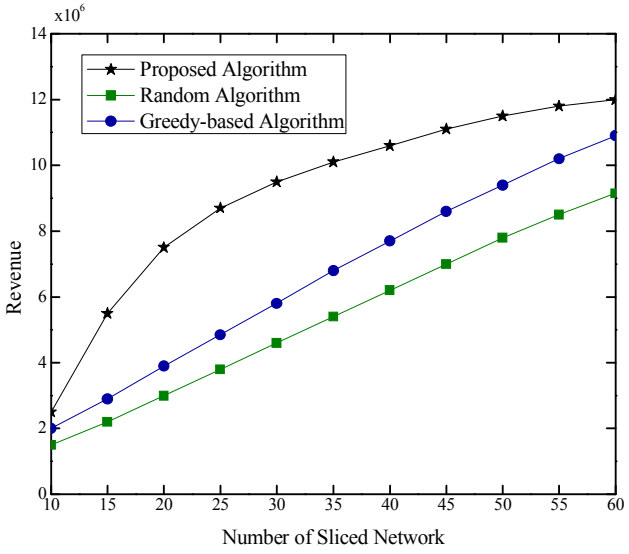


Figure 3. Comparison of total revenue with respect to the number of sliced networks.

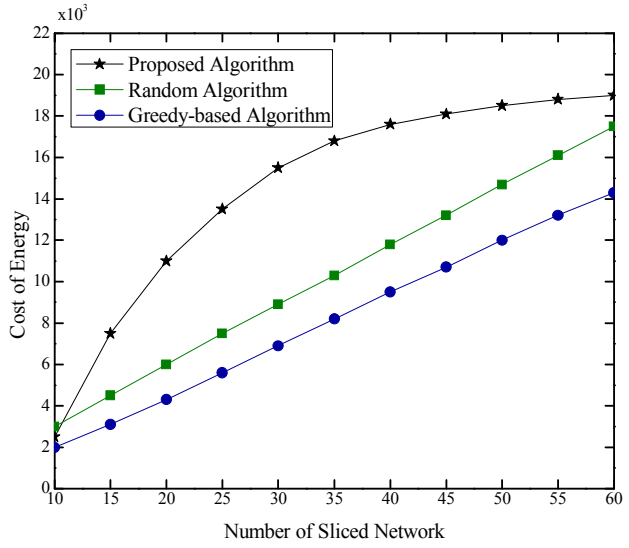


Figure 4. Comparison of total cost of energy with respect to the number of sliced networks.

V. Research Direction

The research timeline conveniently divides into three completed stages, which includes the literature review, modelling, functional analysis, and review the steps. The research review process will include constructive discussion and comparison with possible results found in the literature. The research timeline obtains as follows:

Stage 1: Literature review and initial developing an RL based efficient algorithm for optimizing resources and ensuring security in NS for the lightly dense traffic scenario.

Stage 2: The research will focus on the development of RL based efficient algorithm for dynamic, heavy, dense traffic

scenarios and achieve precisely the optimal policy for NS with policy-iteration based RL approach.

Stage 3: Progressively improve the performance of RL based NS for the next generation wireless communication system (i.e., 5G and beyond 5G) and mitigating the challenges for critical cases.

VI. Potential Challenges

The most significant reward for an action will obtain from a series of trial actions in trial-and-error learning approach of RL to accurately determine a distinct action for sliced network optimization. Moreover, it is not always a reliable way to determine an action, where the learning approach is typically based on the arbitrary choice. Simultaneously, the prolonged learning process also pushes the whole security system of the sliced network into a vulnerable situation.

On the other hand, both policy-based and value-based methods have some key limitations individually, such as an insufficient convergence process, significant variance to weights the performance of the policy, perform iterative computation from finite discrete actions, and difficult to determine the learning rate; whereas actor-critic method has better convergence process [16]. However, high variance at actor part does not provide any trouble to estimate policy gradient due to the assistance of critic part.

VII. Potential Approaches

The potential approaches for the research will address as follows:

Approach 1: Integrating imitation or hierarchical learning approaches with RL to progressively improve its learning speed for fast network optimization and also ensure network security.

Approach 2: Deploying policy-gradient-based actor-critic reinforcement learning algorithm to achieve optimal policy for NS. At this place, the actor generates successive actions, whereas a critic weights the performance of the policy with rewarding the action initiated by the actor.

VIII. Conclusions

The current NS has faced several challenges, including slices isolation, slices network management, mobility management, resource allocation for slices, maintain security in slices, which is required to solve to fulfil the demand for wireless capacity with 5G and beyond communications. In this paper, we have proposed a model-free RL for resource optimization and security enhancement. Additionally, policy-gradient based RL helps manage an optimal policy for NS. Detecting new malicious attacks with accomplishing optimal or near-optimal performance of NS will be the biggest novelty of this approach. However, it still faces some challenges, such as an insufficient convergence process, significant variance to weights, iterative computation from finite discrete actions, and slow learning progress, leads to a good research issue.

Acknowledgement: This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-0-01396) supervised by the IITP (Institute or Information & communications Technology Promotion) and the Korea Research Fellowship Program (2016H1D3A1938180).

References

[1] X. Foukas *et al*., "Network slicing in 5G: survey and challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 94–100, May 2017.

[2] M. Z. Chowdhury, M. T. Hossan, A. Islam, and Y. M. Jang, "A comparative survey of optical wireless technologies: architectures and applications," *IEEE Access*, vol. 6, pp. 9819–9840, 2018.

[3] M. T. Hossan, M. Z. Chowdhury, M. Shahjalal, and Y. M. Jang, "Human bond communication with head-mounted displays: scope, challenges, solutions, and applications," *IEEE Communications Magazine*, vol. 57, no. 2, pp. 26–32, Feb. 2019.

[4] X. Li *et al*., "Network slicing for 5G: challenges and opportunities," *IEEE Internet Computing*, vol. 21, no. 5, pp. 20–27, 2017.

[5] J. O. Lucena *et al*., "Network slicing for 5G with SDN/NFV: concepts, architectures, and challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 80–87, May 2017.

[6] H. Liu, S. Liu, and K. Zheng, "A reinforcement learning-based resource allocation scheme for cloud robotics," *IEEE Access*, vol. 6, pp. 17215–17222, Mar. 2018.

[7] Y. Saleem *et al*., "Clustering and reinforcement-learning-based routing for cognitive radio networks," *IEEE Wireless Communications*, vol. 24, no. 4, pp. 146–151, Aug. 2017.

[8] C. Campolo, A. Molinaro, A. Iera, and F. Menichella, "5G network slicing for vehicle-to-everything services," *IEEE Wireless Communications*, vol. 24, no. 6, pp. 38–45, Dec. 2017.

[9] S. Sharma, R. Miller, and A. Francini, "A cloud-native approach to 5G network slicing," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 120–127, Aug. 2017.

[10] Y. Hao, D. Tian, G. Fortino, J. Zhang, and I. Humar, "Network slicing technology in a 5G wearable network," *IEEE Communications Standards Magazine*, vol. 2, no. 1, pp. 66–71, Mar. 2018.

[11] R. Li *et al*., "Deep reinforcement learning for resource management in network slicing," *IEEE Access*, vol. 6, pp. 74429–74441, 2018.

[12] M. R. Raza *et al*., "A slice admission policy based on reinforcement learning for a 5G flexible RAN," in Proc. of *2018 European Conference on Optical Communication (ECOC)*, 2018, pp. 1–3.

[13] O. Dayan and Y. Niv, "Reinforcement learning: the good, the bad and the ugly," *Current Opinion in Neurobiology*, vol. 18, no. 2, pp. 185–96, Aug. 2008.

[14] M. H. Ling *et al*., "Application of reinforcement learning for security enhancement in cognitive radio networks," *Applied Soft Computing*, vol. 37, pp. 809–829, Dec. 2015.

[15] R. S. Sutton *et al*., "Policy gradient methods for reinforcement learning with function approximation," in Proc. of *Advances in Neural Information Processing Systems*, 2000 pp. 1057–1063.

[16] I. Grondman *et al*., "A survey of actor-critic reinforcement learning: standard and natural policy gradients," *IEEE Transactions on Systems, Man, and Cybernetics*, Part C (Applications and Reviews), vol. 42, no. 6, pp. 1291–1307, Nov. 2012.

Throughput Analysis of Compressive Sensing in Cooperative Opportunistic Radio Network

Shohely Tasnim Anindo[1], Teethiya Datta[2], Sk. Shariful Alam[3]
[123]Department of Electronics and Communication Engineering
[123]Khulna University of Engineering & Technology, Khulna-9203, Bangladesh
anindyodip@gmail.com[1], TeethiyaDatta@gmail.com[2], ssa1996@gmail.com[3]

***Abstract*—Cooperation in opportunistic radio (OR) technology is one of the prominent alternatives to meet up spectrum inadequacy and to enhance Quality of Services (QoS) in wireless communication infrastructure. Opportunistic Users (OUs) are sensor nodes that utilize spectrum when transmission channel is unoccupied, OU senses spectrum and cooperatively takes their decision about data transmission, it provides better probability of detection (P_{d}) about whether Primary User (PU) is active or inactive and lower probability of false alarm (P_{f}). Though for cooperation sensing time increases, it provides better detection, reliable data transmission and maximized throughput. In this paper we have presented comparative studies among different cooperative decision fusion rules and their impact on throughput under different fading channels. The cooperation of different no. of nodes and effect of changing SNR on achievable throughput for OU data transmission are also studied. Finally, optimal tradeoff between sensing time and time frame under different channels and fusion rules are investigated for compressive sensing in opportunistic radio network (ORN).**

***Index Terms*—Cooperation, opportunistic radio, throughput maximization, decision fusion**

I. Introduction

Spectrum utilization has become a great issue as spectrum scarcity problem has been become a challenge for its limited resources. To use underutilized part of spectrum, ORN has become popular now a days [1]. In ORN, OU first senses spectrum then makes decision about data transmission depending on presence of primary users (PU) in spectrum bands. As soon as OU senses the presence of primary users, it vacates the channel within a certain amount of time [2]. Non cooperative sensing indicates that single opportunistic node senses spectrum and makes local decision about its own transmission and transmits data if primary user is found inactive. But in cooperative sensing, a group of sensor nodes make their decisions jointly by applying decision fusion rules among individual local decisions and the final decision is carried out for all nodes. For this, all local decisions are sent to a predefined node among the group of secondary nodes called decision Fusion Centre (FC). The decision is made by FC and FC sends cooperative decision to all nodes. So sensing time increases for cooperation and decision about P_{d} and P_{f} are distinguished more properly than that of no cooperation case. Thus throughput performance of cooperative compressive sensing is better than no cooperation case [3]. Sometimes for cooperation OU nodes are considered to be in defined topology [4]. In this paper, for simplicity no topology is defined for cluster of nodes. Compressive sensing based spectrum sensing is performed where signal detection classification and co-channel interference [5] are avoided.
Fading is fluctuation of signal strength in small period of time [6]. Fading occurs due to multipath propagation, Doppler spread, attenuation and noise of particular environment. There are some model of fading channels to reduce effect of fading [7], most of the cases real time fading follows these fading models. In this paper a study is performed for throughput maximization in cooperative ORN for different channels (AWGN as non fading channel and Rayleigh, Rician and Nakagami as fading channel). The relevance between sensing time and throughput for different channels are studied. More precisely, the objectives of this paper are to investigate maximized throughput in terms of optimal sensing time and time frame and to observe performance of different cooperation methods among OUs for different channels.
These paper has been organized in five sections. In section II cooperative wideband spectrum sensing is shortly discussed, in section III different cooperation rules and fading channels are discussed, in section IV simulation outcomes and finally in section V conclusion is discussed.

II. Spectrum Sensing and Throughput Preliminaries

A. Signal Model

Lets assume a signal $s(n)$, whose frequency band with carrier frequency f_{c} and bandwidth W and sampling frequency of the received signal is f_{s}. When the primary user is in data transmission state, the discrete received signal at the opportunistic user follows [2]

$$y(n) = s(n) + u(n) \tag{1}$$

which is considered as circumstance under hypothesis H_1 where $s(n)$ is transmitted signal and $u(n)$ is noise signal. When the primary user is not in data transmission, the received signal of opportunistic user follows

$$y(n) = u(n) \tag{2}$$

and this case is referred to as hypothesis H_0. Here γ is represented as the signal to noise ratio (SNR) measured at opportunistic receiver for presence of primary users.

$$\gamma = \frac{{\sigma^2}_{\mathrm{s}}}{{\sigma^2}_{\mathrm{u}}} \tag{3}$$

γ is measured at opportunistic receiver (or OU) under hypothesis H_1. Here σ_s and σ_u are variance of transmitted signal and noise signal respectively. Two probabilities are apparent for spectrum sensing: P_d, which is considered under hypothesis H_1, is the probability of correctly detecting the presence of primary signal; and P_f, which is considered under hypothesis H_0, is the probability of falsely declaring the presence of primary signal. For a good detection purpose, high P_d is desired while P_f must be kept as low as possible.

B. Energy Detection

Energy detection is performed after compressive sensing. Energy detectior is consists of filter, square law device and finite time integrator. The received signal of OU is compared with a threshold and decisions are made about the presence of PU. Later the values of P_d and P_f are used for throughput calculation. Lets assume that available sensing time is τ and N is number of samples (N is the maximum integer and is not greater than f_s, and for notation simplicity, we assume (N=τf_s). The test statistic for energy detector is given by

$$T(y) = \frac{1}{N}\sum_{n=1}^{N} |y(n)|^2 \tag{4}$$

Under hypothesis H_0, the test static T(y) is a random variable and its probability density function (PDF)- $P_0(x)$ is a Chi-square distribution with $2N$ degrees of freedom for complex valued case and with N degrees of freedom for real-valued case. If we choose the detection threshold as ϵ, the probability of false alarm is then given by

$$P_f(\epsilon,\tau) = P_r(T(y) > \epsilon|H_0) = Q\left((\frac{\epsilon}{{\sigma^2}_u} - 1)\sqrt{\tau f_s}\right) \tag{5}$$

and the probability of detection is then given by

$$P_d(\epsilon,\tau) = P_r(T(y) > \epsilon|H_1)$$
$$P_d(\epsilon,\tau) = Q\left((\frac{\epsilon}{{\sigma^2}_u} - \gamma - 1)\sqrt{\frac{\tau f_s}{2(2\gamma+1)}}\right) \tag{6}$$

Therefore, for a target probability of detection P_d, the probability of false alarm is dependant on the target detection probability as follows:

$$P_f = Q\left(\sqrt{2\gamma+1}\,Q^{-1}(P_d) + \gamma\sqrt{\tau f_s}\right) \tag{7}$$

On the other hand, for a target probability of false alarm, P_f, the probability of detection is given by

$$P_d = Q\left(\frac{1}{\sqrt{2\gamma+1}}(Q^{-1}(P_f) - \gamma\sqrt{\tau f_s})\right) \tag{8}$$

Finally, for a given pair of target probabilities (P_d, P_f), the number of required samples to achieve these targets can be determined by cancelling out the threshold variable ϵ. The minimum number of samples is given by

$$N_{min} = \frac{1}{\gamma^2}[Q^{-1}(P_f) - Q^{-1}(P_d)\sqrt{2\gamma+1}]^2 \tag{9}$$

C. Throughput of OU

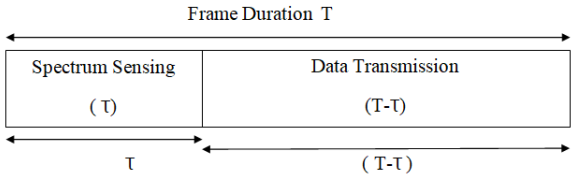


Fig. 1: Structure of single frame in Opportunistic Radio Network [8]

If a frame duration be T and sensing time τ then time required for transmission is (T-τ) as depicted in Fig 1, γ be the SNR found at receiver of OU for hypothesis H_1 when PU is active. f_s be sampling frequency and Throughput of OU, when PU is absent and $P(H_0)$ be probability of hypothesis H_0 after energy detection, throughput for a target probability of detection P_d is [2]

$$R(\tau) = C_0 P(H_0)(1 - \frac{\tau}{T})(1 - Q(\alpha + \gamma\sqrt{\tau f_s})) \tag{10}$$

where

$$\alpha = \sqrt{2\gamma+1}\,Q^{-1}(P_d) \tag{11}$$

Thus, it can be seen that the achievable throughput of the opportunistic network is a function of the sensing time τ.

III. Cooperation of OU

Fading channel: Channel models are utilized to reduce effect of fading [9]. Here, AWGN as non fading channel and among fading channels Rayleigh, Rician and Nakagami channel models are considered to observe throughput performance.
Cooperation: Cooperation is performed among OUs to make integrated decision [10]. For cooperation of different no. of nodes, sensing time increases so apparently time for data transmission decreases [4].

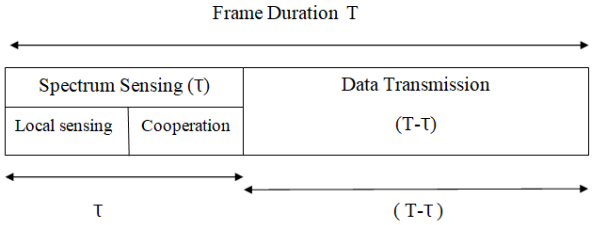


Fig. 2: Frame structure for cooperative sensing in Opportunistic Radio Network [4]

But for cooperation, detection is made better. For cooperative sensing, decision fusion operations are needed to make final decision by fusing individual decisions of sensor nodes. There are four fusion operations [11]. They are:

AND Operation: AND operation works as follows: if all decisions say that there is a primary user, then the final decision declares that there present a primary user. Mathematically

$$P_{d,AND} = P_{d,i}^{N} \tag{12}$$

where N= no of nodes, if $P_{\text{d,AND}}$=1, primary user is present.
OR Operation: OR operation is also simple decision operation like AND operation: if one of the decisions says that there is a primary user, then the final decision declares that there is a primary user. Mathematically,

$$P_{d,OR} = 1 - \Sigma(1 - P_{d,i})^{N} \tag{13}$$

$P_{\text{d,OR}}$=1 then the primary user is present; otherwise, there is no primary user.
Majority Operation: Majority operation is mainly based on majority of the individual decisions. If half or greater than half individual decisions decide that there is a primary user, then the final decision declares that there a primary user is present. Mathematically,

$$P_{d,MAJ} = \sum_{l=\lfloor \frac{N}{2} \rfloor}^{N} (\binom{N}{l} P_{d,i}^{l}(1 - P_{d,i})^{N-l} \tag{14}$$

Optimal Fusion Operation: In this operation sensor nodes are given weighted value to their decisions based on preference, performance and distance of nodes. Mathematically,

$$\Lambda = \sum_{i=1}^{n} [u_i \log \frac{P_{di}}{P_{fi}} + (1-u_i) \log \frac{1 - P_{di}}{1 - P_{fi}}] + \log \frac{P(H_1)}{P(H_0)} \tag{15}$$

if Λ=1, primary user is considered to be present. Decision of highest priority nodes are valued most.

IV. Simulation Outcomes

For analysis of throughput, The signal was taken 64 MHz wideband signal, BPSK modulated and divided into 32 sub bands. Each sub bands are sampled, coded and reconstructed via compressive sensing and energy detection is performed to estimate P_{d}. Probability for which PU transmission is inactive or absent under hypothesis H_0 is considered 0.9. Target probability of detection considered 0.98 and probability of false alarm 0.02. For achieving simple and fair decision under majority operation,odd no of nodes are considered for simulation. Cooperation of 5 nodes and 11 nodes are simulated to observe the effect of increasing no. of nodes on throughput. SNR was taken 10 dB and 15 dB to observe its impact.

A. Against Sensing Time:

Here time frame was kept fixed and it was 100 ms. Throughput are simulated for both 10dB and 15dB SNR for different channels (AWGN, Rayleigh, Rician and Nakagami channel) and different decision fusion operations (AND, OR, Majority and Optimal Fusion). Cooperation of 11 nodes are considered in every case.

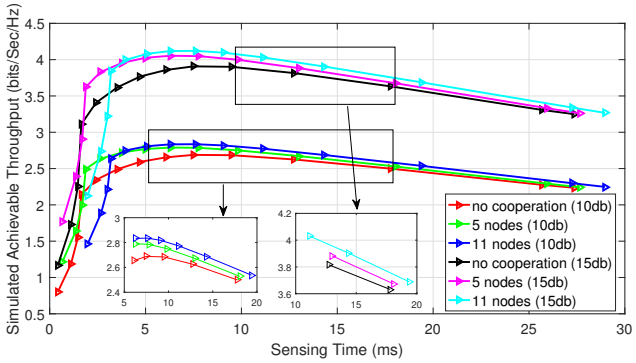


Fig. 3: Throughput Vs Sensing Time for AWGN channel for AND operation (SNR 10dB and 15 dB)

From Fig. 3, achievable throughput found for no cooperation and cooperation of 5 nodes and 11 nodes for AWGN channel AND operation. Here both achievable throughput for 10 dB and 15 dB SNR values are shown. From Fig. 3, it can be seen that, though sensing time for 5 nodes is greater than for no cooperation, and sensing time for 11 nodes are greater than sensing time for 5 nodes, throughput increases for 11 nodes than 5 nodes than no cooperation. Here though sensing time increases, cooperation of nodes provides better result because detection is made better than no cooperation. From simulation achievable throughput found 2.8354, 2.7893 and 2.6886 bits/second/Hz for cooperation of 11 nodes, 5 nodes and no cooperation respectively for SNR value 10 dB. For SNR 15 dB, achievable throughput found 4.1209, 4.0539 and 3.9075 bits/second/Hz for cooperation of 11 nodes, 5 nodes and no cooperation respectively. Again, throughput increases for increased value of SNR.

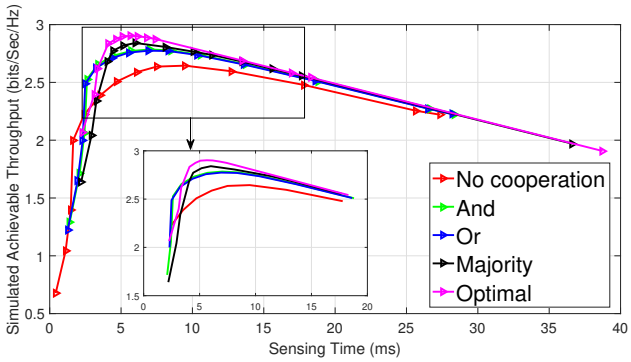


Fig. 4: Achievable Throughput Vs sensing time under Rayleigh channel for different decision fusion operations (AND, OR, Majority, Optimal)

From Fig. 4, simulated achievable throughput found for different fusion operations (AND, OR, Majority and Optimal Fusion) under Rayleigh channel. In every operations, cooperation of 11 nodes are compared with no cooperation case. Maximum achievable throughput for no cooperation found 2.6886 bits/sec/Hz. For AND, OR, Majority and Optimal Fusion operation maximum achievable throughput found 2.8354, 2.8008. 2.8600 and 2.9345 bits/sec/Hz respectively for SNR

10 dB. Optimum sensing time can be determined for different operations. In no cooperation maximum throughput found when sensing time is about 8 ms but for optimal fusion operation optimum sensing time is about 5 ms. Optimal fusion operation provides better throughput than others under same condition in Rayleigh channel.

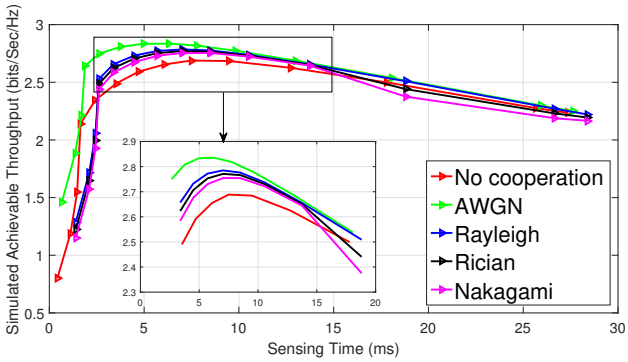


Fig. 5: Achievable Throughput Vs Sensing Time under AND operation for different channels (AWGN, Rayeigh, Rician and Nakagami)

From Fig. 5, simulated achievable throughput found for different channel under AND operation. All these throughput found for different channels are also compared with no cooperation in AWGN channel. The results are taken for SNR 10 dB and cooperation of 11 nodes. Maximum achievable throughput for no cooperation found 2.6886 bits/sec/Hz and for AWGN, Rayleigh, Rician and Nakagami channel maximum achievable throughput found 2.8354, 2.7850, 2.7711 and 2.7543 bits/sec/Hz respectively for 10 dB SNR. From this figure, optimum sensing time can be determined for different channels. In no cooperation maximum throughput found when sensing time is about 8 ms but for cooperation of AWGN channel optimum sensing time is about 5 ms, though cooperation requires more sensing time than no cooperation situation, after achieving cooperative decisions throughput maximization requires less timing than no cooperation case. AWGN channel provides better throughput than others under same condition in AND operation. After 14 ms sensing time, throughput of Rician and Nakagami channel degrades sharply than Rayleigh and AWGN for their fast fading nature.

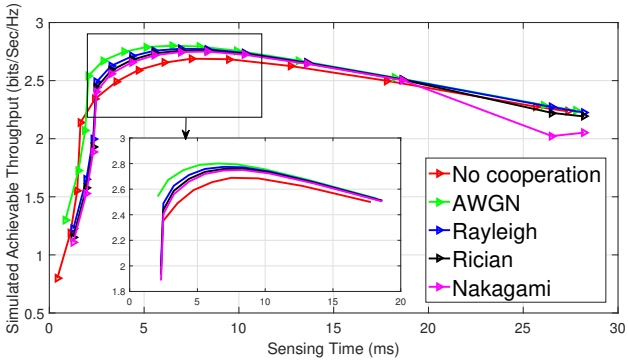


Fig. 6: Achievable Throughput Vs Sensing Time under OR operation for different channels (AWGN, Rayeigh, Rician and Nakagami)

From Fig. 6, simulated achievable throughput found for AWGN, Rayleigh, Rician and Nakagami channel under OR operation. Here also AWGN channel provides better throughput and Nakagami channel provides comparatively inferior performance. Performance of OR operation is almost same as AND operation. Maximum achievable throughput for no cooperation found 2.6886 bits/sec/Hz and for AWGN, Rayleigh, Rician and Nakagami channel maximum achievable throughput found 2.8008, 2.7743, 2.7588 and 2.7502 bits/sec/Hz respectively for 10 dB SNR. But here performance of Nakagami channel degrades more than no cooperation case after about 18 ms sensing time.

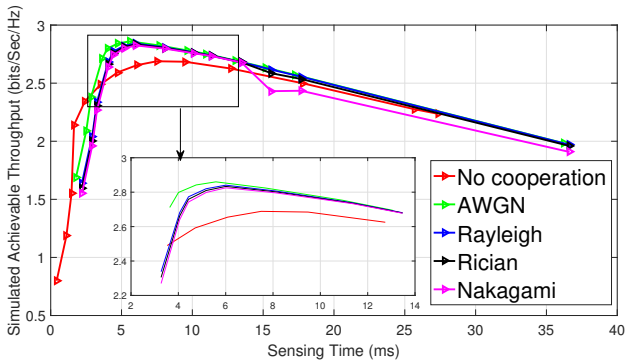


Fig. 7: Achievable Throughput Vs Sensing Time under Majority operation for different channels (AWGN, Rayeigh, Rician and Nakagami)

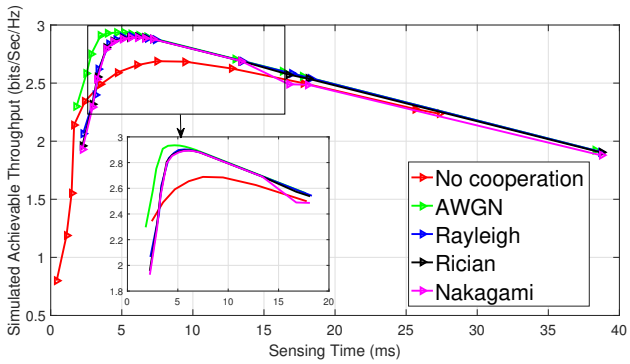


Fig. 8: Achievable Throughput Vs Sensing Time under Optimal Fusion operation for different channels (AWGN, Rayeigh, Rician and Nakagami)

From Fig. 6,7 and 8, performances for different channels for fixed decision fusion operation found. In every case AWGN channel provides better throughput then Rayleigh channel then Rician and finally Nakagami channel.

B. Against Frame length

Simulated achievable throughput against varying time frame for different channels and different fusion operations are analyzed. Here time frame varied from 50 ms to 200 ms and for each case sensing time kept fixed and it is 7 ms. The comparisons are made for different channels when fusion operation is fixed and among different fusion operation under

fixed channel.

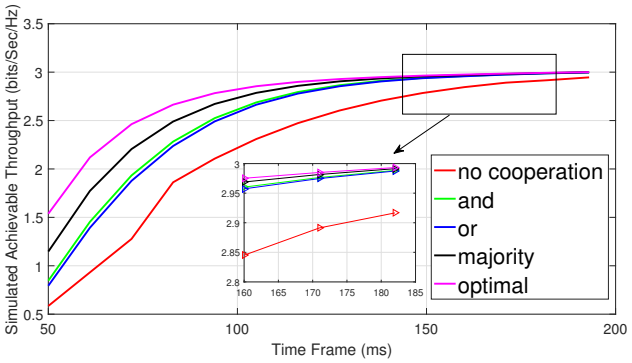


Fig. 9: Achievable Throughput Vs Time Frame under Rayleigh channel for different decision fusion operations (AND, OR, Majority, Oprimal Fusion)

From Fig. 9, simulated achievable throughput found under Rayleigh channel for AND, OR, Majority and Optimal Fusion operation. Here cooperation of 11 nodes for 10 dB SNR for different operations are compared with no cooperation state. Sensing time kept constant 7 ms and time frame varied from 50 ms to 200 ms. Increasing time frame length, throughput also increases. After 150 ms of time frame increasing rate of throughput gets slow for all decision fusion operations. Here Optimal Fusion operation provides better performance than others. From zoomed portion, it can be seen that for increasing time frame from 160 to 200 ms, throughput for Optimal Fusion operation and other operations also increases in slower rate. As in every cases optimal fusion operation provides better throughput than others. For changing time frame, achievable throughput for different channels under Optimal Fusion operation are analyzed.

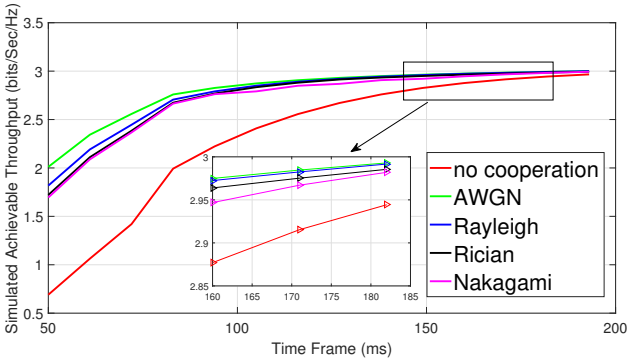


Fig. 10: Achievable Throughput Vs Time Frame under Optimal Fusion operation for different channels (AWGN, Rayeigh, Rician and Nakagami)

From Fig. 10, simulated achievable throughput found for different channel under Optimal Fusion operation. All these throughput found for different channels are also compared with no cooperation case under AWGN channel. The results are taken for SNR 10 dB and cooperation of 11 nodes. Here frame length varied from 50 ms to 200 ms. From Fig. 10, it can be seen that, increasing time frame also increases throughput as time for data transmission increases. From Fig. 10, AWGN channel provides better throughput among four then Rayleigh channel among two other fading channels. After time frame 75 ms, rate of increase of throughput gets slow for AWGN, Rayleigh, Rician and Nakagami channel. So for optimal fusion operation and for cooperation of 11 nodes optimal time frame can be set between 70-80 ms.

In ORN, QoS like throughput increases with the increase of cooperation of no. of OU. Among four fusion operations (AND, OR, Majority and Optimal Fusion), optimal fusion operation provides better throughput. In AND operation, decision is finalized when each nodes provides positive results about the presence of opportunistic nodes. If every nodes decide that there is a presence of PU, detection will be made. In OR operation, if any of nodes decides the presence of other node, detection will be made. As here P_{f} is greater, so performance of OR operation is less than AND operation. In Majority operation decision is made according to decisions of majority. So it provides better performance than AND and OR operation. In Optimal Fusion operation, decisions of nodes are prioritized according to their distance or performance. As a result, reliable nodes decision is valued most. Thus optimal fusion operation provides best result among four fusion operations.

Among Channels AWGN channel provides better throughput compared to others. AWGN channel is ideal noise channel. But in practice, there is real time noise present and that affects achieving maximized throughput like ideal cases. Rayleigh channel is slow fading channel where Rician and Nakagami are fast fading channel. When there is line of sight propagation component present between transmitter and receiver, Rician channel can be utilized, if there is multiple echo component present, Nakagami channel can be modeled. As in cognitive radio network for cooperative compressive sensing for estimating throughput for opportunistic users, there is no dominant line of sight component present and fading is slow. So Rayleigh channel can be modeled for these purpose compared to Rician and Nakagami fading channel.

Throughput performance against time frame provides optimum time frame for each situations. Above 150 ms time frame, increase in throughput gets slower and almost gets saturated. Optimal sensing time found around 5 to 8.7 ms for different decision fusion operations. Optimal sensing time varies from channel to channel and from fusion scheme to scheme.

V. Conclusion

In this paper, cooperation of different no. of opportunistic nodes in ORN are studied in terms of achievable throughput, sensing time and variable frame length. A comparative study on throughput performance has been carried out for different decision fusion operations and different fading channels. Trade-off between optimal sensing time and achievable throughput for variable frame length are presented for defined simulation setup for compressive sensing in cooperative ORN. Future contribution in performance analysis can be made by

estimating computational burden and throughput maximization can be made by evaluating mini slot sensing, other spectrum sensing approaches etc.

REFERENCES

[1] B. Wang and K. R. Liu, Advances in cognitive radio networks: A survey, *IEEE Journal of selected topics in signal processing, vol. 5, no. 1, pp. 523, 2011.*

[2] Y. C. Liang, Y. Zeng, E. Pehand A. T. Hoang, Sensing throughput tradeoff for cognitive radio networks, *in 2007 IEEE International Conference on Communications, pp. 53305335, IEEE, 2007.*

[3] A. W. Ahmad, H. Yang, and C. Lee, Maximizing throughput with wireless spectrum sensing network assisted cognitive radios, *International Journal of Distributed Sensor Networks, vol. 11, no. 6, p. 195794, 2015.*

[4] A. S. Kadhim and H. M. Alsabbagh, Throughput analysis for cognitive radio (cr) systems, *International Journal of Computer Networks & Communications, vol. 4, no. 4, p. 211, 2012.*

[5] S. Haykin, D. J. Thomson, and J. H. Reed, Spectrum sensing for cognitive radio, *Proceedings of the IEEE, vol. 97, no. 5, pp. 849 877, 2009.*

[6] F. F. Digham, M.-S. Alouini, and M. K. Simon, On the energy detection of unknown signals over fading channels, *IEEE transactions on communications, vol. 55, no. 1, pp. 2124, 2007.*

[7] H. Zhou, Throughput capacities and optimal resource allocation in multiaccess fading channels, 2003.

[8] S. S. Alam, L. Marcenaro, and C. S. Regazzoni, Improved throughput performance in wideband cognitive radios via compressive sensing, *in 2013 8th EUROSIM Congress on Modelling and Simulation, pp. 585590, IEEE, 2013.*

[9] G. Bhalerao and R. Zope, Ber improvement of ds-cdma with rake receiver using multipath fading channel, *International Journal of Electronics, Communication and Soft Computing Science & Engineering (IJECSCSE), p. 1, 2013.*

[10] Alam, Sk Shariful, M. Lucio, and Carlo S. Regazzoni. "Opportunistic Spectrum Access of Sparse Wideband in Stand-Alone and Cooperative Cognitive Radio Networks." *2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity). IEEE, 2015.*

[11] Al-Zadid Sultan Bin Habib, Shishir Mallick, Abu Shakil Ahmed, Sk. Shariful Alam, Abu Saleh Ahmad. " Performance Appraisal of Spectrum Sensing in Cognitive Radio Network." *2018 4th International Conference on Electrical Engineering and Information Communication Technology (iCEEiCT), iCEEiCT, 2018.*

Design of a High Voltage Closed Loop Charge Pumping System with a self-regulation mechanism

Alif Md. Asif Khan
Circuit & System Design
Ulkasemi Pvt. Limited
alif.khan.sebpo@gmail.com

Tahmida Islam
Circuit and System Design
Ulkasemi Pvt. Limited
Tahmida1993@gmail.com

Md Wasif Absar
Circuit and System Design
Ulkasemi Pvt. Limited
wasifabsar@gmail.com

Md Rubel Sarker
Circuit and System Design
Ulkasemi Pvt. Limited
Rubelsarkar161@gmail.com

Md Shariful Islam
Circuit and System Design
Ulkasemi Pvt. Limited
shariful.sojib04@gmail.com

***Abstract*— This paper proposes a complete design and simulation of a charge pumping system using a voltage regulator. The system has been implemented using a 6 stage Dickson Charge Pump with an operating frequency of 582 MHz and a maximum pumped-up voltage of 5.32 volts. The voltage level can be fixed to lower user-defined values utilizing the proposed regulatory components in the system. The generated output voltage of the charge pump is compared with a reference voltage produced from a low leakage on-chip Voltage Regulator thus forming a closed loop system. This feedback system allows the user to get the preferred output voltage from the Charge pump system. In this work, the targeted output voltage is 2.65 volts. The whole system is designed using CMOS 180nm process technology.**

Keywords—Charge Pump, Voltage Regulator, Bandgap Voltage, Biasing Current, Comparator, CMOS technology

I. INTRODUCTION

A Charge pump is an electronic circuit used for on-chip generation of higher voltages by transferring charges among capacitor networks utilizing clock phases. For certain applications such as writing or erasing data in non-volatile memories like flash memories or EEPROM, generating actuation voltages for MEMS devices, producing higher voltages in high gain photo-detectors and LED drivers, etc. charge pump is used.[1] [2]. Charge pump has specific applications in analog circuits for running analog switches and generating voltages to provide high dynamic range. [3]

There has been some considerable works in the past on the design and analysis of efficient charge pumps. A novel voltage regulation scheme using variable clock frequency for high voltage charge pump was proposed in [4] suited for neural implants. Two modified charge pump topologies were proposed in [5] which have higher efficiencies due to the elimination of reversion loss in conventional cross-coupled charge pump topologies. A novel tripler charge pump using switched bootstrapping technique with higher efficiency and faster startup was presented in [6]. For low power applications several design modifications were put forward in [7]-[10] and some comparative analysis of different charge pumps were conducted in [11]-[13].

With an endeavor to model state of the art analog circuits and investigate their performance we propose a design of a regulated system capable of controlling the pumping mechanism of the charge pump.

In this paper, Section II describes the different building blocks used for designing the whole system, Section III highlights the design trade-offs in Charge Pump, Section IV shows the simulation results of the proposed integrated circuit and Section V concludes our paper as well as discusses the future scope of our work.

II. FUNCTIONAL BLOCKS & DESIGN ARCHITECTURE

The diagram below shows the orientation of functional blocks and design architecture. Here, a clock system provides the frequency for charge pump stages and operational amplifier is used to turn the clock on or off. A reference voltage is needed for the comparator which can be generated from Bandgap reference and Voltage regulator circuit. A brief explanation regarding the operation of those blocks and connectivity is shown in following sections.

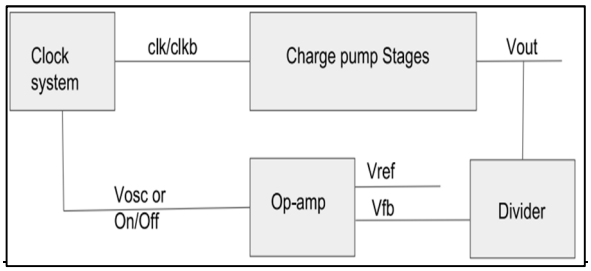


Fig. 1. System Overview of Proposed Charge Pump System

A. Charge Pump

For pumping up the voltage from the supply, charge pump uses multiple stages of capacitors which are charged and discharged by means of switches based on periodic input signals. At first stage the capacitor is charged to the supplied voltage and then at the next stage this stored charged along with the supplied voltage is applied to the load. The supplied voltage across the load increases proportionally by the number of stages used in the charge pump. In our design we have used a six stage Dickson Charge Pump and its output voltage is governed by the following equation:

$$V_{out} = V_{in} + N\left[\left(\frac{C}{C+C_s}\right).V_{\emptyset} - V_D\right] - V_D - \frac{NI_{out}}{(C+C_s)f} \quad (1)$$

where,

N=no of stages,

C= capacitor connected at each node,

Cs=stray capacitance of the device,

V_D=voltage drop across the gate drain connected MOSFET,

Vφ=voltage swing at each node due to capacitive coupling from the node.

I_{out}= Current with load.

The following figure shows the unit block of 6 stage Charge Pump.

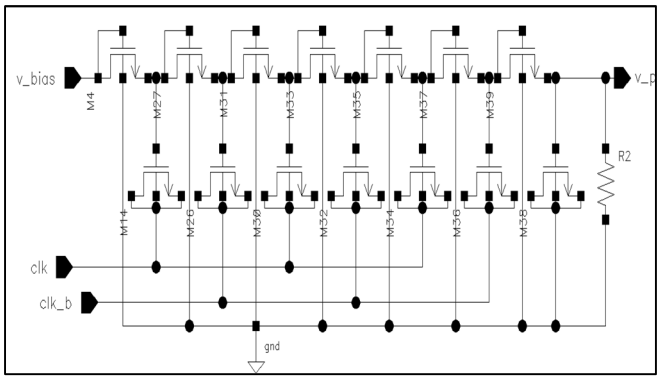


Fig. 2. Charge Pump stage.

B. Clock Generation

In order to generate the clock signals for the charge pump, voltage controlled oscillators (VCO) were used as it gives the advantage of controlling the oscillating frequency unlike the ring oscillators. The oscillating frequency of the VCO is related to the input voltage by the following equation:

$$f = \frac{I_D}{NC_{tot}V_{DD}} \quad (2)$$

where, N=No of stages & C_{tot}=sum of input and output capacitance of each stage. The figure below shows voltage controlled oscillator.

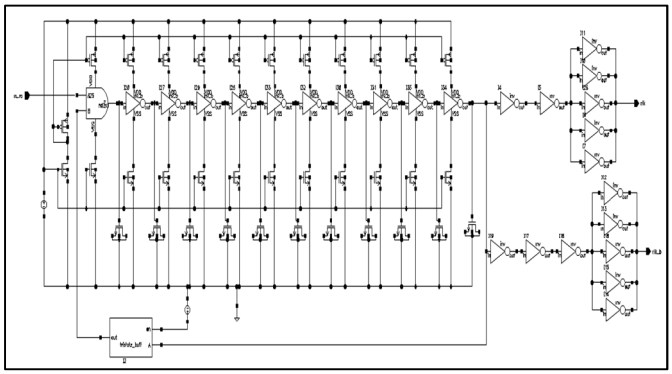


Fig. 3. Voltage Controlled Oscillator.

A frequency control pin has been used to regulate the oscillation frequency. Current in the branches directly varies with the frequency. By using current mirroring technique, same current is provided to all the branches ensuring a much stable operation. The output buffers are used to increase the driving strength and the tri-state buffer is used to make the edges of the clock sharper. Normally, generated oscillations had lower slew rate, but the tristate buffer improved the output.

C. Biasing Current Generator

To generate fixed amount of current to the bandgap voltage reference circuit, voltage trimmer and the operational amplifier of the system, a current generator independent of supply voltage variation has been used. Figure below shows a beta multiplier circuit where the circuit operation is controlled by the enable pin. Beta multiplier circuit has two stable operating points and one of them is when zero current flows in the circuit. Hence to operate the circuit at the desired operating point a startup circuit has been used. The output current for the circuit can be calculated from the following equation:

$$I_{REF} = \frac{2}{R^2 KP_n \frac{W_1}{L_1}}\left(1 - \frac{1}{\sqrt{k}}\right)^2 \quad (3)$$

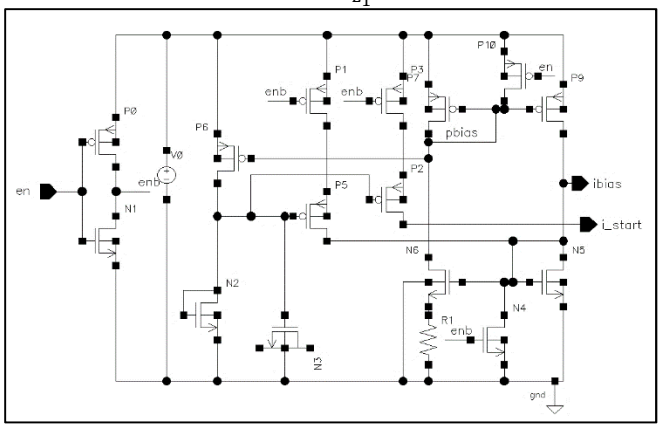


Fig. 4. Biasing Current Generator

As, this type of bias current generator has issues regarding high leakage current, multiple enabling input has been introduced. This helps the MOSFETS in the main generator circuit to be at totally OFF state during non-operating stages thus ensuring very less amount of leakage flowing through the circuit. Also, it allows us to eliminate the floating voltage from gates and keeping them at known values.

D. Bandgap Voltage Reference Circuit

This circuit is used to generate a constant voltage which is resistant to supply voltage, temperature and process technology variations. Combining the CTAT (Complementary to Absolute Temperature) behavior of the Bandgap energy of silicon and the PTAT (Proportional to Absolute Temperature) behavior of the thermal voltage a temperature resilient voltage reference circuit is designed.

If V_{BE} is considered as Base to Emitter voltage for BJT then,

$$\Delta V_{BE} = V_{BE_1} - V_{BE_2} = V_t \ln\left(\frac{I_1}{I_{S_1}}\right) - V_t \ln\left(\frac{I_2}{I_{S_2}}\right) \quad (4)$$

Again, $$V_{R_2} = \Delta V_{BE} = V_t \ln\left(\frac{I_1 I_{S_2}}{I_2 I_{S_1}}\right) \quad (5)$$

$$I_2 = \frac{\Delta V_{BE}}{R_2} = \frac{V_t}{R_2}\ln\left(\frac{I_{S_2}}{I_{S_1}}\right)) \quad (6)$$

Now, $$V_{ref} = V_{BE2} + V_{R2} + V_{R0,R1}$$

$$V_{ref} = V_{BE2} + V_{R2} + I_2 R_0 \quad [R_o = R_1]$$

Using (4), (5), and (6), it can be written as,

$$V_0 = V_{BE_1} + (1 + \frac{R_1}{R_2})V_t \ln(\frac{I_{S_2}}{I_{S_1}}) \quad (7)$$

The above derived equation is used to set the values of resistors in the circuit.

A differential amplifier was used in the Bandgap Reference circuit to keep same voltages on both sides of the circuit. Differential amplifier causes much noise, especially in Bandgap circuits. In this circuit, an extra low pass filter improved the effect of noise on the circuit.

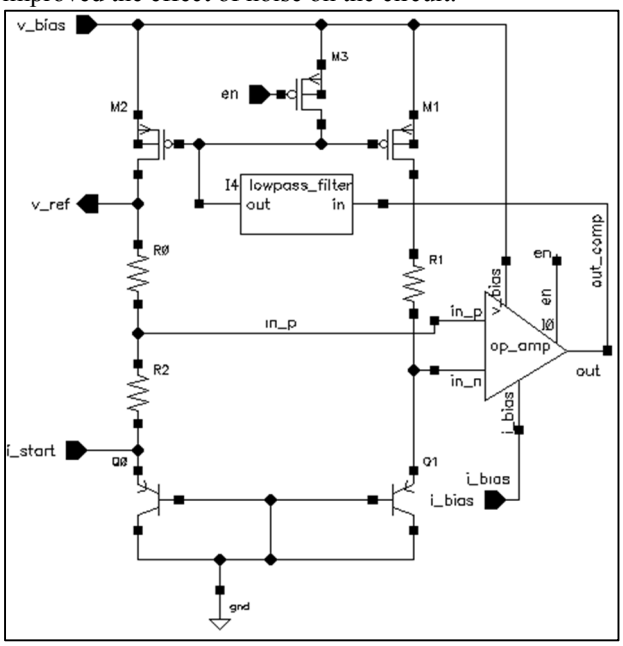


Fig. 5. Bandgap Reference Voltage Generator

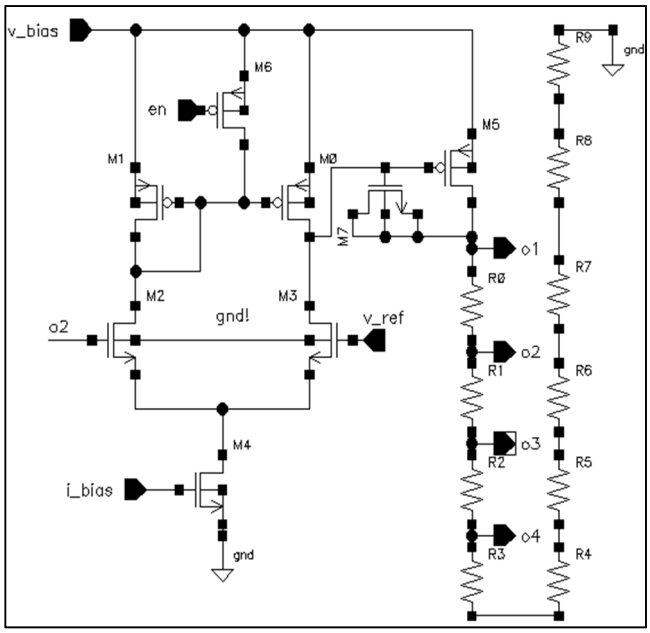


Fig. 6. Voltage Trimmer Circuit

E. Voltage Trimmer Circuit

In order to get different voltage levels from the output of the voltage reference circuit, a voltage trimmer is used. It adds flexibility of using the system at multiple voltage levels. The comparator keeps the voltage level of the two inputs at the same value where its output is eventually trimmed at different voltage levels. User can choose any suitable voltage as the reference value from this Voltage Trimmer circuit.

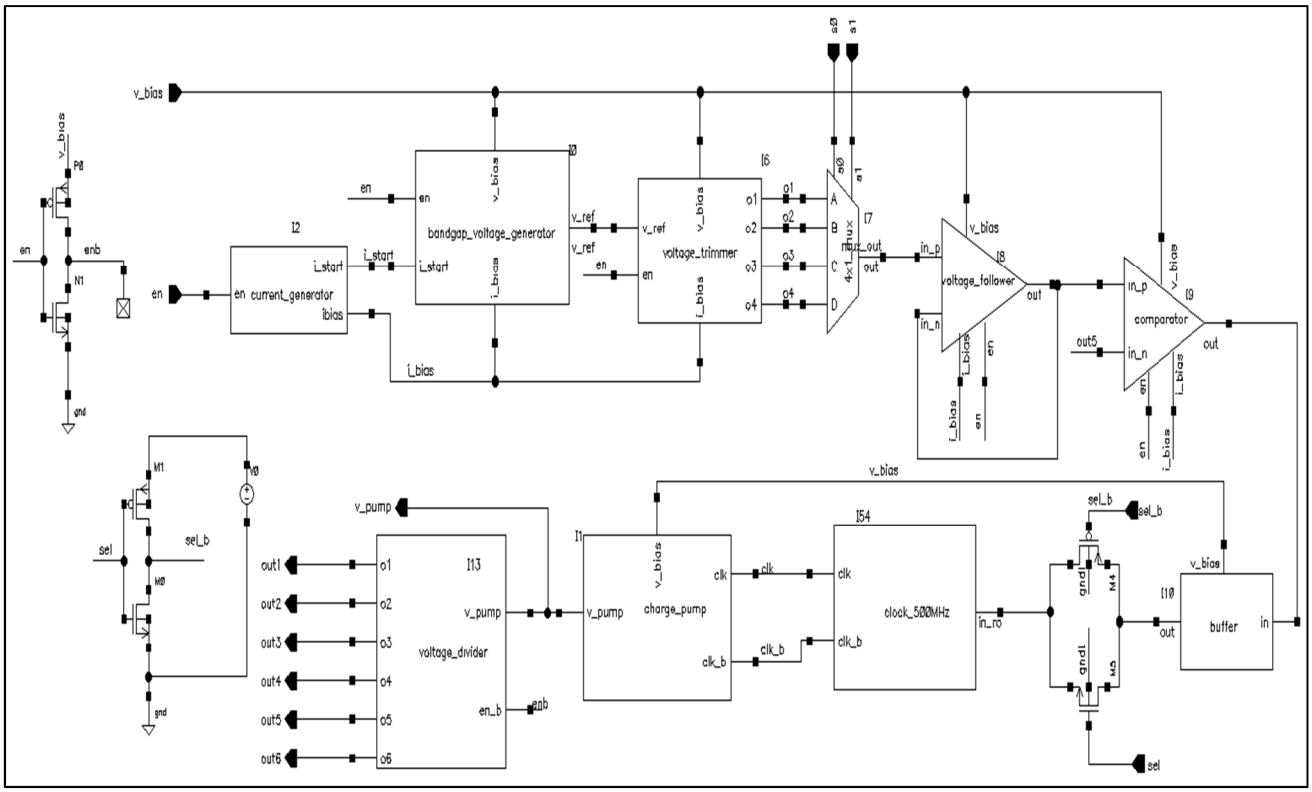


Fig. 7. Top Level schematic of High Voltage Charge Pump System

F. *Top Level of System*

The figure above shows the schematic of the proposed system. The output voltage of the Charge Pump goes into a Voltage Divider to generate a comparable voltage for the comparator. Another input of the Comparator is a user defined value which comes from the output of the Voltage Trimming circuit. Initially feedback voltage from the Voltage Divider is lower than the reference voltage which compels the comparator to give a logic high at its output enabling the oscillators. With the non-overlapping clock signals created by the oscillator, Charge Pump starts pumping up the voltage. The pumped up voltage results into an increased feedback voltage which when surpasses the value of the reference voltage generates a logic low at the output of the comparator. As a result the oscillator is turned off and the charge pump will consequently stop pumping up the voltage until the feedback voltage is lower than the reference voltage. In that case the comparator will give logic high and enable the charge pump again. From this process a fixed scalable voltage can be achieved from the system

III. Design Tradeoffs in a Charge Pump

The behavior of a charge pump usually depends on a few decisive parameters which have been discussed in detail in the following section.

A. *Effect of stage capacitance*

Increasing the value of stage capacitance increases the storage capacity of the pump and thus increases the output voltage and current. On the contrary much delay is added to the circuit and thus makes the circuit slower. Almost 400ns of time was recorded to get a stable voltage level while using capacitors of 20pF. But the settling time has increased comprehensively when capacitors' values were increased to 80pF.

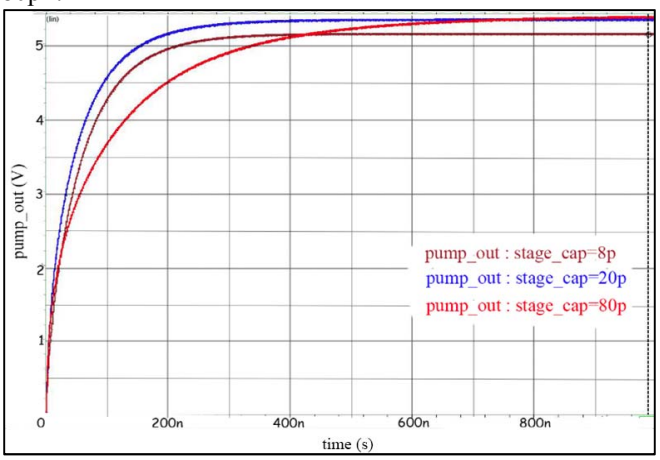


Fig. 8. Effect of stage capacitance for charge pump

Decreasing the value of the stage capacitors reduces the maximum output voltage to 5.13v.

B. *Effect of load capacitance*

The load capacitor when reduced to a lower value makes the ripples more distinct in the signals. In fig. 9, the effect of varying the output capacitor has been shown. Adding more load capacitance to the circuit can smoothen the output curve at a greater scale but at the expense of adding more delay to the output which has been illustrated in fig. 9. Here the load capacitance was increased from 10pf to 50pf which consequently delayed the signals reaching their maximum value.

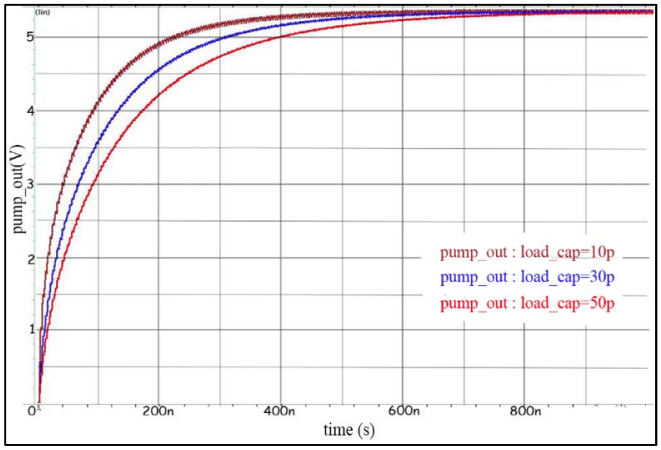


Fig. 9. Effect of load capacitance for charge pump

C. *Effect of load Resistance*

Reducing the load resistance increases the current across the load and consequently reduces the output voltage. The load resistance also has an inverse effect on the ripple of the signals. At resistance value of 2.8k, current increased to 1.45mA and voltage decreased to 4.3V. Load resistance was increased from 2.8k to 20.8k which dropped the current to a value of 270uA but the voltage increased to 5.61V. Also the ripples were minimized.

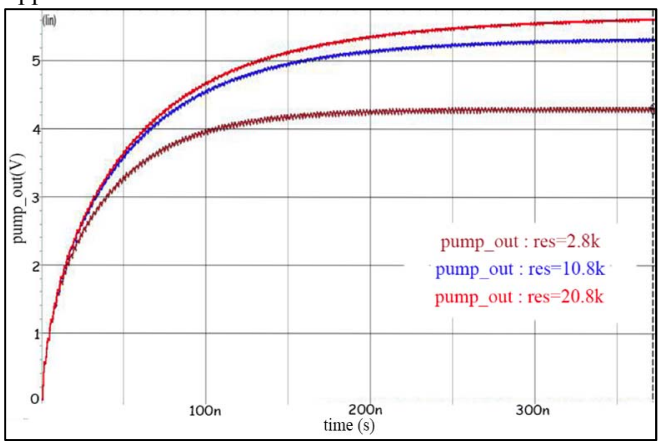


Fig. 10. Effect of load resistance on output voltage

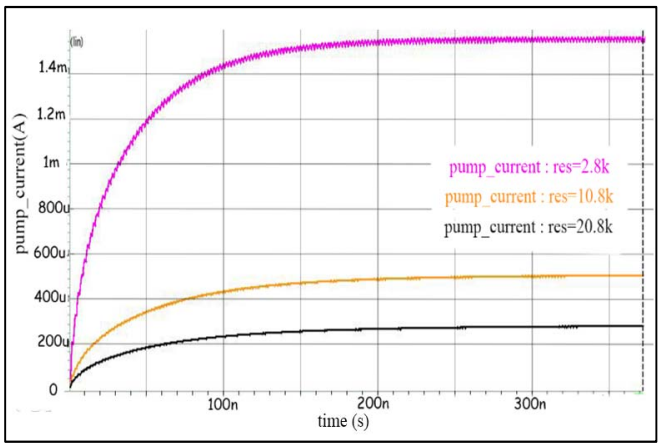


Fig. 11. Effect of load resistance on output current

D. Effect of Frequency

Frequency is always a dominant factor for a charge pump. When the time period was increased from 800ps to 2ns, the output takes almost 185ns to become stable. It also has an inverse effect on the ripple of the signals which can be seen in the figure below.

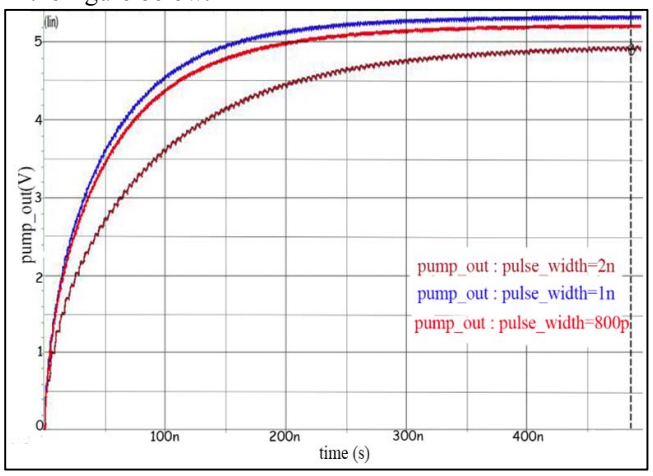


Fig. 12. Effect of frequency on output voltage

Increasing the frequency too much hinders the proper charging and discharging of the stage capacitors thus resulting into fewer charge accumulation and degrading the performance of the charge pump. It is shown from the fig. 12 that, output voltage got degraded and also ripples were added at the frequency of 500 MHz. Frequency also effects the output current in the same manner. Fig. 13 shows the effect of frequency on output current.

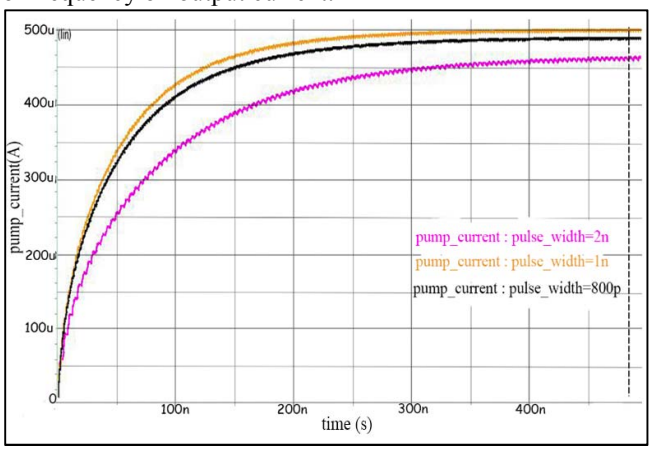


Fig. 13. Effect of frequency on output current

IV. Simulation Results

This section discusses the simulation results of the different blocks considering multiple conditions. It validates the authenticity of the proposed system.

In fig. 14. voltage profile at different nodes of the Current Generator is shown. The enable signal is used to start the current generator circuit and a startup circuit is used to inject voltage into the current sinking NMOS network turning them ON. The current generator circuit has been used to generate a current of 44uA. It is observed that, current flowing through startup circuit is about 25uA for a short period of time until a stable current supply has been ensured.

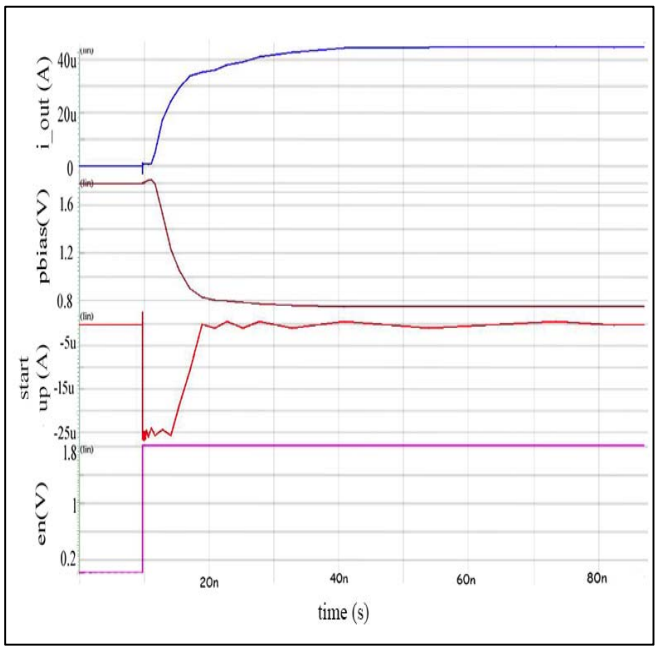


Fig. 14. Waveform of biasing current and startup current obtained from Current Generator

Fig. 15. represents results for Bandgap Reference Generator. When the enable signal is given, the startup circuit injects voltage into the reference circuit which initiates its operation. The operational amplifier used helps stabilizing the voltage on both sides of the circuit and a constant voltage of 1.24V is generated. The graph below shows an almost accurate temperature independence of the circuit. For a temperature sweep of -40 to 149 degree Celsius the circuits shows a voltage variation of 0.009V which is shown in fig. 16.

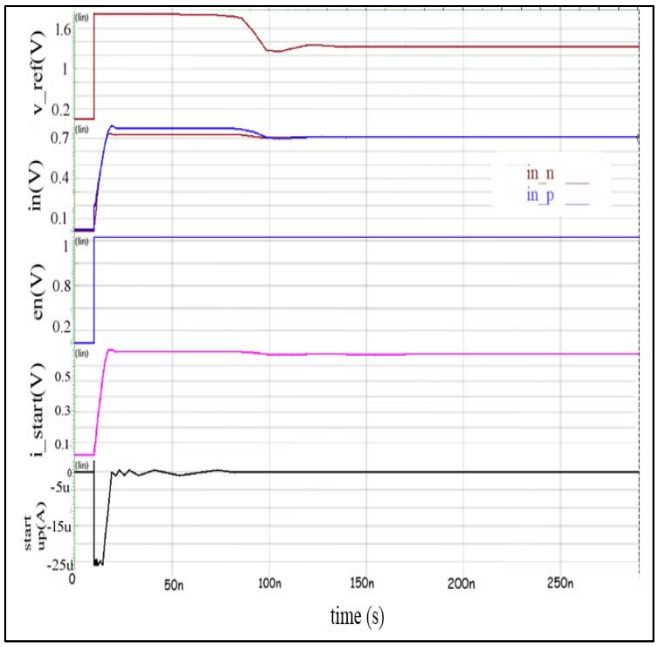


Fig. 15. Transient analysis of Bandgap Reference voltage

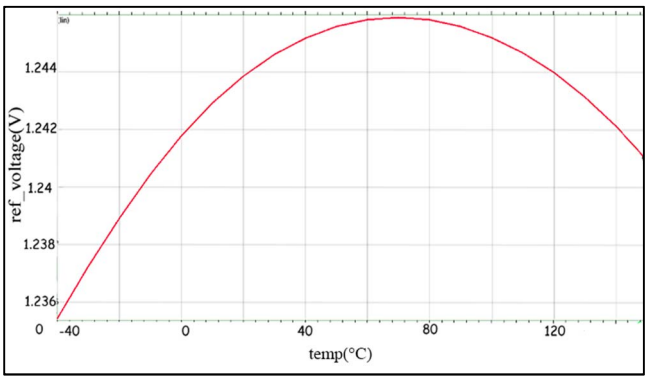


Fig. 16. DC analysis of Bandgap Reference voltage sweeping temperature

It states that, Reference Generator is independent of temperature variation. Once again very low leakage is ensured as startup circuit only draws current while enabling the main circuit.

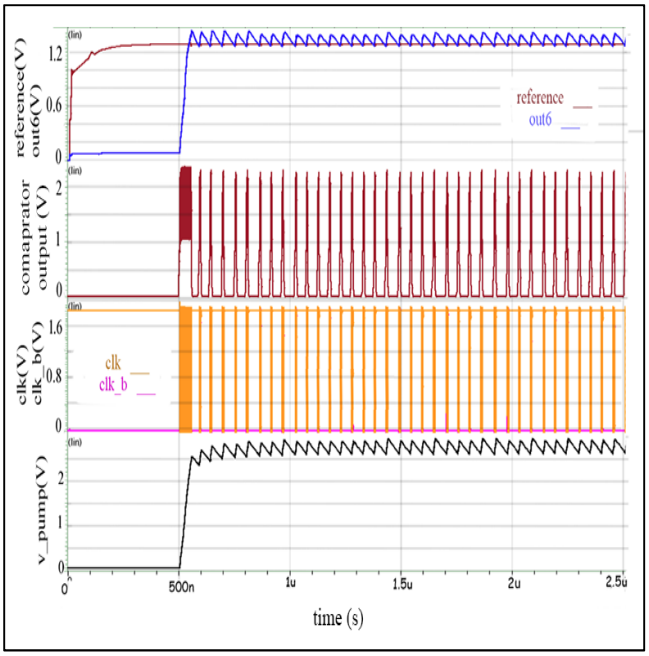


Fig. 17. Output wave shape of Charge Pump along with self-regulated mechanism

In order to control the clocks of the charge pumps we have used the voltage from the bandgap reference and a divided value of the pumped up voltage. When the derived output voltage of the charge pump goes below the reference voltage the clocks are switched on to activate the charge pump and the voltage is pumped up to a desired value of 2.64 volts.

V. Conclusion

The paper reported a low leakage biasing circuit by introducing multiple enabling inputs and flexible operation of the system at multiple voltage levels by using a voltage trimmer. All design and simulation are done under Cadence – Virtuoso environment. The results of the developed system were analyzed under different conditions which validate the purpose of the proposed method. Further work will include the improvement of the charge pump for lower voltage and higher frequency. Higher frequency introduces more noise. So the effect of noise will be a concern for future work.

REFERENCES

[1] D.S. Hong, M.N. El-Gamal, “Low operating voltage and short settling time CMOS charge pump for MEMS applications”, Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS '03

[2] Mustapha El Alaoui, Fouad Farah, Karim El Khadiri, Hassan Qjidaa, Abdellah Aarab, Rachid E Alami, Ahmed Lakhassassi, “Analysis and design of dickson charge pump for EEPROM in 180nm CMOS technology”, 2018 International Conference on Intelligent Systems and Computer Vision (ISCV)

[3] R. St. Pierre, ‘‘Low-power BiCMOS op amp with integrated current mode charge pump’’, IEEE J. Solid State Circuits, vol.35, no.7, 2000, pp.1046-1050.

[4] Alfian Abdi, Hyung Seok Kim, Hyouk-Kyu Cha, “A High-Voltage Generation Charge-Pump IC using Input Voltage Modulated Regulation for Neural Implant Devices”, IEEE Transactions on Circuits and Systems II: Express Briefs (Early Access)

[5] Boyu Shen, Matthew L. Johnston, “Zero Reversion Loss, High-Efficiency Charge Pump for Wide Output Current Load Range”, 2018 IEEE International Symposium on Circuits and Systems (ISCAS)

[6] Younis Allasasmeh, Stefano Gregori, “Fast-Startup High-Efficiency Tripler Charge Pump in Standard 0.18-µm CMOS Technology”, 2018 IEEE International Symposium on Circuits and Systems (ISCAS)

[7] Xiaoxue Jiang, Xiaojian Yu, Jie Chen, “A low-voltage charge pump with improved pumping efficiency”, 2017 IEEE International Symposium on Circuits and Systems (ISCAS)

[8] Steve Ngueya W., Julien Mellier, Stephane Ricard, Jean-Michel Portal, Hassen Aziza, “Ultra Low Power Charge Pump with Multi-Step Charging and Charge Sharing”, 2016 IEEE 8th International Memory Workshop (IMW)

[9] Rahma Aloulou, Hassene Mnif, Frederic Alicalapa, Mourad Loulou, “An improved MOS charge pump circuit for low voltage operations and wireless sensor applications”, ICM 2011 Proceeding

[10] Labonnah F. Rahman, M.B.I. Reaz, M.A.M. Ali, “A Low Voltage Charge Pump Circuit for RFID Tag EEPROM”, 2011 Fourth International Conference on Emerging Trends in Engineering & Technology

[11] Younis Allasasmeh, Stefano Gregori, “A performance comparison of dickson and fibonacci charge pumps”, 2009 European Conference on Circuit Theory and Design

[12] Harshita Dadhich, Vijendra Maurya , Kumkum Verma, Sanjay Jaiswal, “Design and analysis of different type of charge pump using CMOS technology”, 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)

[13] Vipul Bhatnagar, Sujata Pandey , Pradeep Kumar, “Analysis and implementation of ripple and area efficient charge pump circuits”, 2015 2nd International Conference on Signal Processing and Integrated Network

Effects of Wave Function Penetration on Gate Capacitance and Threshold Voltage of p-Channel Double Gate Junction Less Transistor

Md. Mohsinur Rahman Adnan[1]*, Mohammad Shahidur Rahman[1]
[1] Department of Computer Science and Engineering, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh
*e-mail: adnanmohsin27@gmail.com

***Abstract*—Wave function penetration phenomenon had been studied for a p-channel Double Gate Junction Less Transistor in terms of Capacitance-Voltage (C-V) characteristics. It has been found that the gate capacitance differs visibly for the two cases (i.e. with and without penetration) only in the accumulation mode of device operation. As a result, it is safe to assume that the occurrence of threshold condition is independent of this phenomenon for the studied device, which had been established by extracting threshold voltage from the simulated C-V curves.**

***Index Terms*—Junction Less Transistor, Self-Consistent Schrödinger-Poisson Solver, Wave Function Penetration, C-V characteristics, Threshold Voltage.**

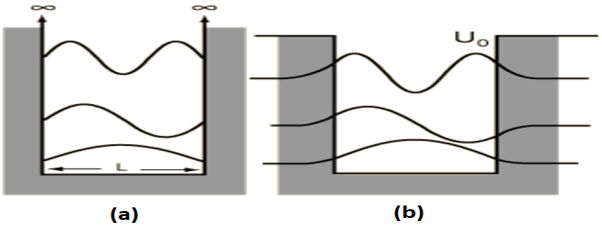


Fig. 1: Wave Function Penetration (WFP) effect in Double Gate Junction Less Transistor (DGJLT). The structure is basically a quantum well with 1D confinement. Being a finite potential well, it is safe to assume some penetration into gate oxide occurs as shown in (b), instead of ideal case with infinite barriers as depicted in (a). [9]

I. Introduction

The need for quantum mechanical modeling of Field Effect Transistor (FET) devices' characteristics was first described and practically implemented by Stern [1]. To model the ultra-small devices with micro-to-nanometer range feature lengths, Stern employed self-consistent methodology which solved Schrödinger and Poisson equations iteratively till a complete description of charge distribution and potential profile inside the FET structure was found. With the scaling of FET devices to ever decreasing sizes, self-consistent modeling has become a popular method to model their electrostatics now-a-days. In his model, Stern assumed that the FET structures were quantum confined i.e. well (1D) or wire (2D) with infinite potential barriers surrounding the channel [1]. As a result, the carrier wave function was assumed going to zero abruptly at channel-oxide interface. But this is not really a valid scenario as the potential barrier arising from the difference in band structure of channel and oxide materials' is actually only a few electron volts high [2]. As a result, assuming no penetration of carrier wave function into the gate oxide of the device gives an incomplete picture of the charge distribution and potential profile inside the FET structure [2]. So, in order to obtain a more complete description of charge distribution, potential profile and other electrostatic properties like gate capacitance and threshold voltage, a few previous works [2], [3], [4] have allowed carrier wave function to penetrate into gate oxide. All of them studied Inversion Mode (IM) MOSFET, so there is a scarcity of such modeling for other FET devices in literature.

Recently, a new FET device called the Junction Less Transistor (JLT) has been proposed [5]. While IM-MOSFET has its carriers conducting at channel surface just below the channel-oxide interface [3], the JLT device's carrier conduction mostly happens at or near the channel center [5], which is the furthest point from channel-oxide interface inside the device structure. Specially during sub-threshold and near threshold operation, the bulk conduction mechanism of JLT can not be ignored [5]. As a result, it is quite possible that considering wave function penetration effect while self-consistently modeling JLT electrostatics is not so important as in the case of IM-MOSFET modeling.

Indeed, a recent study of Capacitance-Voltage (C-V) characteristics of n-channel Double Gate (DG) JLT device has demonstrated that the occurrence of threshold condition in n-JLT is independent of penetration into gate oxide by carrier wave function [6]. The study also showed that wave function penetration phenomenon has no effect on the gate capacitance in full and partial depletion modes of n-JLT operation [6]. Even though definite proof of theory now exists for n-JLT, no such study was carried out for p-JLT yet. As the results for deep submicron n- and p-MOSFETs were found to be different in [2], it is necessary to carry out a similar study for p-JLT. So, the present work aims to first simulate the p-DGJLT C-V characteristics both allowing and disallowing wave function penetration and then study gate capacitance and threshold voltage to find out if they are as affected as in the case of n-JLT by the occurrence of this phenomenon.

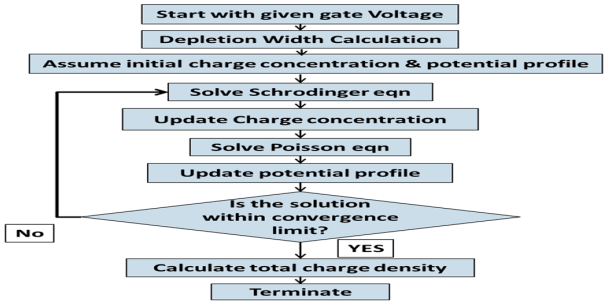


Fig. 2: Self consistent solver [1].

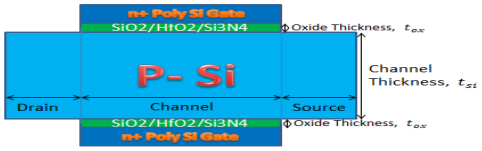


Fig. 3: Double Gate Junction Less Transistor (DGJLT) structure.

II. Simulation Details

A MATLAB based 1D Schrödinger-Poisson solver had been created following the steps described in [1] to simulate the C-V curves, and is depicted in Fig.2. Necessary modification for the physics of JLT as described in [7] had been used in building the simulator. Necessary parameters of gate insulator (i.e. K value, conduction band offset and band gap) were obtained form [3]. The valance band offset was calculated by subtracting the conduction band offset from the bandgap offset. DOS and Quantization effective masses of holes and degeneracy information of Si were obtained form [2]. The effective mass of hole in oxide is assumed to be $0.5m_0$, m_0 being electron's rest mass [2]. Whenever wave function penetration had to be ignored the channel-oxide valance band offset was set to a very high value (e.g. 4000 eV), following [4]. The threshold voltage was extracted depending on the work of [8], which suggests that when mobility variation is ignored, the position of the maximum peak of $|\frac{dC_{gate}}{dV_{gate}}|$ vs. V_{gate} curve gives the threshold voltage V_{th}. For p-channel JLT, absolute value $|\frac{dC_{gate}}{dV_{gate}}|$ had been plotted to observe the change only, ignoring its direction. The range of uncertainity for all reported threshold voltage values is limited to within ± 0.1 V.

III. Results

The studied structure is shown in fig.3. The channel is made of p-Si while the gate is made of n+ poly Si. The parameters are chosen as doping density, $N_A = 10^{19}/cm^3$, channel thickness, $t_{si} = 10nm$, oxide material, SiO_2 and Effective Oxide Thickness (EOT), $t_{ox} = 2nm$ and have been maintained for all simulations unless otherwise specified.

To find the effect of changing oxide material, along with SiO_2, other gate oxides used were HfO_2 and Y_2O_3. Fig.4 shows the C-V curves. The $|\frac{dC_{gate}}{dV_{gate}}|$ vs V_{gate} curves are shown in Fig.5.

To find the effect of changing channel thickness, it was varied to 6, 8 and 10 nm. Fig.6 shows the C-V curves. The $|\frac{dC_{gate}}{dV_{gate}}|$ vs V_{gate} curves are shown in Fig.7.

To find the effect of changing oxide thickness, it was varied to 1, 1.5 and 2 nm. Fig.8 shows the C-V curves. The $|\frac{dC_{gate}}{dV_{gate}}|$ vs V_{gate} curves are shown in Fig.9.

To find the effect of changing channel doping density, it was varied to $10^{19}/cm^3, 5*10^{18}/cm^3$ and $10^{18}/cm^3$. Fig.10 shows the C-V curves. The $|\frac{dC_{gate}}{dV_{gate}}|$ vs V_{gate} curves are shown in Fig.11.

For the defined device structure, the metal to semiconductor work function (also known as the flat-band voltage) [7] is $-1.1V$. In the C-V curves, saturation behavior occurs beyond this voltage. This is also the voltage beyond which gate capacitances differ due to penetration. So, from observing the results it can be said that regardless of penetration the gate capacitances are visibly same in the C-V curves up until the accumulation mode of operation (i.e. saturation behavior) [7].

The maximum peak in the $|\frac{dC_{gate}}{dV_{gate}}|$ vs. V_{gate} curves occurs at same V_{gate}=V_{th} regardless of penetration.

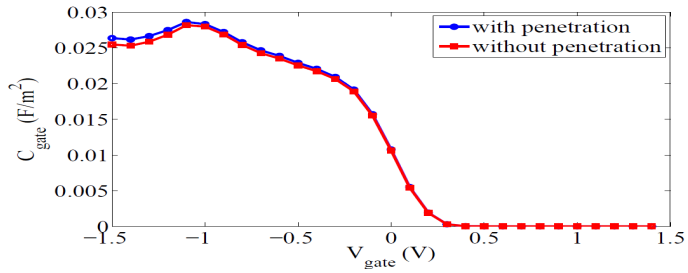
IV. Conclusion

In p-JLT, initially a space charge region arising from the acceptor doping covers the whole channel. As a gate voltage is applied and gradually increased in positive direction, the depletion region reduces and free holes start to gather at the channel center. With sufficient holes gathered at the channel center to nullify the depletion charges, a string like path of neutral-Si from source to drain is created [5]. After this threshold condition has occurred, applying a suitable drain to source voltage will make conduction possible. As channel center is the furthest point inside the device form channel-oxide interface so it is expected that wave function penetration into gate oxide should have no effect on threshold voltage, which has been proven from the results of this work.

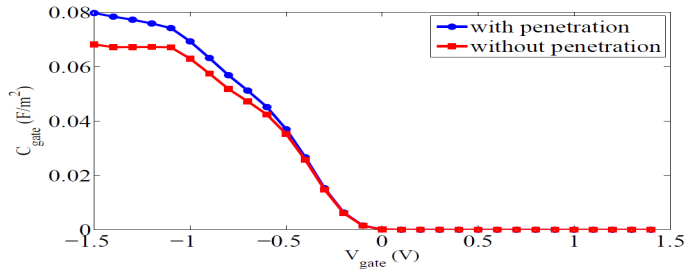
Moreover, the C-V curves show the wave function penetration to have no perceivable effect on gate capacitance in full or partial depletion mode of device operation. The gate capacitance is defined as,

$$C_{gate} = \frac{q\delta(p - N_A^-)}{\delta V_{gate}}$$

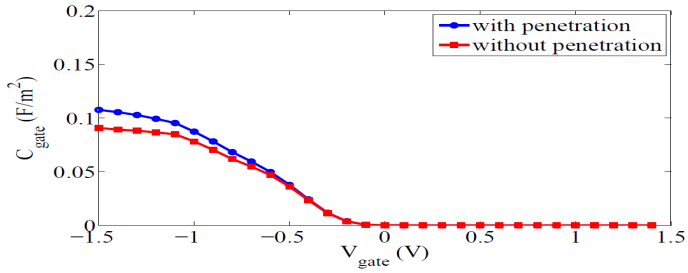
where, q=charge of a hole, qp=mobile charge arising from holes(+ve) and qN_A^-=depletion charge arising from acceptors (-ve). In this equation, only the mobile charge, qp is variable. This term is affected by wave function penetration only when the holes are present near channel-oxide interface. Only then the probability of holes tunneling into gate oxide takes on a moderate value. This does not happen in full or partial depletion mode of operation. Only beyond flat-band, when all of p-JLT channel is available for conduction [5], then the mobile charge term is affected by wave function penetration. So, the observations in this work comply with both theoretical notions and the results for n-JLT.



(a) SiO_2

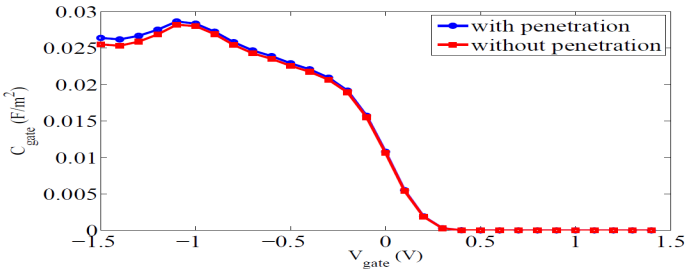


(b) Y_2O_3

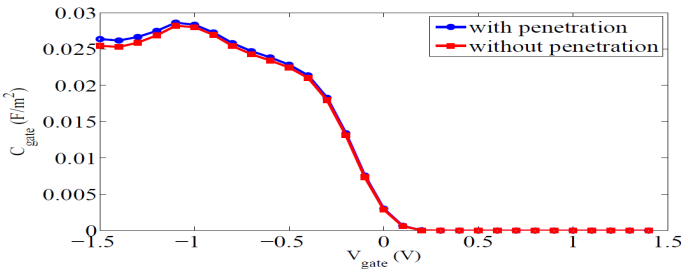


(c) HfO_2

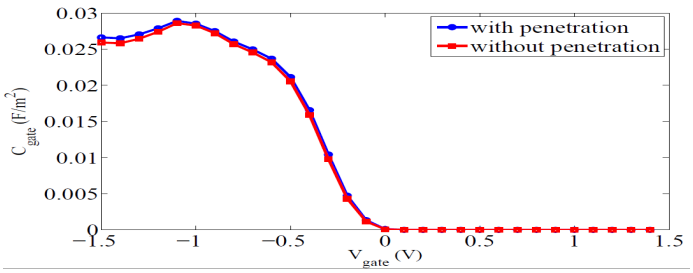
Fig. 4: C-V characteristics with and without oxide penetration with varying oxide material.



(a) Channel thickness 10 nm

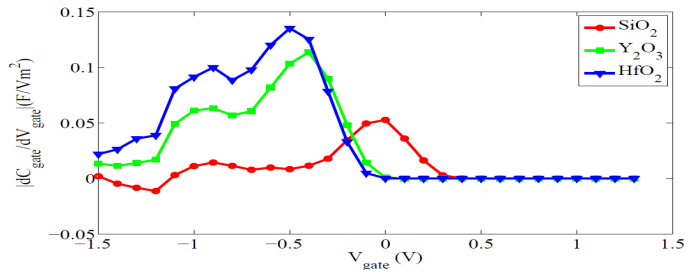


(b) Channel thickness 8 nm

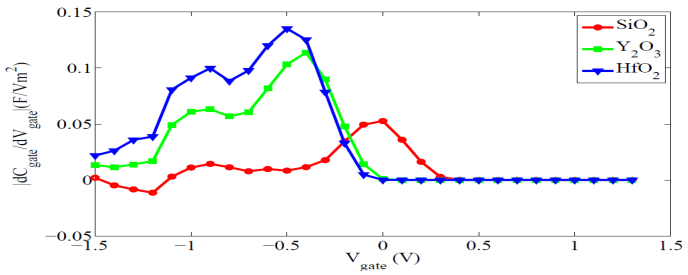


(c) Channel thickness 6 nm

Fig. 6: C-V characteristics with and without oxide penetration with varying channel thickness.

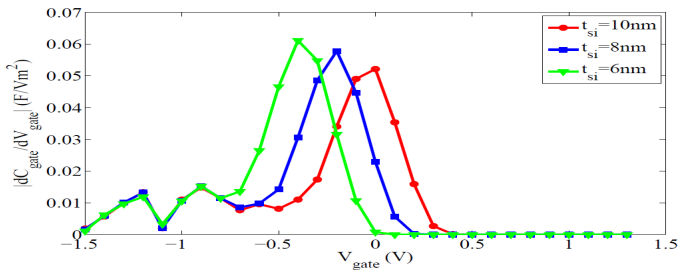


(a) Without Oxide Penetration

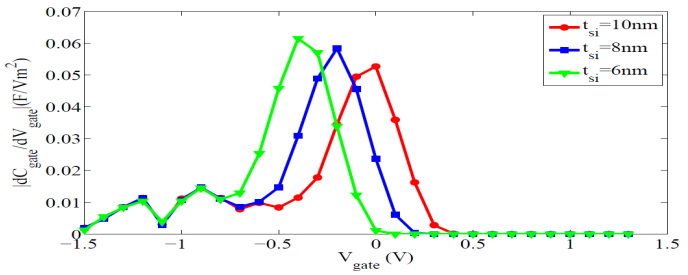


(b) With Oxide Penetration

Fig. 5: Threshold voltage extracted without and with oxide penetration with varying oxide material. The maximum peaks occur at the same gate voltages (V_{th}) i.e. 0 V for SiO_2, -0.4 V for Y_2O_3, and -0.5 V for HfO_2 regardless of penetration.

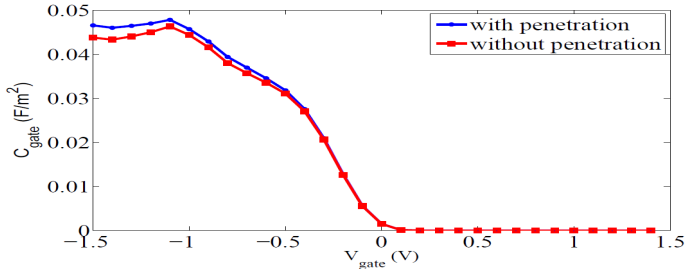


(a) Without Oxide Penetration

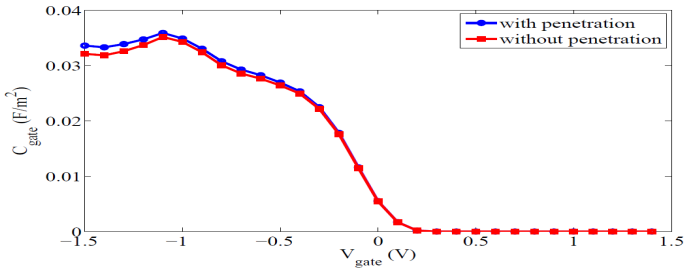


(b) With Oxide Penetration

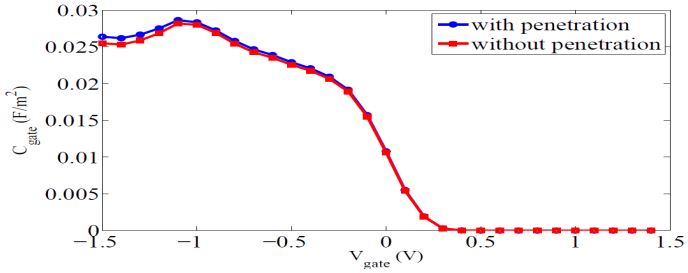
Fig. 7: Threshold voltage extracted without and with oxide penetration with varying channel thickness. The maximum peaks occur at the same gate voltages (V_{th}) i.e. 0 V for 10 nm, -0.2 V for 8 nm, -0.4 V for 6 nm regardless of penetration.



(a) Oxide thickness 1 nm

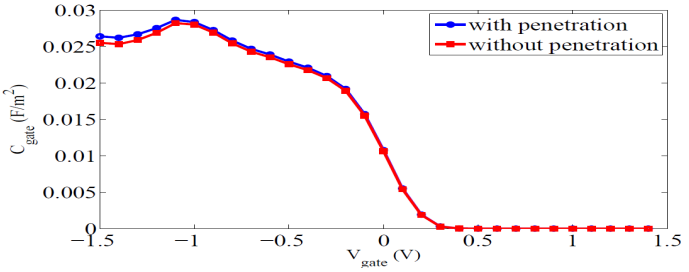


(b) Oxide thickness 1.5 nm

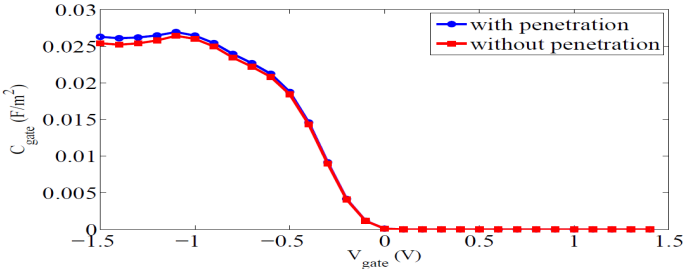


(c) Oxide thickness 2 nm

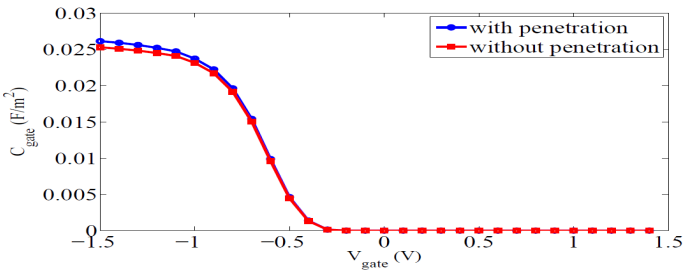
Fig. 8: C-V characteristics with and without oxide penetration with varying oxide thickness.



(a) Doping density $10^{19}/cm^3$

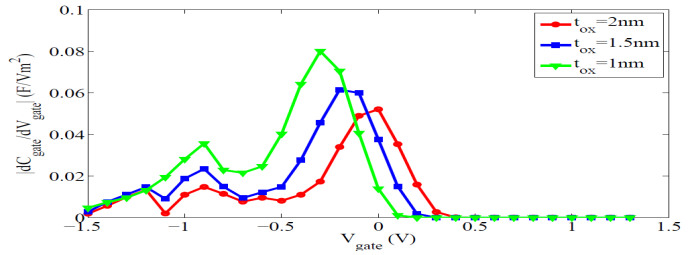


(b) Doping density $5 * 10^{18}/cm^3$

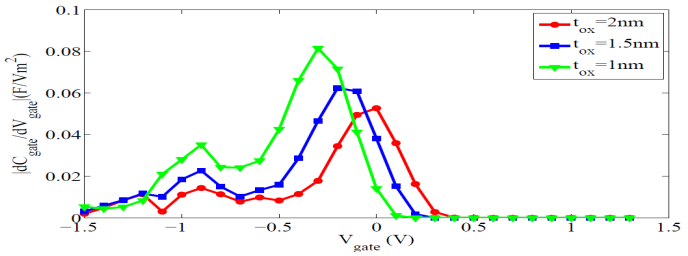


(c) Doping density $10^{18}/cm^3$

Fig. 10: C-V characteristics with and without oxide penetration with varying doping density.

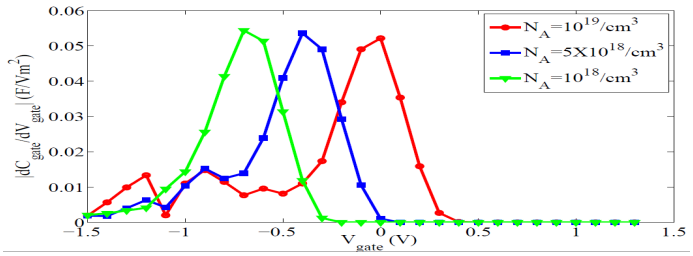


(a) Without Oxide Penetration

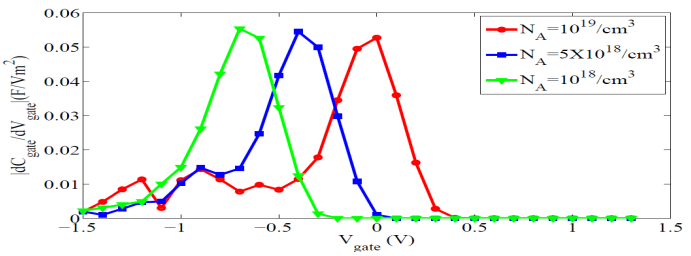


(b) With Oxide Penetration

Fig. 9: Threshold voltage extracted without and with oxide penetration with varying oxide thickness. The maximum peaks occur at the same gate voltages (V_{th}) i.e. -0.3 V for EOT 1 nm, -0.2 V for EOT 1.5 nm and 0 V for EOT 2 nm regardless of penetration.



(a) Without Oxide Penetration



(b) With Oxide Penetration

Fig. 11: Threshold voltage extracted without and with oxide penetration with varying doping density. The maximum peaks occur at the same gate voltages (V_{th}) i.e. 0 V for $10^{19}/cm^3$, -0.4 V for $5 * 10^{18}/cm^3$ and -0.7 V for $10^{18}/cm^3$ regardless of penetration.

REFERENCES

[1] Stern F., "Self-consistent results for n-type Si inversion layers". *Physical Review B*, Vol:5, Issue:12, pp:4891-99, June 1972.

[2] A.Haque and M.Z.Kauser, "A Comparison of Wave Function Penetration effects on Gate Capacitance in Deep submicron n-and p-MOSFET.", *IEEE Transaction on Electron Devices*, Vol. 49, No. 9, September 2002.

[3] A.I.Khan, M.K.Ashraf and Q.D.M.Khosru, "Effects of wave function penetration on gate capacitance modeling of nanoscale double gate MOSFETs", IEEE International Conference on Electron Devices and Solid-State Circuits, 2007, pp.137-140.

[4] M.K.Alam, "Modeling of Charge Quantization and Wave Function Penetration Effects in Ultra Thin Body Double Gate MOSFET." M.Sc. Thesis, Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, 2007.

[5] J.P.Colinge, C.W.Lee, A.Afzalian, N.D.Akhavan, R.Yan, I.Ferain, P.Razavi, B.ONeill, A.Blake, M.White, A.M.Kelleher, B.McCarthy and R.Murphy, "Nanowire transistors without junctions," *Nature Nanotechnology*, Vol. 5, No. 3, pp. 225-229, 2010.

[6] M.M.R. Adnan and M. S. Rahman, "Impact of Wave Function Penetration into Gate Oxide on Gate Capacitance and Threshold Voltage of n-Channel Double Gate Junction Less Transistor", IEEE International Conference on Advancement in Electrical and Electronic Engineering, 2018.

[7] M.M.R.Adnan and Q.D.M.Khosru, "A Simple Analytical Model of Threshold Voltage for P-Channel Double Gate Junctionless Transistor", IEEE International Conference on Electron Devices and Solid-State Circuits, 2018, pp.1-5.

[8] T.Rudenko, A.Nazarov, I.Ferain, S.Das, R.Yu, S.Barraud, and P.Razavi, "Mobility enhancement effect in heavily doped junctionless nanowire silicon-on-insulator metal-oxide-semiconductor field-effect transistors", *Applied Physics Letters*, vol.101, 2012.

[9] "Particle in a Finite Walled Box." Accessed: April 12, 2019. [Online], Available: http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/pbox.html

Design of CMOS On-Chip Millimeter-Wave Transformer Coupled Balun and Power Divider-Combiner with Optimal Amplitude and Phase Imbalance

Md. Sazzad Hossain[1], Minoru Fujishima[2], Takeshi Yoshida[2], Shuhei Amakawa[2], Md. Mamunur Rashid[1]
Dept. of Electrical and Electronic Engineering[1], Semiconductor Electronics & Integration Science[2]
The International University of Scholars, Bangladesh[1], Hiroshima University, Japan[2]
{mshossain.icd@, mrashiddm@}gmail.com[1], {fuji@, tyoshida@dsl. ,amakawa@}hiroshima-u.ac.jp[2]

***Abstract*— Millimeter-wave passive elements are one of the vital part to distribute power through integrated circuitry. And transformer coupled passive devices are used not only for signal transmitting but also making chip area smaller. A transformer coupled planar and stacked type both could be used for balun and power divider-combiner design. This type of balun provide small amplitude imbalance ~2dB and 180^0 phase imbalance between 57-120 GHz bandwidth (BW) and successively shrinking the layout area of $0.0061mm^2$. On the other hand, transformer coupled two way power divider-combiner (PD-C) occupies $0.0088\ mm^2$ (7.5 Times Smaller than Wilkinson PD-C) with amplitude variation ~1.87dB and phase angle variation 3^0 with larger BW (50-120 GHz). Again four way PD-C occupies $0.0122mm^2$ (5 Times Smaller than Wilkinson PD-C) having amplitude variation ~2dB and phase angle variation 2.8^0 with a larger BW 50-120 GHz. Proposed stacked small transformers with a center trapped baluns and power divider-combiner have been implied on two millimeter-wave amplifiers on 55nm DDC CMOS.**

Keywords—Millimeter-Wave Transformer; Transmission Line; Transformer Coupled Balun; Transformer Coupled Power Divider-Combiner; Two-way and Four-Way Power Divider-Combiner.

I. INTRODUCTION

Millimeter-wave passive elements are one of the vital part in integrated circuit design. Transmission line (TX-Line) and capacitors have been used for coalesce with other active section. Balun and power divider-combiner are the key elements for power distributing and combining in-phase or out-phase. Along with several type of baluns [1] [2] [3] [4] and power divider-combiner [5] [6], proposed transformer coupled balun and power divider-combiner (PD-C) have been designed with twelve metallic stack of 55nm DDC CMOS. Initially transmission line has been designed with ground floor, ground wall and signal line (fig.1). Withdrawing the signal line part of TX line, the transformer coils have been added with signal I/O. A center-trapped transformer balun coil has been attached with differential ended active element. In fig. 2, a transformer coupled balun and PD-C have been inserted for signal distribution and matching network setting.

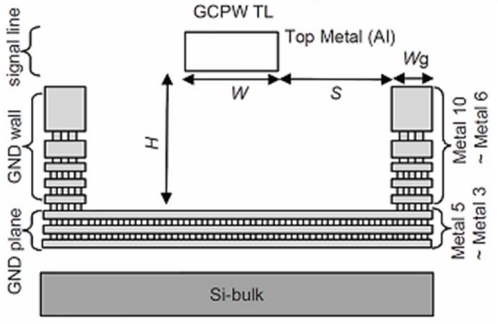


Fig. 1. Cross view of TX Lines

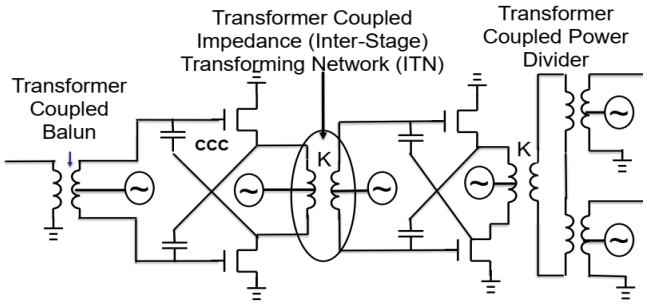


Fig. 2. Transformer coupled ITN with differential amplifier

II. REVIEW OF SOME TEST DESIGN

Millimeter-wave amplifier contain several passive elements for satisfying the property of lower insertion loss, good input and output matching, higher gain and output power, low noise figure, good figure of merit etc. For L-C-L coupled circuit, MIM capacitor, Transmission Line, MOM capacitor, Inductor and resistor are the passive elements for design circuitry. Balun, power divider and combiner also contain transmission lines to mitigate its properties. An ideal balun is a three-port device with a matched input and differential outputs. It is most succinctly described all by the required (ideal) S-parameters (eqn. 1).

$$\angle S12 = -\angle S13 = \angle S21 = -\angle S31$$

$$\mathrm{dB}(S12) = \mathrm{dB}(S13) = \mathrm{dB}(S21) = \mathrm{dB}(S31)$$

$$S_{11} = -\infty$$

$$\text{Amplitude Balance} = 20\log_{10}\left|\frac{S_{31}}{S_{21}}\right| \quad \text{(eqn. 1)}$$

$$\text{Phase Balance} = \text{ang}\left(\frac{S_{31}}{S_{21}}\right)$$

$$\mathrm{IL}_{\mathrm{power}} = 20\log_{10}(|S_{12}|^2 + |S_{13}|^2)$$

A. Ratrace and Marchand Balun

Transmission line with HFSS designed simulation (1-500 GHz) has been converted into RLGC .dat file, fixed for a

center frequency (60GHz & 79GHz) and then inserting the values of R, L, G, C into these equation. The propagation constant, $\gamma = \sqrt{((R + j\omega L)(G + j\omega C))}; \beta = \text{imag}(\gamma); \lambda = 2\pi/\beta$. Rat-race balun also behaves as differential signals $(0^0, 180^0)$ having $\frac{\lambda}{4}$ and $\frac{3\lambda}{4}$ signal line length and layout area 0.138mm^2. The schematic and layout area is comparatively large than capacitively coupled transmission line balun, Marchand balun & transformer coupled balun (fig.2). Marchand balun is TX line type balun, which has left ported input and right ported output with equal length of transmission lines. Marchand balun behaves as differential signals $(0^0, 180^0)$ having separately $\frac{\lambda}{4}$ signal line length and layout area 0.169mm^2. Transformer type balun is considerably good from rat-race and marchand balun in case of keeping lower Amplitude and phase imbalance and layout area is also smaller than marchand and ratrace balun. And it sustains larger BW of 180^0 phase differences.

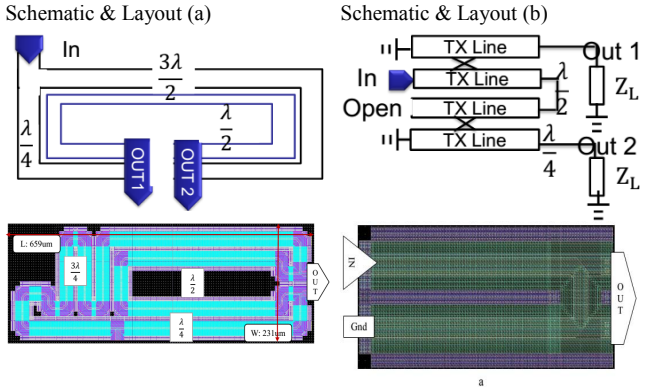


Fig. 2. Schematic and Layout of Ratrace Balun

B. *Wilkinson Power Divider-Combiner*

Power Divider-Combiner (PD-C) is a specific class of circuit that can achieve isolation between the output ports while maintaining a matched condition on all ports. Network theorem of this divider can't satisfy all three conditions at the same time. It satisfies the first two (matched and reciprocal properties) but can't satisfy last one ideally (Loss-less) in eqn. 2.

$$[S] = -\frac{j}{2} * \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{(eqn. 2)}$$

$S_{11}, S_{22}, S_{33} = 0$, These terminals are matched.

$S_{23}, S_{32} = 0$, Output terminals are isolated.

$S_{21} = S_{31}$, Equal power need to be achieved.

$$S_{21} = S_{31} = -3\text{dB} = 10\log_{10}\left(\frac{1}{2}\right), \text{ Ideal wilkinson PD} - \text{C yields}$$

Wilkinson PD-C is widely used for power splitting or combining due to its input/output matching and high isolation characteristics. However, it often requires multiple $\lambda/4$ transmission lines as shown in fig 3, leading to a large circuit area especially at the GHz RF frequency. Recently, several papers have reported Wilkinson PD-C designs with size reduction, by using slow-wave structures [7], lumped/quasi-lumped elements [8], and general π-network structures [9]. However, these circuits still require a large area at the GHz frequency and cannot be easily extended for general N-way Wilkinson power splitting or combining.

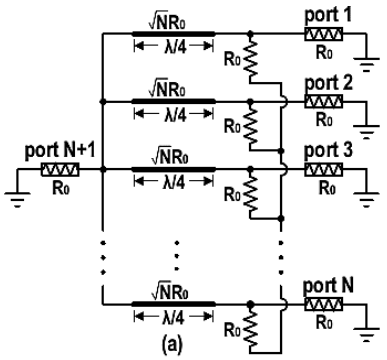


Fig. 3: N-way Wilkinson power divider-combiner (PD-C)

However, S parameter of two way PD-C has been extracted from Momentum cell simulation and converted into ABCD matrices. RLGC parameters has been extracted from these equations to set 100 Ω TX line for 4-Way power divider & combiner. As well as 50Ω, 71Ω resistances TX line value can be evaluated from HFSS model. For two-way PD-C at $Z_0 = 71\Omega$ in the output section. And theoretically the output port signal phases $\angle S_{21} = \angle S_{31}$ and $\text{dB}(S_{21}) = \text{dB}(S_{31})$ are come to represent. And for four-way power divider and combiner $Z_0 = 100\Omega$ is to be settled at the output section. But the signal line width of 100 Ω TX line is so narrow to design whether the capacitance decreases literally and change the performance of propagation constant. Besides it increase the layout area larger with insertion loss (fig. 4).

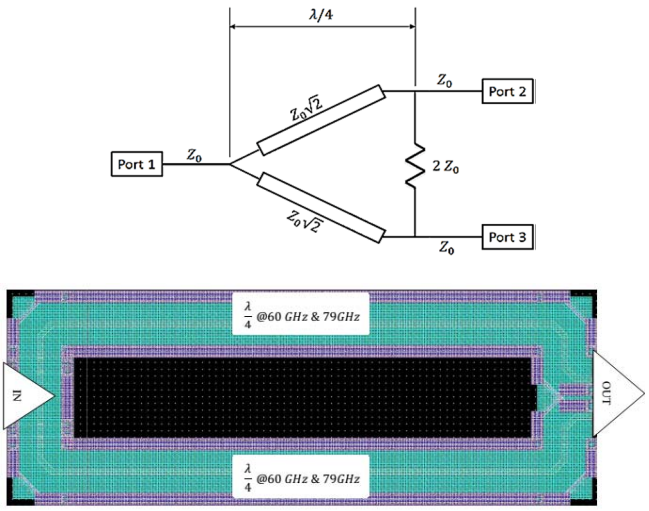


Fig. 4: Schematic & Layout of two-way Wilkinson PD-C

III. PROPOSED MILLIMETER-WAVE TRANSFORMER COUPLED BALUN AND POWER DIVIDER-COMBINER

A. *Ideal Transformer Coupled Balun*

In classical transformers, there are two electrically separate windings of wire coils around the transformer's core and it's behave as a voltage Balun. The primary winding receives the input signal, and the secondary winding puts out

the converted signal. The core that they are wound on air core or a magnetically neutral material like a porcelain support, also it contains a material which is good magnetic conductor like ferrite in modern high-frequency (HF) balun. The electrical signal in the primary coil is converted into a magnetic field in the transformer's core. When the electrical current through the primary reverses, it causes the established magnetic field to collapse. The collapsing magnetic field then induces an electric field in the secondary winding. Transformer coupled Balun behaves as 3 ports (fig.5), the input signal divided into two phases such as in phases (0^0) and in opposite phase (180^0) differential ports signal and then it applies through a differential circuit. The secondary coil has been taken center trapped for gate signaling. And the layout area is so smaller than others balun structural design.

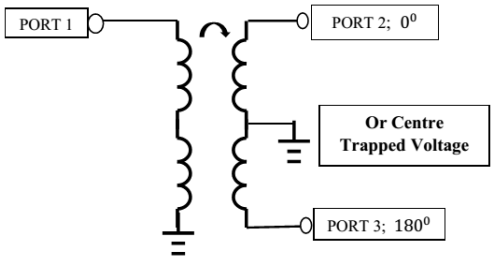


Fig. 5: Ideal Transformer Coupled Balun

B. *Practical Transformer Coupled (Flipped HFSS) Balun*

In transformer coupled practical cell, both flipped or planar topology has been used for 180^0 phase difference and same amplitude signal distribution. Here, flipped transformer coupled balun has been designed with top two metal.

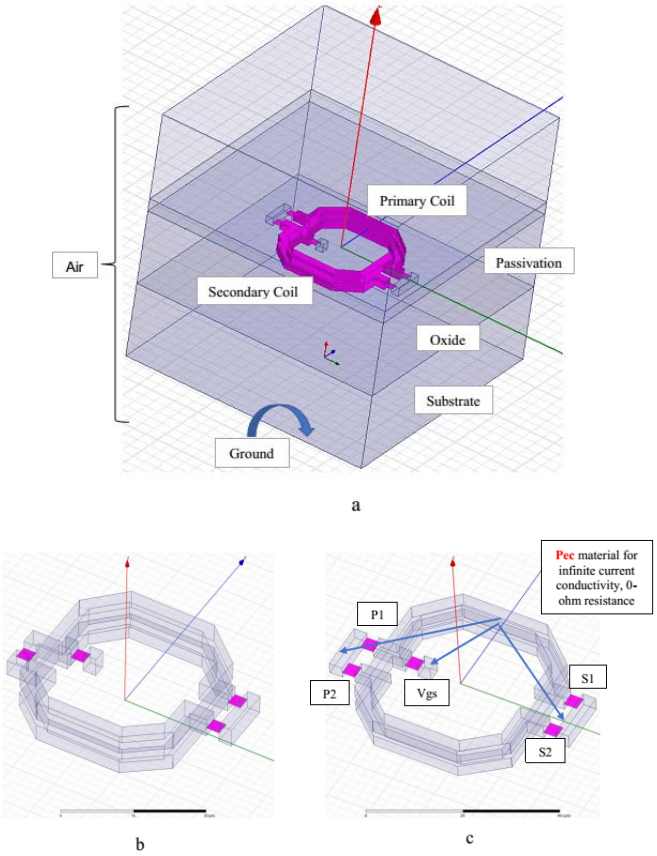


Fig. 6: (a) mm-Wave transformer (b) 4-Port Transformer Balun (c) 5-Port Transformer Balun

Primary coil has been set as Al Metal (2.8 um, TNA) and Secondary coil has been set as Cu metal (3.3 um, LDA). Figure (6a) behaves as mm-wave transformer balun with ground, substrate, oxide, passivation and in air medium. It has two coils: Primary & Secondary with 5-ports (4-port containing source and 1 port is for gate signal). Figure (6c) represents the 5-port transformer coupled balun and fig (6b) represents 4-port transformer coupled balun having short connected primary port.

C. *Practical Transformer Coupled (Flipped Cadence) Balun*

Here, flipped transformer coupled Balun (fig. 7) has been designed with top two metal. Primary coil has been set as Metal 11 (Cu, LDA) and Secondary coil has been set as Metal 12 (Al, TNA). XPWA layer has been used for improving quality factor. From TX Line in fig.2, Ground floor contain metal 3 to metal 5 (LAC, LAD, LAE) with via 4 & 5 (NAD, NAE) on 55nm CMOS. And ground wall contain metal 6 to metal 11 (LAF, LBA, LBB, LBC, LCA, LDA) with via 6 to via12 (NAF, NBA, NBB, NBC, NCA, NDA, TWA). From transformer coupled balun, it's five port S parameters having port 1 (P_1), port 2 (P_2), port 3 (S_1), port 4 (S_2) and port 5 (V_{gate}). For checking balun properties, port 2 (P_2) and port 5 (V_{gate}) have been grounded. And it gives in phase 0^0 and out of phase 180^0 signal at the secondary output section. In figure 13 & 14 below transformer balun property is shown. At 60GHz and 79GHz the phase angle variation is $\angle S21 - \angle S31 = 180^0$, V to W band: $50 - 110\text{GHz} = 180^0$. Here in table (8), transformer coupled balun properties have been compared with another topology.

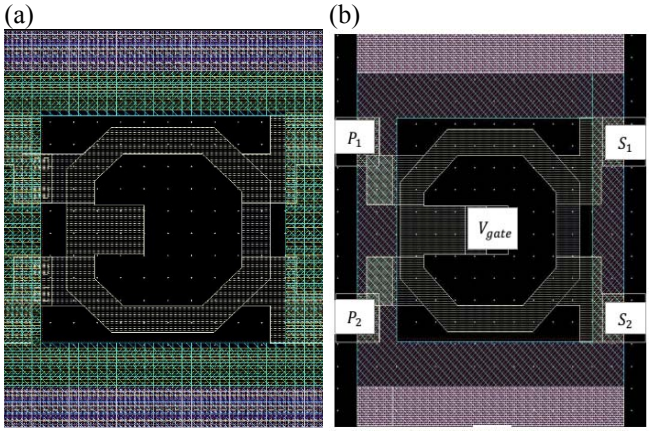


Fig. 7: Layout & Momentum Cell of Flipped Transformer Coupled Balun

D. *Ideal Transformer Coupled Power Divider-Combiner*

Transformer coupled PD-C (fig.8) is same as wilkinson type devices. It represents transformer-based coupling 2, 4, 8…n way dividing and combining with symmetric length and width of primary & secondary coil. Based on different parallel combining of transistor, the length of signaling can be distributed equally. From transformer coupled PD-C, equal power and in same phase signal will transmit through output terminals at OUT1+ & OUT2+.

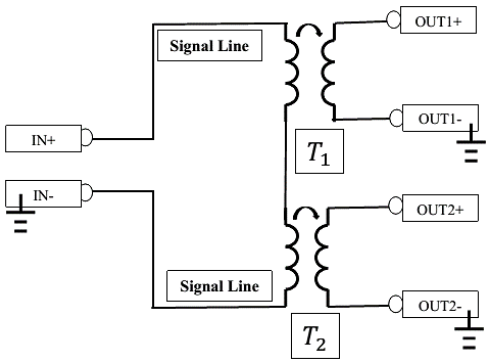


Fig. 8: Ideal Transformer coupled power divider and combiner

E. *Practical Two-Way Transformer Coupled Power Divider-Combiner*

In transformer coupled two-way PD-C with current flow animation (fig. 9), the primary input and secondary output have two-way current distribution section (Out1 & Out2). Primary section top metal (Al, TNA), thickness: 2.8um and Secondary Section metal 11(Cu, LDA), thickness: 3.3um. The layout area of transformer Coupled PD-C ($72\text{um} \times 123\text{um} = 0.0088\ \text{mm}^2$) which is 7.5times smaller than Wilkinson PD-C ($157\text{um} \times 435\text{um} = 0.0683\text{mm}^2$).

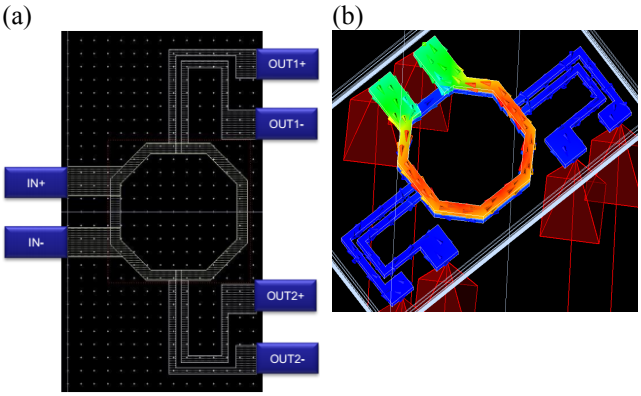
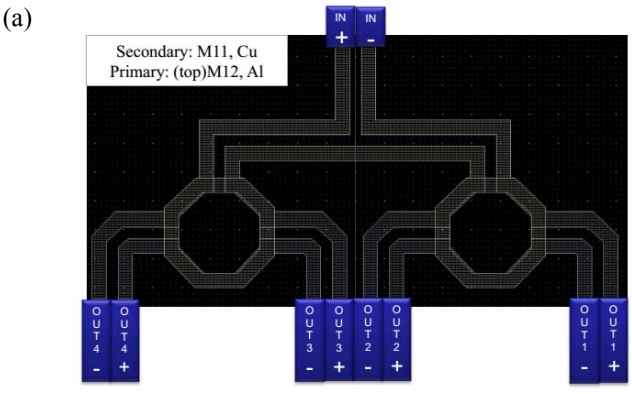


Fig. 9: Two-way transformer coupled power divider-combiner, Layout (a) Current Flow Animation (b)

F. *Practical Four-way Transformer Coupled Power Divider-Combiner*

Four-way transformer coupled PD-C with current flow animation (fig. 10), the primary input and secondary output have four-way current distribution section (Out1, Out2, Out3 & Out4). Primary section metal (Al, TNA), thickness: 2.8um. Secondary section metal (Cu, LDA), thickness: 3.3um.



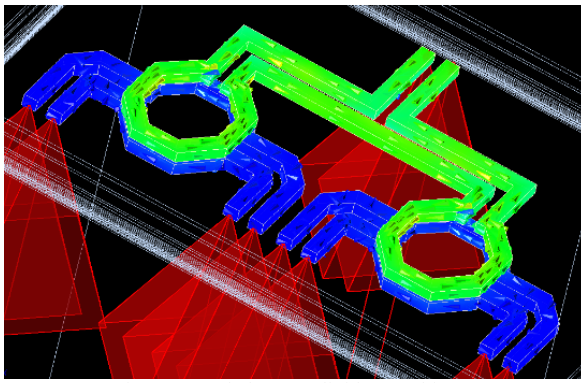


Fig. 10. Four-way transformer coupled power divider-combiner, Layout (a) Current Flow Animation (b)

The layout area of transformer four-way Coupled Power Divider-combiner ($72\text{um} \times 169\text{um} = 0.0122\ \text{mm}^2$) which is 5.5times smaller than Wilkinson PD-C ($157\text{um} \times 435\text{um} = 0.0683\text{mm}^2$). Depend on the signal line position of Balun at the end port, this transformer coupled PD-C output section can be set at longer different symmetric length. Two Way & Four Way PD-C are almost same amplitude & phase angle between $50\text{GHz} - 120\ \text{GHz}$. Finally, it can be said that these transformers coupled PD-C are basically designed for two major purpose, (1) layout area minimizing & (2) signal line matching without using matching network between this bands. S parameter characteristics of four-way transformer coupled PD-C has been shown below here. Here, four-way transformer coupled PD-C (fig. 10) works between 1-150GHz and at 79GHz, phase variation is $\angle S21 \sim \angle S31 \sim \angle S41 \sim \angle S51 = 2.8^0$ and amplitude variation is $\text{dB(S21)} = \text{dB(S31)} = \text{dB(S21)} = \text{dB(S31)} \sim 2\text{dB}$.
Transformer type PD-C is considerably good from wilkinson in case of keeping lower amplitude and phase imbalance and layout area is so small. It sustains larger BW of 0^0 phase differences.

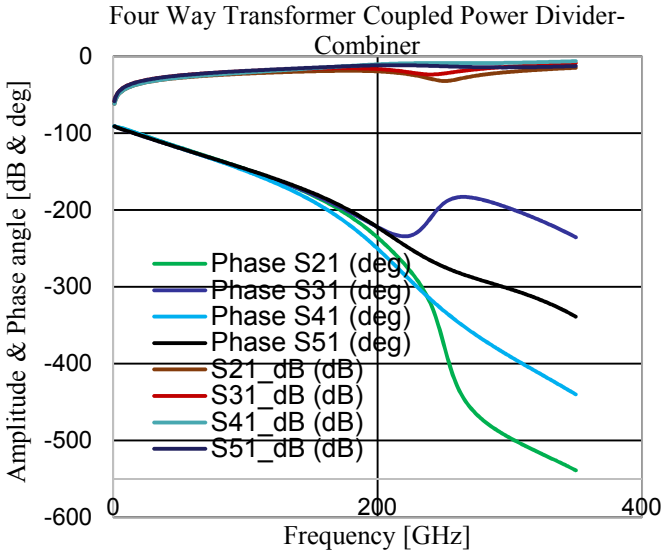


Fig. 11. Four-way transformer coupled power divider-combiner, Amplitude & Phase Differences (a) Return loss (b)

IV. RESULT AND DISCUSSION

From fig.12 & 13, HFSS simulation result, the S parameter property of transformer balun $\angle S21 - \angle S31 = 180^0$ and amplitude are kept same $[\text{dB(S21)} = \text{dB}(\angle S31)]$ between

30GHz to 250GHz. And it has amplitude imbalance (AI): ~1 dB between 30-250 GHz and phase imbalance (PI) < 2^0. Similarly, from cadence simulation it contains 180^0 phase difference and BW limit to 1-120 GHz, amplitude imbalance (AI): ~2 dB between 1-120 GHz and phase imbalance (PI) < 3^0.For transformer coupled balun, HFSS data comparatively give better performance than momentum (cadence platform) design. From fig. 14, the amplitude differences have been compared with transformer, ratrace (lower AI BW zone: 77-89 GHz, lower PI BW: 60-90GHz, return loss<-20 dB) and marchand (lower AI & PI BW: 72-80 GHz, return loss<-15 dB) balun. Here, ratrace balun gives low insertion loss through all BW than other types of balun while layout area is so large. On the other hand, marchand and transformer balun take less chip area but comparatively transformer coupled balun occupies more less area. It provides a bit more insertion loss at low frequency (1-50GHz) however it takes less area than marchand.

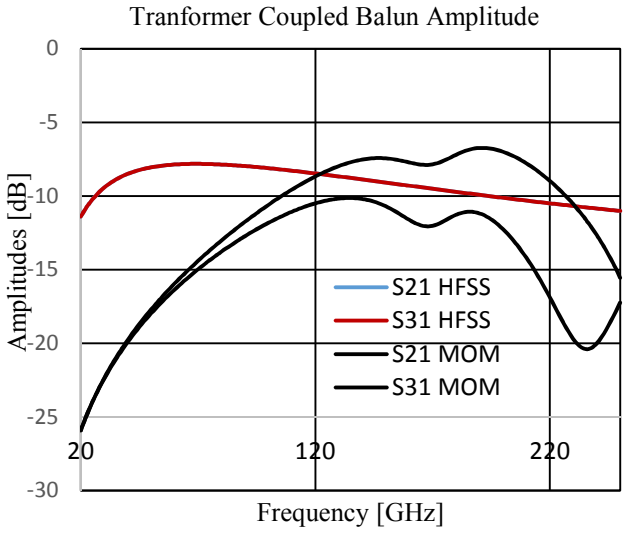


Fig. 12. Amplitude variation of HFSS and Momentum results

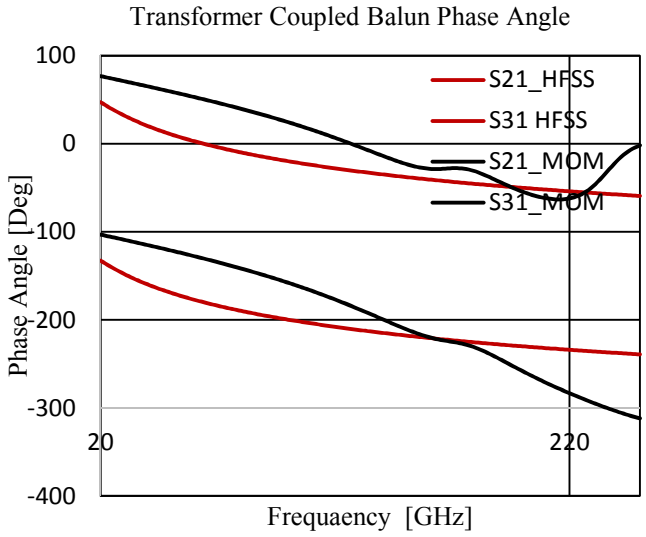


Fig. 13. Phase variation of HFSS and Momentum results

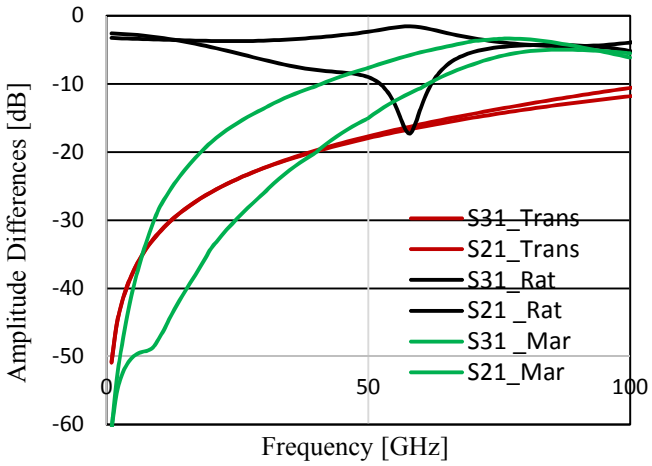


Fig. 14. Amplitude difference of different kind of baluns

In fig. 15, transformer balun provide better performance in phase difference within a large BW than other type of baluns. Transformer coupled balun amplitude imbalance ~2dB and Phase differences perfectly in 180^0 with larger bandwidth (50-120 GHz) occupying area 0.0061 mm^2 which is 23 times smaller than ratrace & 3 times smaller than marchand balun. In table 1, several balun properties have been illustrated and compared.

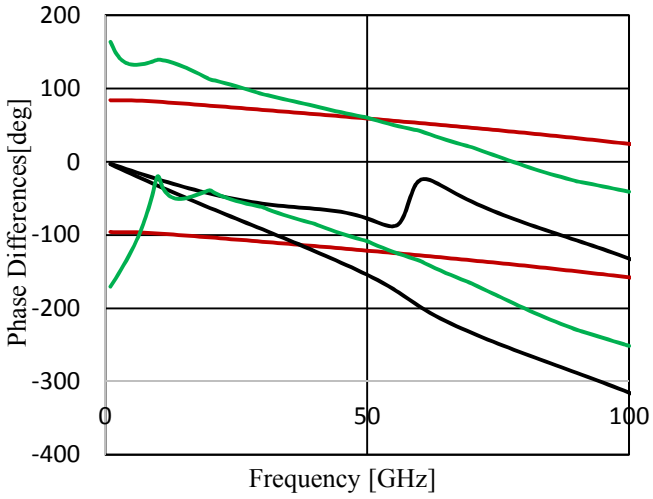


Fig. 15. Phase difference of different kind of baluns

TABLE I. Balun Properties Comparison [10]

Ref.	Topology	Tech.	Freq. (GHz)	AI (dB)	PI (Deg)	IL (dB)	Surface (mm^2)
Lee, 04	Marchand	GaAs	26-55	<1	$<5^0$	<2	0.51
Din, 08	Marchand	0.18um Bi-CMOS	45-75	<0.5	$<6^0$	<2.5	0.09
Din, 05	Ratrace	0.13um Bi-CMOS	57-71	<0.6	$<10^0$	<3.2	0.28
Ham, 05	Line-based	InGaP/GaAs	15-45	<1	$<5.5^0$	<7	0.04
Fel, 07	Transformer	0.13um CMOS	55-65	<3	$<10^0$	3	0.05
Leite, 12	Transformer	0.13um Bi-CMOS	60-120	<1	$<3^0$	<3	0.03
This Work	**Transformer**	**55nm CMOS**	**57-120 30-250**	**~2 ~1**	**$<3^0$ $<2^0$**	**7~11 5~7**	**0.0061**

From fig. 16, the amplitude differences have been compared with transformer and wilkinson PD-C and in comparison with phase differences wilkinson gives better performance than other types but the layout area is so large (fig.17). On the other hand, transformer PD-C take less area but transformer coupled two–way PD-C occupies less area 0.0088mm^2 which is 7.5 times smaller than Wilkinson PD-C and amplitude variation is $\text{dB(S21)}\sim\text{dB(S31)} = 1.87\text{dB}$. Transformer coupled PD-C provides a little bit more insertion loss at low frequency (1-60GHz) although it takes more less area. From fig. 17, two-way transformer coupled power divider-combiner provides 0^0 between 50-120GHz, phase variation is $\angle\text{S21}\sim\angle\text{S31} = 3^0$.

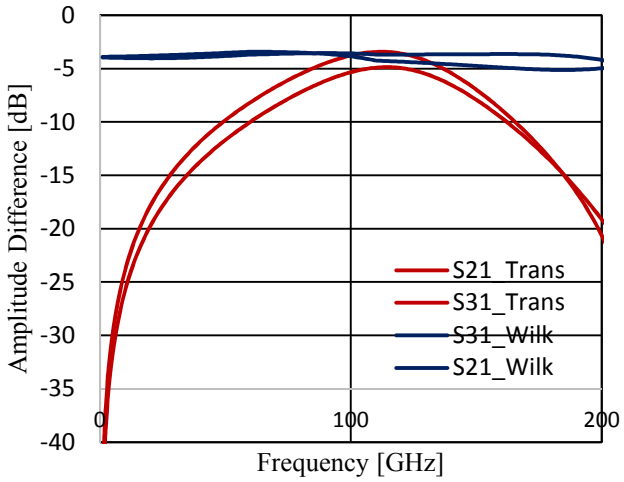


Fig. 16. Amplitude difference of different kind of power divider-combiner

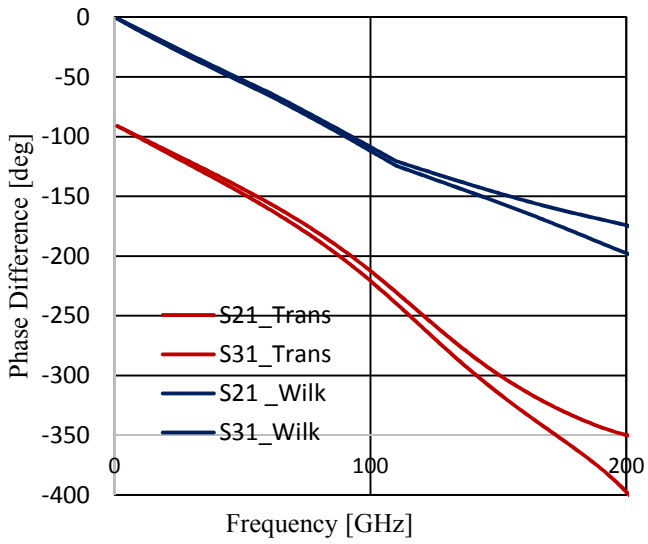


Fig. 17. Phase difference of different kind of power divider-combiner

V. CONCLUSION

Transformer coupled passive elements such as balun and power divider-combiner are comparably good choice for mm-wave circuit design. It contains smaller chip area and easy to integrate with active parts. Considering the total chip layout area shrinking, Transformer coupled devices provide a good performances than others. A small stacked transformers having center trapped balun, power divider-combiner have been designed with 55nm DDC CMOS process. The originality of these paper is maintaining an mm-wave smaller layout design with optimal phase and amplitude imbalance avoiding some pitfalls in low frequency regime.

ACKNOLEDGEMENT

The author would like to thank the member of Frontier Integrated System Laboratory, AdSM, Hiroshima University, along with MIE Fujitsu Semiconductor Inc. Japan, for cadence 55nm DDC CMOS IC design platform and their design rules.

REFERENCE

[1] R. Michaelsen, et al., "A modified marchand balun configuration with tunable phase balance,"*IEEE Microw. Wireless Compon. Lett.*, vol. 23, no. 2, pp. 66–68, 2013

[2] L. Boglione and J. Goodman, "Analysis of tunable marchand baluns," in*Proc. IEEE Radio Frequency Integrated Circuits Symp.(RFIC)*,2014,pp.115–118

[3] P.-H. Tsai, et al., "Broadband balanced frequency doublers with fundamental rejection enhancement using a novel compensated marchand balun," *IEEE Transection Microwave Theory Technology*, vol. 61, no. 5, pp. 1913–1923, 2013.

[4] L. Xu, et al., "A miniaturized marchand balun in cmos with improved balance for millimeter wave applications," *IEEE Microw. Wireless Compon. Lett.*,vol.24,no.1,pp. 53–55, 2014

[5] Liang-Hung Lu et. al," A Miniaturized Wilkinson Power Divider With CMOS Active Inductors" IEEE Microwave and wireless components letters, vol. 15, no. 11, nov. 2005

[6] Mehdi Askari et. al, "Optimization of on-chip spiral inductors using coupled NN and TLBO algorithms for low-loss lumped-element CMOS power dividers" Journal of Intelligent & Fuzzy Systems 30 (2016) 2029–2036

[7] J. Wang, et al, "Miniaturized Microstrip Wilkinson Power Divider with Harmonic Suppression," in IEEE Microwave and Wireless Components Letters, vol. 19, no. 7, pp. 440-442, July 2009

[8] Hai Hoang Ta and Anh-Vu Pham, "Compact Wilkinson power divider on multilayer organic substrate," Microwave Symposium Digest (MTT), 2012 IEEE MTT-S International, Montreal, QC, Canada, 2012, pp. 1-3.

[9] Xiaolong Wang, et al., "Wilkinson Power Divider with Complex Isolation Component and Its Miniaturization," in IEEE Transactions on Microwave Theory and Techniques, vol. 62, no. 3, pp. 422-430, March 2014.

[10] Doug Jorgesen and Christopher Marki, "A Tutorial on Baluns, Balun Transformers, Magic-Ts, and 180° Hybrids", Marki Microwave Inc., Morgan Hill, CA, 2014.

Designing an Electro-Mechanical Ventilator Based on Double CAM Integration Mechanism

Md. Rakibul Islam[1], Mohiuddin Ahmad[2], Md. Shahin Hossain[3], Muhammad Muinul Islam[4], and Sk. Farid Uddin Ahmed[5]
[1,3,4]Department of Biomedical Engineering, Khulna University of Engineering & Technology (KUET), Khulna-9203, Bangladesh
[2]Department of Electrical and Electronic Engineering, KUET, Khulna-9203, Bangladesh
[3,5]Khulna Medical College, Khulna-9100, Bangladesh
Email: rakibulbme14@gmail.com, ahmad@eee.kuet.ac.bd, shahin_srmna@yahoo.com, mmi@bme.kuet.ac.bd

***Abstract*—This paper proposes a simplified structure of microcontroller based mechanical ventilator integrated with a Bag-Valve-Musk (BVM) ventilation mechanism. Here, an Ambu bag is operated with computer-aided manufacturing (CAM) arm that is commanded via a microcontroller and manual switches by sending a control signal to the mechanical system and according to this control signal, the mechanical computer-aided manufacturing (CAM) arm simultaneously compresses and decompresses the Ambu bag. It is a self-inflating bag and like a one-way valve around its inlet and outlet corner. By compressing the Ambu bag it delivers air and by relaxing, it takes air from the environment through a mechanical scavenger. The control signals are designed with three modes named adult mode, pediatric mode, and child mode based on the respiratory rate. The device is in assist controlled mode by dint of fixing the tidal volume for all unique control signals. The control signal is visualized by a platform known as BIOPAC student's lab system. The proposed device is portable, compact, low weight, and efficient performable. It can be supplied around the rural area hospitals for immediate medication with cost efficiency and risk avoidance. Anyone can operate it as no need to study or training of ventilation rules like ICU ventilator. The proposed system is safe, riskless, and repairable. The angle, volume, and respiratory measurement have found 95%, 92%, and 90% accuracy respectively. By applying this portable ventilator system immediate attention can be taken up in rural or general hospitals and in ambulances.**

Keywords—Ambu bag, Automation, computer-aided manufacturing arm, positive pressure ventilation, Portable, Emergency ventilation system

I. INTRODUCTION

In this era of modern science, the uses of microcontrollers are increasing day by day. Integrating with it the automation systems are improving. A manual resuscitator generally known as Bag-valve-musk (BVM) ventilator is used in the mechanical ventilating system. Manual compression, as well as relaxation of an Ambu bag, is inaccurate and painful for the operator with an unknown volume of air. Reducing this specific problem an automation system is necessary for well delivery of air for the person who is unable of taking a breath. Ambu stands for Artificial Manual Breathing Unit. It was invented in 1956 by German engineer doctor Holger Hesse and his partners. In 1559 an Italian professor of anatomy named Matteo Realdo Colombo recounted a method called tracheotomy's method [1].

For attainment a better output, an approximate mechanism is maintained for controlling Ambu bag. In this paper, a specific mechanism is used to get the required output from the system. A normal adult (18 years aged or more) person generally takes 12-20 breath within a minute where a child (1- 12 month aged) has 30-60.

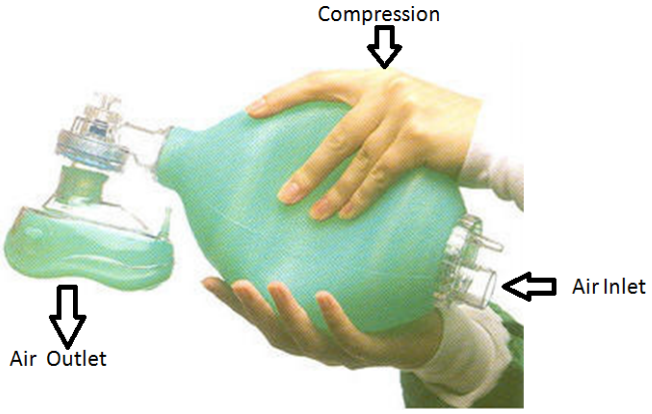


Fig. 1. Manual operation of the Ambu bag system.

It is clear that the respiratory rate of a child is faster than an adult. Out of them, pediatric (6-11 years aged) persons have the respiratory rate (RR) of 18-25 per minute. Authors in [2] describe the detection and monitor of a metabolite of pneumonia at the early stage. A silicon Ambu bag is chosen for its auto relaxation property. This property takes zero energy for relaxation and thus automation system makes flexible. At an emergency condition in ambulance sometimes need 90 percent or more oxygen for the patient recovery. So portability and low power consumption is a great fact for patient care.

With the rapid growth of medical technology, the configuration architecture of the ventilators is becoming complex day by day due to the integration of mechanical, electrical, electronic, and computational technologies [3]. Authors in [4] describe a roller-chain method where this chain compression makes a force on an Ambu bag. This method is acceptable to some drawbacks. There is some possibility of breaking chain thus be dangerous in an emergency condition and need huge torque to compress it. In paper [4], another mechanism named Cam-arm mechanism (CAM) is shown. The cam-arm mechanism is better than the roller chain mechanism on account of its noise reduction property. Our prototype is designed on the basis of two handmade CAM-arm functions that use optimal torque with noise reduction, low power consumption, portability augmentation, and better output. Lung-protective ventilation is treated as a simple supportive therapy and also as a prophylactic therapy to attenuate patient self-inflicted lung injury [5].

The paper is organized as follows: Section II describes the materials and methodology. This section describes the framework of the proposed method; the existing system, proposed system, CAM arm mechanism, BVM mechanism, corresponding mathematical formulation, and circuit diagram. Section III gives hardware implementation and its relative advantages and disadvantages. Section IV gives the results and discussion of the proposed system. Finally, section V concludes our paper with some future directions.

II. Materials and Methodology

A. Flowchart of the Proposed Research

The basic framework of the proposed research work is shown in Fig. 2.

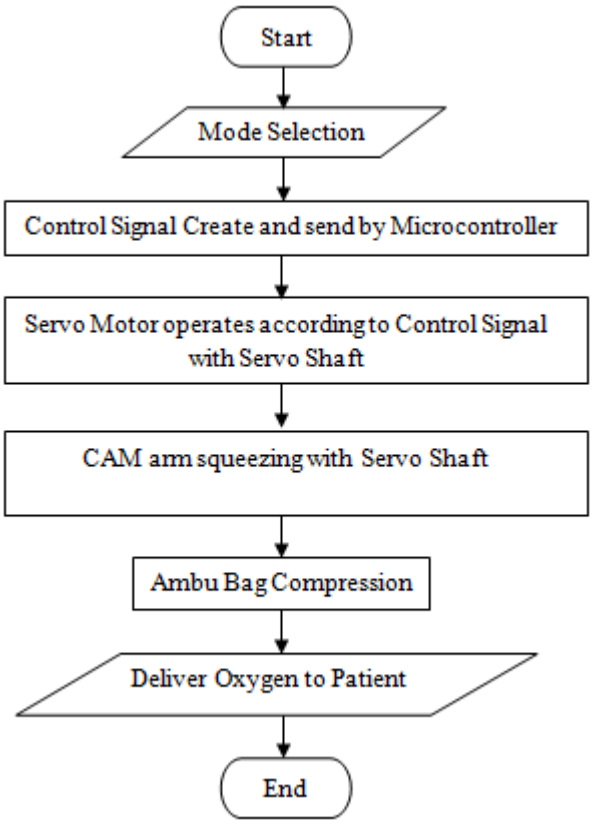


Fig. 2. Basic flowchart of the proposed work. The flowchart represents three modes at a time (Adult mode, pediatric mode, and child mode).

B. Existing System

There are several existing systems of operating an Ambu bag. Such as Ruler chain mechanism, CAM mechanism, Rack and pinion mechanism, Lead screw Mechanism. They have some advantages and disadvantages. Another existing ventilator system is integrated with the vocal alarming system to reduce noise that generates from ventilators in the hospital [6]. Ruler chain mechanism is preferred in previous for its availability and high torque gain with DC motor. When it operates, noise generates that can cause sound pollution in the care environment. Tele-controlling mechanism exists and integrated with the system [7, 8]. After that, another mechanism preferred called CAM mechanism eradicates the noise or additional sound. Actually, this mechanism does not generate any noise like ruler chain mechanism. Rack and Pinion mechanism makes the platform of gear for more precise controlling. Lead screw mechanism uses a threaded screw to operate a mechanical arm with a nut that compresses and relax the Ambu bag simultaneously. Authors in [9] show the hyperbaric oxygen therapy using mechanical ventilator where the pressure is proportional to volume.

C. Proposed System

Our proposed system is based on CAM principle having some additional feature of controlling precisely with portability, battery operated, reduced size and noise removal. Several modes selected for control operation such as child mode, pediatric mode, and adult mode. These modes are selected according to the lung capacity and breath-rate of the subject as well as inspiration and expiration (I:E) ratio. All of these parameters are generally depend on the age of the subject. The automation requires a microcontroller, some mechanical switches, and two servos. Servos are positioned face to face oppositely with the 180-degree angle. The shafts of the motor are attached with a mechanical pulley that rotates according to motor direction. A nonflexible cot wire connects one arm and one pulley and same to the second one. When the motor rotates, the torque creates tension in the cot wire through the pulley that results in pulling a mechanical arm to compress an Ambu bag.

D. Computer-Aided Manufacturing Arm Mechanism

A CAM (Computer-Aided Manufacturing) arm is a rotating or sliding piece in a mechanical linkage that is used especially to convert rotary motion into linear motion which is shown in Fig. 3. The main reason for selecting a cam arm is the servo system as it has a high torque of approximately 15 kg-centimeters. The arm is nonflexible, hard and weight bearable. The cam function is graphically shown in Fig. 4.

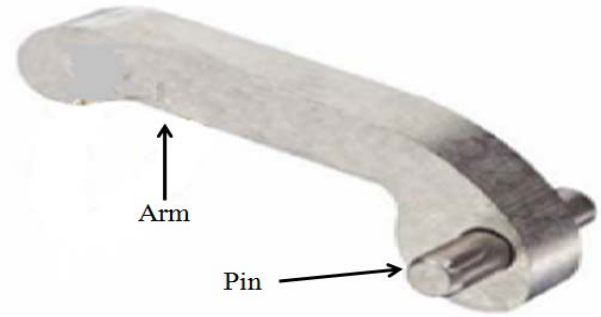


Fig. 3. Basic CAM arm with its pin attachment.

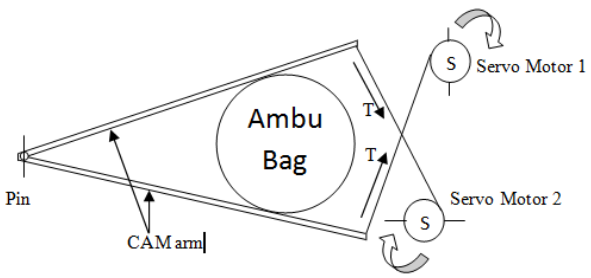


Fig. 4. CAM function with the proposed system (T means tension).

E. Bag-Valve-Mask (BVM) Mechanism

The Bag-valve-mask was also known as manual resuscitator is a mechanical system that intelligently delivers air to a single direction when it is compressed and takes air from the environment when relaxation takes place [1]. In a proper way, BVM ensures air and oxygen from the environment to the patient. There are three inlets in the Ambu bag used in this project. One is for the environmental air inlet. The second one is for the connection pathway of oxygen reservoir and the third one is for the loss minimization. Out of that, there is a single pathway for emergency air circulation. There is a single outlet of the Ambu bag that is used to connect the patient for air inhalation. There is a one-way valve in the outlet of a BVM Ambu bag. Out of that, there is also a single air pathway at the outlet side of it for emergency exhalation. Some emergency patients need a high volume of oxygen approximately tends to hundred percent oxygen otherwise dangerous hypoxia occurs where hypoxia is the oxygen deficiency in the abiotic environment.

F. Mathematical Formulation

Child mode (1-12) month aged child has a tidal volume of (4-6) milliliter per kilogram. Pediatric mode (6-11) years an aged person has a tidal volume of (5-8) milliliter per kilogram. Adult mode (18 or more) years an aged person has a tidal volume of 7 milliliters per kilogram. So, for example generally, the tidal volume of 23 years and 61 kilogram weight adult person is equal to (7 ml/kilogram×61 kilogram) or 427 ml. In this way, we can calculate the general tidal volume of all aged person.

The diameter of the pulley is 4.3 centimeter, so the circumference of the pulley is 27 cm approximately. As we have two servos operating on 180 degrees or two right angles at a time, so the cot wire can oscillates within (27/2)×2 = 13.5×2=27 centimeters. Arm length is 40.50 centimeter. Relax angle or initial angle is 37.94 degrees or 5.1/7.7 radian and compress angle or final angle is 18.2 degrees or 2.45/7.1 radian for an adult. Now we want to find out the angle between two CAM arms because the output volume depends on the angle between two arms.

We know that the arc length of a circle is $S = r\theta$, where the angle θ is in radian and r is the radius. We have three circles, two pulleys which act like a circle and another imaginary large circle is considered where CAM arm is the radius of that large circle. As we have two pulleys, we take $S_p = S_{p1} + S_{p2}$ *where* S_{p1} is the arc from the first servo motor and S_{p2} from the second servo motor. For 180 degree or π radian rotation of pulley the imaginary circle covers a small arc Sa. So it can be said that S_a is equal to S_p, then $r_a\theta_a = r_p\theta_p$. Finally, we get Eq. (1).

$$\theta_a = \frac{r_p \theta_p}{r_a} \tag{1}$$

We know that as large as the radius of the pulley, it covers large arc S_a of the largest circle shown in Fig. 5 (left). Therefore, $r_p \propto \theta_a$ and we get Eq. (2).

$$r_p = k_1 \theta_a \tag{2}$$

Where k_1 is the proportional constant that depends on r_p and θ_a. Another proportional constant k_2 is found where θ_a is proportional to the output volume (V) of the Ambu bag results for compression on it by the mechanical CAM arm as written in Eq. (3).

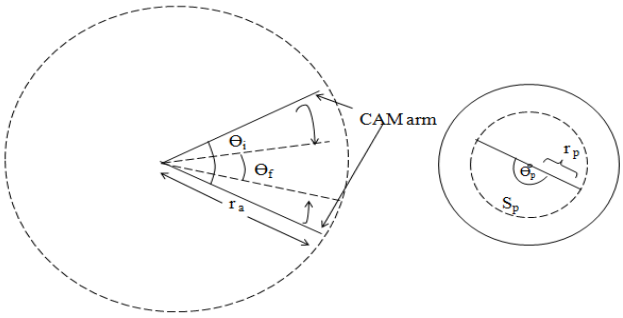


Fig. 5. (Left) Large imaginary circle to measure of $\theta_a = \theta_i - \theta_f$ and (Right) Pulley structure to calculate different angles.

$$\text{So, we can write } \theta_a \alpha V \quad \text{or } \theta_a = k_2 V \tag{3}$$

From Eq. (2) & Eq. (3), the relationship between pulley radius and oxygen output volume is obtained that is shown in Eq. (4).

$$r_p = k_1 k_2 V = kV \tag{4}$$

Since k is constant, therefore, the radius of the pulley is proportional to the volume output of the Ambu bag.

In the above analysis, θ_a is the difference between an initial angle and final angle, θ_i is the initial angle, θ_f is the final angle, V is the volume output that means the tidal volume, r_a is the radius of the large circle shown in Fig. 4(left) that means the length of the CAM arm, r_p is the radius of both pulleys. Here, S_a is the arc covered by the large imaginary circle described before.

G. Proposed Circuit Diagram

The circuit diagram consists of resistors, microcontroller, servo motors, switches, and power supplies. The input and output pins are set to the digital pin of the microcontroller. A 12-volt power supply is added to get high torque from the servo motor. The electronic circuit diagram of the proposed system is shown in Fig. 6.

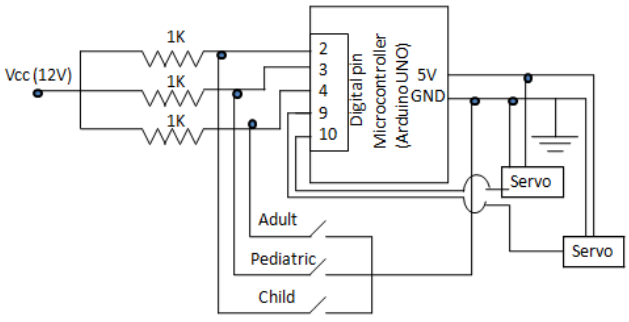


Fig. 6. Electronic circuit diagram of the proposed system

III. System Architecture and Hardware

A. System Architecture

This architecture shows the combination of all hardware part of the system that is stable with a hardware frame. Oxygen cylinder is used with a control opening. An oxygen tube is connected with an oxygen cylinder and Ambu system. Our environment contains about 23 percent oxygen. So we can take it through another inlet to reduce loss. The system architecture is shown in Fig. 7.

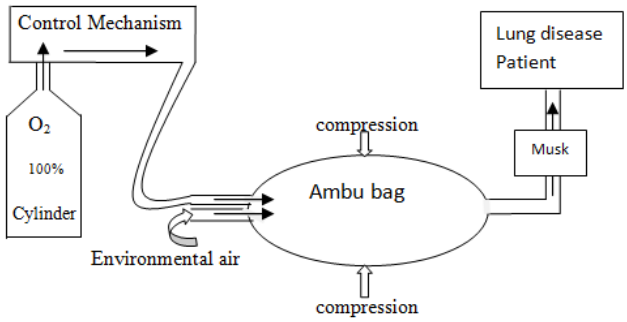


Fig. 7. The architecture of the proposed system making a way of oxygen delivery.

B. Hardware Implementation

The actual hardware implementation of this project is shown in Fig. 8. The hardware implementation is one of the most important as it is a challenge to reduce mechanical error. The hardware design consists of the following: microcontroller, wood frame, CAM arm made by PVC, servo motor, switches, pulleys, DC 12 volt battery, Ambu bag, jumper wires, etc. The top, front and right side view of the hardware implementations are given below in the figures respectively.



Fig. 8. Experimental setup with BSL SS11LA airflow transducer (right) and with a peak flow meter (left) of the proposed system.

C. Facilities and Advantages

There are some attractive facilities of the Ambu bag with this proposed system. Clinicians offer patients to use age wise different Ambu bag. Such as adult Ambu bag, the silicon Ambu bag is the self-inflating system. Here we need only to compress it. After compression when decompress it inflates itself. This is the most attractive feature of silicon Ambu bag. Ambu bag can be sterilized easily although it is non-autoclave-able. There are two emergency airways at the inlet and outlet side. Another attractive feature of Ambu bag with this proposed system is that instead of using three Ambu bag (such as for Adult, Pediatric and Child Ambu bag) one Ambu bag is used to perform different modes by controlling it with a microcontroller. The adult BVM system is about 1650 ml that is enough to deliver the minimal tidal volume to all aged patient using the control mechanisms. In the future, some more additional advantage can be integrated with this proposed system to improve it.

D. Drawbacks, Hazards, and Solutions

Misassemble of Ambu bag sometimes occurs for the cause of using an Ambu bag for so many days as the manufacturing material have fracture properties. To avoid this problem Ambu bag can be replaced. The second one is the Bacteria and Germ attack from the environment while filtering is not used. The third reason is the ventilation difficulty of the BVM system because of not replacing after long-term use. The fourth one is improper BVM selection. All of this problem can be solved by using proper hardware implementation and technique development. After all the entire germ attack is the main problem of an Ambu bag. Germ attack can be reduced by using a filter at the inlet of the bag and the filter may be changed every day after use per patient. However, in a developing country like Bangladesh has no in-house clinical engineer in any of the ICU which represents the inferior type of clinical engineering interaction with life support ventilator [3, 10].

IV. Results and Discussions

We have measured three of our output data with respect to the input. They are data of angle, data of tidal volume (TV) and data of respiratory rate (RR). The command angle is generally measured from the large circle that is shown in Fig. 4 (left) without any Ambu bag and the output angle is measured with the Ambu bag. This is why the Ambu bag is like a load in the system of the device. So we found the efficiency of the whole system while we got a different angle at with load and without load for the same input. The angle errors are shown in Table I. this angle error results for the insufficient torque of the servo motor.

In Table II, we have discussed the tidal volume output. All the separate command angles that described in Table I, we measured the command volume in the same way. Then after Ambu bag insertion, we found some variation in the output tidal volume. As our servo was DC servo with 180-degree angle highest rotation, this is the highest output for our hardware device, rather than we can get a more accurate result by changing the hardware structure. In Table III respiratory rates are discussed. We chose some respiratory rate according to the age of a patient named command respiratory rate or command RR and beside them, the observed respiratory rate is displayed. The negative percentage refers to the high observation than input command. The tidal volume graph of an adult, pediatric and child mode respectively collected from Biopac Student Lab system as a portion of spirogram is described by Fig. 9, Fig. 10 and Fig. 11 [11]. An adult person has a low frequency of RR than a pediatric boy or a child. On the other hand, adults have a high amplitude of air volume that is clearly notified from the output graph. The average peak to peak value describes the tidal volume that is found 395, 243 and 71 for adult, pediatric and child mode respectively. For some experimental result, we have used some standards displayed in the command. We can alter the command standards and add several modes more to find out more accurate results that would reduce errors. These are our final outputs.

TABLE I. MEASUREMENT OF INITIAL ANGLE AND FINAL ANGLE FOR DIFFERENT MODES

Mode	Command angle (degree)	Initial angle (degree)	Final angle (degree)	Output (difference of angle)	Error
Adult	19	37.9	19.7	18.2	4.71 %
Pediatric	16	37.9	23.1	14.8	6.91 %
Child	10.5	37.9	27.8	10.1	5.66 %

TABLE II. MEASUREMENT OF OUTPUT VOLUME AND ERORR CALCULATION FOR DIFFERENT MODES

Mode	Command volume (ml)	Output volume (ml)	Volume error
Adult	430	395	8.13 %
Pediatric	260	243	6.53 %
Child	65	71	-8.85 %

TABLE III. MEASUREMENT OF RESPIRATORY RATE CALCULATION FOR DIFFERENT MODES

Mode	Command RR (breath/minute)	Output RR (breath/ minute)	RR error
Adult	20	21	- 5 %
Pediatric	30	34	- 13.33 %
Child	60	67	- 11.67 %

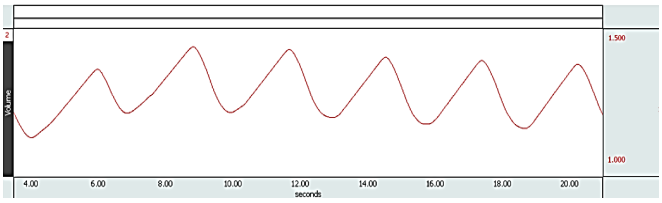


Fig. 9. Spirogram of tidal volume for an Adult male collected from BSL integration with our device. (Tidal volume is 395 ml approximately found)

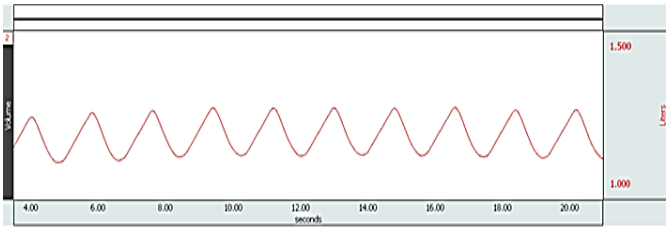


Fig. 10. Spirogram of tidal volume for a pediatric boy collected from BSL integration with our device. (Tidal volume is 243 ml approximately found)

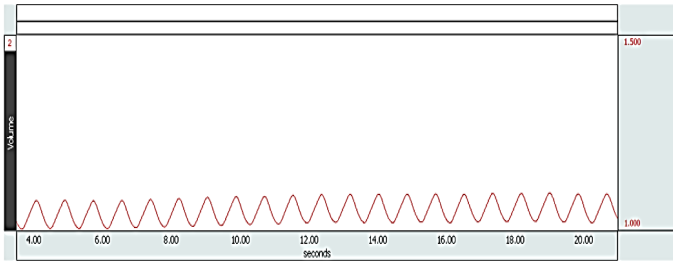


Fig. 11. Spirogram of tidal volume for a child collected from BSL integration with our device. (Tidal volume is 71 ml approximately found)

The command values of angles in Table I are found by the calculation in the programming section of the microcontroller and the output value of angles are the measured value. Based on these angles and structure of Ambu bag, the command volume and respiratory rate are identified and output of them are the measured parameters that are collected from Biopac BSL software [11]. Between the calculated and measured value, the efficiency or error is found out easily while the measured values are the outputs.

V. CONCLUSIONS

Our goal is to implement such a ventilator system that is stable, portable, high efficient, low powered or battery operated, easy to use, riskless, repairable and reusable. Although we could not fulfill all facilities we want, we efficiently show the output at so near to the exact output values. In the future, we can develop another technology to find out the tidal volume and respiratory rate automatically using sensors that will be implemented with the ventilator system and according to the physiological parameters, the control signal that means tidal volume can be generated. After all mechanical ventilator is a growing technology in this world now for the automation basics like telemedicine applications. Smart ventilators may be a large field for both in clinical cure or research field in the near future.

REFERENCES

[1] A. Khoury, S. Hugonnot, J. Cossus, A. De Luca, T. Desmettre, F. S. Sall and G. Capellier, "From Mouth-to-Mouth to Bag-Valve-Mask Ventilation: Evolution and Characteristics of Actual Devices—A Review of the Literature," BioMed Research International, pp. 1-6, 2014.

[2] S.-W. Chiu, J.-H. Wang, K.-H. Chang, T.-H. Chang, C.-M. Wang, C.-L. Chang, … K.-T. Tang, "A Fully Integrated Nose-on-a-Chip for Rapid Diagnosis of Ventilator-Associated Pneumonia," IEEE Transactions on Biomedical Circuits and Systems, vol. 8, no. 6, pp. 765–778, 2014.

[3] M. A. Hossain, A. K. M. Rahman, M. R. Rahman, M. R. Islam, and M. Ahmad, "Role of clinical engineering to reduce patient's risk factors in life support ventilator," Proc. 2nd International Conference on Electrical Information and Communication Technologies (EICT), 2015.

[4] A. M. Al Husseini, H. J. Lee, J. Negrete, S. Powelson, A. T. Servi, A. H. Slocum, and J. Saukkonen, Design and Prototyping of a Low-Cost Portable Mechanical Ventilator, Journal of Medical Devices, vol. 4, no. 2, 2010.

[5] L. Brochard, A. Slutsky, and A. Pesenti, "Mechanical Ventilation to Minimize Progression of Lung Injury in Acute Respiratory Failure," American Journal of Respiratory and Critical Care Medicine, vol. 195, no. 4, pp. 438–442, 2017.

[6] S. Mojdeh, A. Sadri, M. Nabi, H. Emadian and M. Rahimi, "Designing the vocal alarm and improving medical ventilator," Iranian Conference of Biomedical Engineering, 2010.

[7] H. Seddik and M. A. Eldeib, "A wireless real-time remote control and telemonitoring system for mechanical ventilators," Cairo International Biomedical Engineering Conference, 2016.

[8] J. G. Seifert, D. S. Hedin, R. J. Dahlstrom, and G. D. Havey, "Telemedicine enabled remote critical care ventilator," Annual International Conference of the IEEE Engineering in Medicine and Biology, 2010.

[9] R. Felber and M. S. Candidate, "Automation of ventilator control for hyperbaric oxygen therapy," IEEE 30th Annual Northeast Bioengineering Conference, 2004.

[10] L. Vignaux, D. Tassaux and P. Jolliet, "Evaluation of the user-friendliness of seven new generation intensive care ventilators," Intensive Care Medicine, vol. 35, no. 10, pp. 1687–1691, 2009.

[11] Biopac student lab. https://www.biopac.com/curriculum/l12-pulmonary-function-i/

Performance evaluation of an optical drop filter based on simple homogeneous photonics crystal resonator

Md. Aminur Rahman
Dept. of Electrical and Electronic Engineering
Ahsanullah Univeristy of Science and Technology
Dhaka, Bangladesh
aminur.rahman.eee@gmail.com

A.K.M Ehtesanul Islam
Dept. of Electrical and Electronic Engineering
Ahsanullah Univeristy of Science and Technology
Dhaka, Bangladesh
ehtesan@aust.com

***Abstract*— In this paper we proposed a novel drop filter structure based on Photonic Band Gap (PBG) for the 1550nm communication wavelength. The filter is designed and then simulated using Finite Element Method (FEM) and Plane Wave Expansion Method (PWEM) method. The filter shows 91% efficiency at the mentioned wavelength. We also analyzed how the target wavelength can be varied by changing various aspects of the structure. The design utilizes all homogeneous method which is easier to fabricate and introduce in photonic integrated circuit.**

Keywords—Photonic Integrated Circuit, drop filter, photonic band gap (PBG), photonic crystal

I. INTRODUCTION

Photonic crystal (PC) structure gained tremendous popularity in the last decade. Though the one dimensional photonic crystal (1DPC) was proposed in 1887, after the introduction of 2D (2DPC) and 3D Photonic(3DPC) crystal in 1987 the potential of PC structure were fully realized [1 ,2]. Due to the lack of computational power in the early stage not many structure was analyzed. As the computational power increased analysis of different structure is easier than ever before.

Electromagnetic wave can be controlled by manipulating periodic dielectric structure. Most commonly used design to guide EM wave is metallic pipe waveguide and dielectric guide [3]. They are used mostly in the infrared, visible or microwave range. Waveguide which are based on total internal reflection are restricted by heavy losses in the bends. But by creating defect or engineered structure in a Photonic band gap (PBG) material, EM wave can be confined as such it does not show high radiation losses in the bend. A linear defect in a PBG material can be used as a waveguide. This type of waveguide work as the defect results in a guided band in-between TM band gap. So, the conventional optical fiber use index guiding whereas the PBG structure use band gap confinement to confine light. Band gap confinement works better as this minimizes losses, nonlinearities, and any other unwanted properties of the material [4].

In all optical computing, photonic integrated circuit or an optical communication system PCs can be used extensively due to their longer life period, better confinement, high speed and small size [5-10]. In a dense wavelength division multiplexing system (DWDM) due to the high spectral selectivity, low loss and high quality factor of PC structure, the conventional strip based micro-ring resonator can be replaced by PCs [11-14].

PBG structure can be classified broadly as homogeneous and heterogeneous structure. In a heterogeneous structure, different sections use different reflective index or the reflective index of the material can be actively changed (in case of electro-optic/ magneto-optic/ acousto-optic materials) [15,16]. Hetero structure is useful on its own ground but it is difficult to fabricate. On the other hand, homogeneous structure has material with single reflective index which is easy to fabricate.

Many different structure based on PBG have been proposed, analyzed and fabricated. L-shaped bends, T shaped power splitter, [18,19] X shaped, quasi square, tri quarter square etc. [20]. from the above structures it is various category of filter like that band pass, add/drop or device like demultiplexer in WDM system are studied extensively. In most of the filters, they showed various way to manipulate the filter characteristics. Tuning the wavelength changing the add/ drop frequency, increasing the transmission coefficient can be done by changing dielectric constant, radius of the rod, changing lattice constant etc.

In this paper we have investigated a drop filter design created by introducing linear gap in the structure to create the waveguide and then creating a simple structure in the gap to create resonant structure, which performs as a drop filter in the 1550nm range. The next section introduces the structure used in the simulation. In section 3 we describe the simulation result. Section 4 describe the working principle of the structure. Section 5 summarize the finding in the paper.

II. GEOMETRY OF THE FILTER STRUCTURE

The drop filter is designed to be operated in the 1550nm wavelength range. A square dialectic rod lattice with a lattice constant of 'a' and dielectric rod with radius "r=.2*a" is chosen for the structure. To design the filter, first the photonic band gap of the square lattice was determined to find out the range of frequency the filter can be used for. Using the lattice parameter and r=.2*a, relative permeability ε_r = 16, relative permittivity, μ_r= 1 using Plane Wave Expansion Method (PWEM) dispersion diagram was determined (Fig. 1). While calculating different band it was also considered the material in-between the dialectic rod is air.

For transverse electric (TE) polarization the dispersion diagram shows there is a complete photonic gap for the range $0.2451 \leq a/\lambda \leq .3871$. The complete photonic gap (PBG) can be seen in Fig. 1 as yellow region. Inside this region the square lattice will not pass any TM polarized EM wave. (there is no allowed mode in this range of frequency). This is the primary concept in band gap confinement. By creating line defect in perfect lattice structure guided mode can be crated in between the photonic band gap. Two waveguide made out of line defect was created to pass TM polarized light. The EM excitation travel in the air in between the lattice structure. The proposed filter structure (Fig. 2) is created in-between the waveguides.

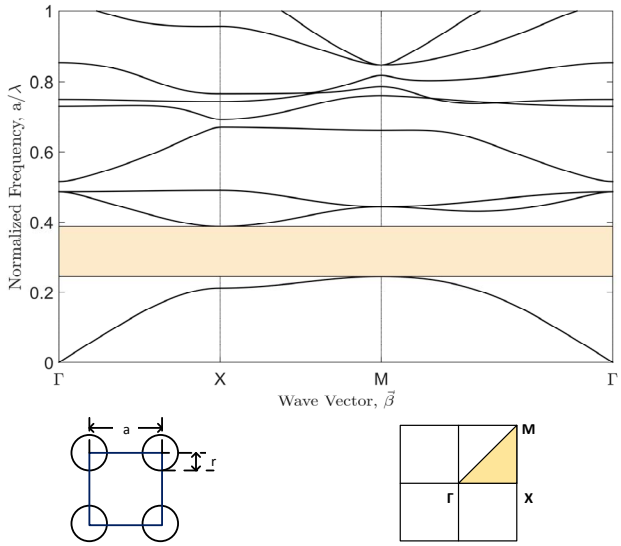


Fig. 1. Top: Dispersion diagram of a square lattice (n=4) and r=.2*a. The yellow region shows a complete photonic band gap for TE mode. Bottom left: The square lattice Bottom right: Irreducible Brillouin zone (IBZ) showing the key points of symmetry

The coupling rods (yellow square in Fig. 2) as the name suggests couples the EM energy from the top waveguide to the resonator cavity. In our design, we employed the simplest of structure by using only a single inner dielectic rod (red square) with a radius of 4*r . Here 'r' is the radius of the dielectric lattice rod

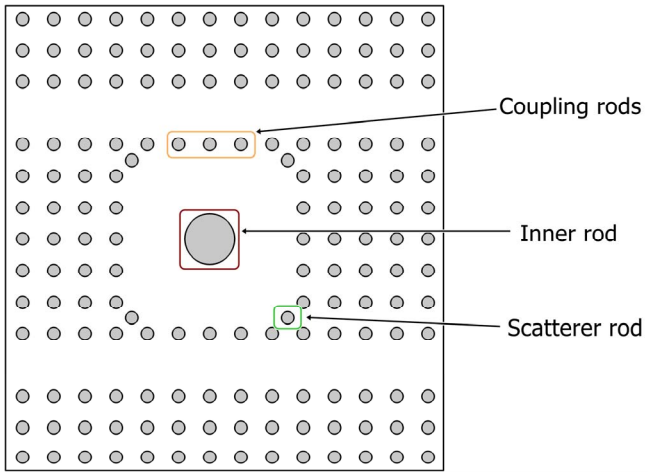


Fig. 2 Proposed resonent filter

PWEM is a popular medthod employed to determine the theoritical dispersion diagram of a periodic dielectric structure [21,22]. PWEM solves Maxwell's equation in Fourier space. Maxwell's equations for time harmonic case can be written as :

$$\nabla \times \left(\frac{1}{\varepsilon_r}\nabla \times \vec{H}\right) = k_0^2 \mu_r \vec{H} \quad (1)$$

$$\nabla \times \left(\frac{1}{\mu_r}\nabla \times \vec{E}\right) = k_0^2 \varepsilon_r \vec{E} \quad (2)$$

The equations assumed a) field strength is small b) material is macroscopic and isotropic c) material dispersion is ignored d) material is transparent e) structure does not vary with time f) there are no free charges or currents.

The above equations are written in the form of a eigen value problem. These equations can be written as a three dimentional (3D) eigen value problem in fourier space as a matrix so that it can be solved numerically.

$$[\vec{K}\times]\frac{1}{[\varepsilon_r]}[\vec{K}\times]\vec{u} = -k_0^2[\mu_r]\vec{u} \quad (3)$$

$$[\vec{K}\times]\frac{1}{[\mu_r]}[\vec{K}\times]\vec{s} = -k_0^2[\varepsilon_r]\vec{s} \quad (4)$$

Here $k_0^2 = \omega\sqrt{\varepsilon_0\mu_0}$, u and s corespond to the H and E field in fourier space .

By solving equation (3) and (4) the dispersion diagram was calcualted.

PWEM is the most popular method used to genarate the dispersion diagram but to simulate propagation mode and transmission or reflection spectra this method cannot be used. In order to simulate the poparating mode we employed a Finite Element Method (FEM) based solver, which produce the tranmission /reflection spectra and field pattern of the structure. A Perfectly Matched Layer (PML) was utilized around the srtucture to stop reflections at the boundary.

III. Simulation and Results:

The concerned filter is depicted in Fig. 3. The filter is made with array of dielectric rod with square lattice structure. The filter is composed of 4 port A, B, C and D. An electromagnetic excitation is introduced in port A and then observed at port B, C and D. By comparing the incident wave to the wave at B, C and D port the transmission spectra can be determined. Transmissivity is defined as the ratio of the square of the amplitude of transmitted intensity at the output ports. To that of an input port. The electromagnetic excitation was varied from 1530nm to 1595nm to generate the transmission spectra at different port.

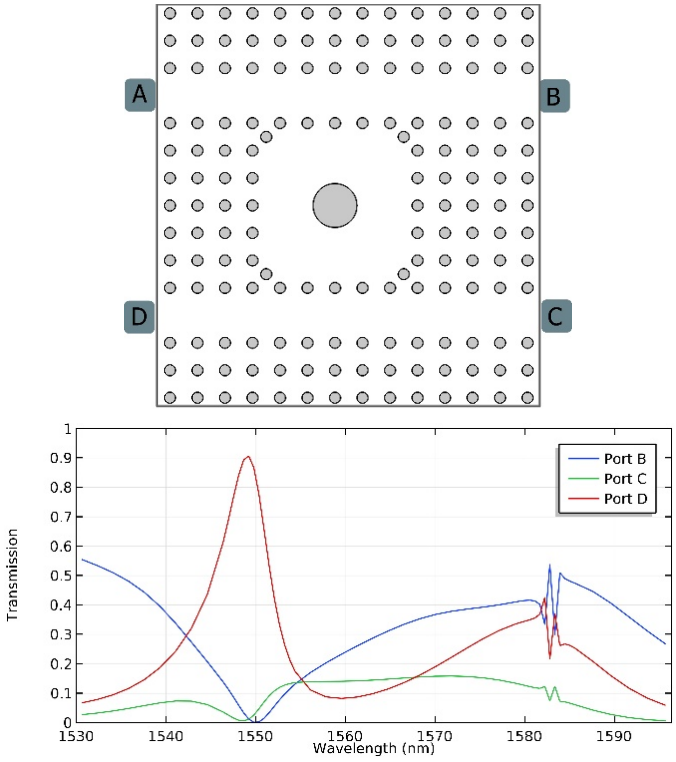


Fig. 3. Top: geometry of the filter Bottom: transmission spectra of the filter at B, C, D port.

Using the parameter mentioned on the previous section the transmission spectra can be seen in Fig. 3. The peak transmission was found 90.7% at a wavelength of 1549nm at port D. At the peak point transmittance at port B is 1.27% and at port C is 1.17%. Some power is also lost in port A. The back- reflection occurs mainly due to the coupling mismatch at port A [23]. The quality factor, $(Q = \lambda_0/\Delta\lambda)$ at 1549nm is 217. The bandwidth is calculated at Full Width Half Maximum (FWHM).

So it is clear that the drop port of the filter is Port D. For that purpose, for the rest of this paper we will focus our study on port D. By varying different parameter, the transmittance can be varied or the peak point can be shifted to another targeted wavelength. By varying the inner rod radius, the lattice rod radius or by changing the dielectric constant of the rods in general different attribute of the transmission spectra can be varied. Changing the inner rod radius results in the peak transmittance, which can be seen in Fig. 4. But it can also be seen that the peak value is blue shifted when the radius is reduced bellow 4*r and various artifacts appears when the inner radius is 4.4*r. Also with different simulation it was also found with a radius of less than 3*r or above 5*r the resonant structure completely loose its resonant property and the filter is no longer functions.

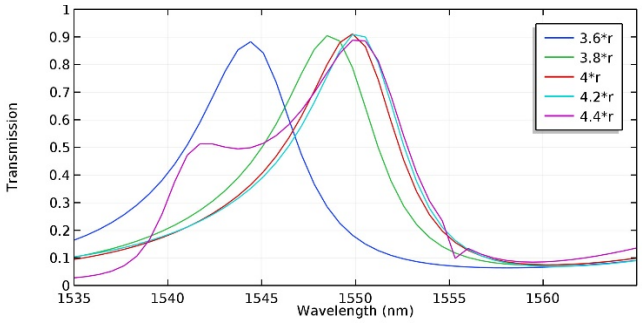


Fig. 4. Transmission spectra at port D while varying the inner radius

While varying inner rod by a small amount from 4*r does not change the output by much, changing the radius of the lattice rod shift the peak point to the right or to the left.

Effect of changing the radius of dialectic rod can be seen in Fig. 5. By increasing the radius of the dialectic rod the spectrum shifts to the right. Also the peak value of the transmission increase slightly when increasing the radius of the rod. But after increasing the radius up to a point the peak transmission levels off and then starts to decrease.

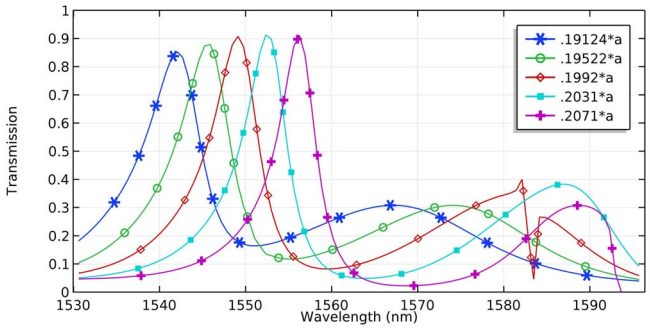


Fig. 5. Transmission spectra at port D while varying the radius of the dielectric rod

In Fig. 6. effect of changing the refractive index can be seen. By varying the refractive index from n=3.9 to n=4.4 the peak transmittance remains the same but the spectra red shifts. Changing the refractive index means changing the material itself. So by using different material the peak can be shifted without changing the peak transmittance.

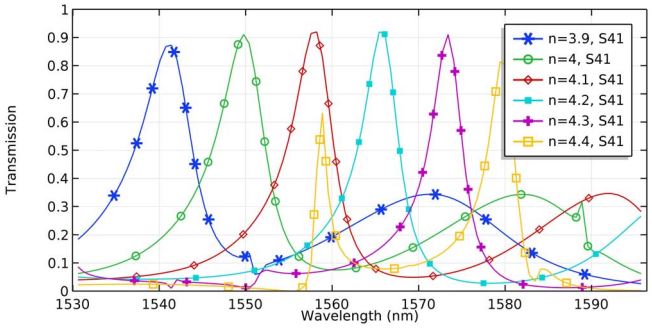


Fig. 6. Transmission spectra at port D while varying the refractive index.

IV. WORKING PRINCIPLE OF THE DROP FILTER:

The drop filter works by coupling a certain range of frequency in the cavity and rejecting the rest. The coupling is done by the coupling rod in the structure. Fig.7 shows the normalized electric field at the point of resonance. The color bar next to the figure show the scale of the relative electric field. Also in the figure the highest electric field is seen near the coupling rod. The input coupling rod couples the EM energy from A–B waveguide path. The coupled energy is then dropped to the C-D waveguide path. The light is confinement in the intended path which can be seen by the near zero electric field (dark blue part in Fig. 7).

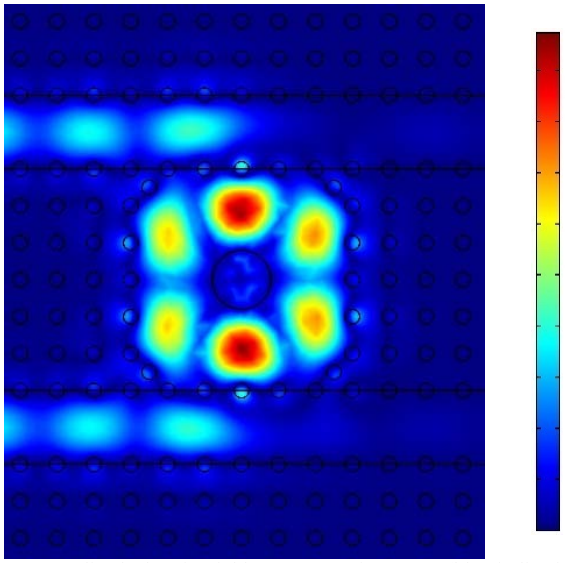


Fig. 7. Normalized Electric Field at resonant frequency (blue indicating no electric field and red indicate high electric field)

From the time average power flow diagram in Fig. 8 (Top) it can be seen when the power circulates clockwise around the inner rod the maximum power is dropped in port D. On the contrary when the structure couples power as such that the power flow on both side of the inner rod Fig. 8 (Bottom) the power flows at port B, C and port D. The coupling rod couples some energy which is divided between ports C and D. But most of the power flows through port B

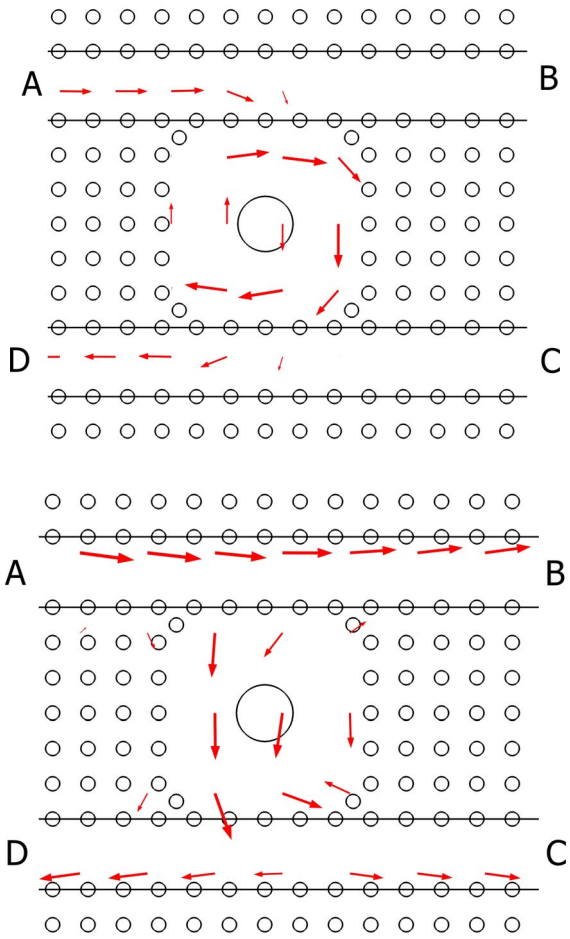


Fig. 8. Time average power flow of the filter Top: at peak resonant wavelength Bottom: at a non resonant wavelength (Length of the arrow represents relative time avgegare power flow in log scale)

V. Conclusions

In this study a simple homogeneous drop filter was simulated to show its various performance aspects. Plane Wave Expansion Method (PWEM) was performed to find out the photonic band gap of the lattice structure. Then using the parameter found in the study a novel structure was created which performs as a drop filter at a communication wavelength of 1550nm. From Finite Element Method (FEM) study the drop transmittance was found close to 91% at 1550nm which can be tuned to different wavelength by varying different parameter of the structure.

As this structure is very compact (8 µm x 9µm) it can also be included in photonic integrated circuit. As the dispersion diagram is normalized and due to the scalability of photonic band gap (PBG) structure the filter can be easily designed for lower or higher wavelength range.

References

[1] Yablonovitch, Eli. "Inhibited spontaneous emission in solid-state physics and electronics." Physical review letters 58.20 (1987): 2059.

[2] John, Sajeev. "Strong localization of photons in certain disordered dielectric superlattices." Physical review letters 58, no. 23 (1987): 2486.

[3] Mekis, Attila, J. C. Chen, I. Kurland, Shanhui Fan, Pierre R. Villeneuve, and J. D. Joannopoulos. "High transmission through sharp bends in photonic crystal waveguides." Physical Review Letters 77, no. 18 (1996): 3787.

[4] J.D. Joannopoulos, R.D. Meade, J.N. Winn, Photonic Crystal: Modeling of Flow of Light, Princeton University Press, Princeton, NJ, 1995.

[5] S. H. Kim, H. Y. Ryu, H. G. Park, G.-H. Kim, Y.-S. Choi, and Y. H. Lee, "Two-dimensional photonic crystal hexagonal waveguide ring laser", Applied Physics Letters, 81 (2002) 2499-2501.

[6] V. D. Kumar, T. Srinivas and A. Selvarajan, "Investigation of ring resonators in photonic crystal circuits", Photonics and Nanostructures, 2 (2004) 199-206.

[7] [7] Z. Qiang, W. Zhou, and R. A. Soref, "Optical add-drop filters based on photonic crystal ring resonators", Optics Express, 15 (2007) 1823-1831

[8] Z. Ma, K. Ogusu, "Channel drop filters using photonic crystal Fabry–Perot resonators, Optics Communications", 284 (2011) 1192-1196.

[9] A. Dideban, H. Habibiyan, H. Ghafoorifard, "Photonic crystal channel drop filter based on ring-shaped defects for DWDM systems", Physica E, 87 (2017) 77-83.

[10] A. Taalbi, G. Bassou, M. Y. Mahmoud, "New design of channel drop filters based on photonic crystal ring resonators", Optik, 124 (2013) 824-827.

[11] F. Monifi, A. Ghaffari, M. Djavid, and M. S. Abrishamian, "Three output port channel drop filter based on photonic crystals", Applied Optics, 48 (2009) 804-809.

[12] Hamed Alipour-Banaei, Farhad Mehdizadeh, Mahdi Hassangholizadeh-Kashtiban, "A new proposal for PCRR-based channel drop filter using elliptical rings", Physica E, 56 (2014) 211-215.

[13] R. Bendjelloul, T. Bouchemat and M. Bouchemat, "An optical channel drop filter based on 2D photonic crystal ring resonator", Journal of Electromagnetic Waves and Applications, 30 (2016) 2042-2410.

[14] V. Kannaiyan, S. K. Dhamodharan, R. Savarimuthu, "Investigation on modified quasi-square PCRR based demultiplexer for WDM applications", Optical and Quantum Electronics, 48 (2016) 393.

[15] Z. Rashki, S. Javad S. M. Chabok, "Novel design of optical channel drop filters based on two-dimensional photonic crystal ring resonators", Optics Communications, 395 (2017) 231-235.

[16] M. A. Mansouri-Birjandi, A. Tavousi, M. Ghadrdan, "Full-optical tunable add/drop filter based on nonlinear photoniccrystal ring resonators", Photonics and Nanostructures– Fundamentals and Applications, 21 (2016) 44-51.

[17] F. Mehdizadeh, H. Alipour-Banaei and S. Serajmohammadi, "Channel-drop filter based on a photonic crystal ring resonator", Journal of Optics, 15 (2013) 075401

[18] Ghaffari, Afshin, Faraz Monifi, Mehrdad Djavid, and Mohammad S. Abrishamian. "Photonic crystal bends and power splitters based on ring resonators." Optics Communications 281, no. 23 (2008): 5929-5934.

[19] Djavid, Mehrdad, Faraz Monifi, Ali Ghaffari, and M. S. Abrishamian. "Heterostructure wavelength division demultiplexers using photonic crystal ring resonators." Optics communications 281, no. 15-16 (2008): 4028-4032.

[20] Robinson, S., and R. Nakkeeran. "Investigation on two dimensional photonic crystal resonant cavity based bandpass filter." Optik-International Journal for light and electron optics123, no. 5 (2012): 451-457.

[21] Troullier, Norman, and José Luís Martins. "Efficient pseudopotentials for plane-wave calculations." Physical review B 43, no. 3 (1991): 1993.

[22] Raymond C. Rumpf, Lecture series on Computational electromagnetics, available at http://emlab.utep.edu/academics.htm

[23] Zhang, Juan, Hao Liu, Yipeng Ding, and Yang Wang. "A novel photonic crystal ring resonator configuration for add/drop filtering." Photonics and Nanostructures-Fundamentals and Applications 30 (2018): 14-19.

A High Speed and Low Power 32nm FinFET Dynamic Latch Comparator Design

Mir Muntasir Hossain
Department of Electrical and Electronic Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
mirmuntasir.eee@gmail.com

Satyendra N. Biswas
Department of Electrical and Electronic Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
sbiswas.eee@aust.edu

***Abstract*—A comparator establishes the bridge to connect the realms of digital and analog circuits. Moreover, the need to fulfil the consistent demand of low power and high speed small-scale devices has brought the MOS transistor based circuit design near to the edge of feasible design limit. In recent times, with the advancement in fabrication technology, FinFET has emerged as the most competent alternative with the prospective to overcome the short channel effect challenges of MOSFET. In this paper, we have presented a dynamic latch comparator circuit using 32nm PTM model of FinFET and performed our simulations with LTspice and MATLAB. Our designed comparator has shown an impressive propagation delay time of only 7.25 ps and average power consumption of 45.64 µW at 0.8V dc supply. Also, its performance comparisons with other contemporary designs verified the supremacy of our proposed comparator. Moreover, it seems that in the upcoming years, FinFET has the potential to bring forth radical transformation in device modeling and supersede conventional MOSFET oriented design.**

Keywords—dynamic latch comparator, FinFET, PTM, low power, high speed

I. INTRODUCTION

Analog to Digital Converter (ADC) and Digital to Analog Converter (DAC) are the two cardinal doorways between the digital and analog world [1]. They have numerous applications in diverse field of areas like memory circuits, switching power regulators, data converters, oscillators, RF communication, etc. [2], [3]. The comparator is the core element for both of the converters. Due to the speed limitation and massive power requirement, static converters have been superseded by dynamic converters [4]. A dynamic latch comparator meets the essential requirements of circuit simplicity, high input impedance, smaller chip area, high frequency operation, reduced power dissipation, low supply voltage operation, noise immunity, lower offset, etc. [5], [6]. Over the years, many researchers have been working relentlessly to develop more and more efficient design of comparators. In [7], to acquire faster speed, an NMOS input comparator has been suggested; however, the power consumption increased dramatically. In order to increase the speed, auxiliary transistors have been included by B. Goll et al., although it enhanced the area and power requirement [8]. For lowering the power absorption and improving the speed performance, a delayed clock pulse based design is proposed in [9]. Although the gain amplified with the deferred clock pulse, but the kickback noise issue affected the circuit seriously. Other than voltage boost or additional transistors for conventional body driven comparators, an ultra-low voltage design has been proposed in [10] for further delay reduction, but still the delay remained high.

The ubiquitous usage of portable devices has escalated the demand of low power and high speed circuitry in recent years. For the last couple of decades, CMOS technology has solely dominated the circuit design industry. But as the technology is continuously shrinking, several short channel effects like the carrier's quantum mechanical tunneling through the gate oxide, sub-threshold leakage current due to drain induced barrier lowering (DIBL), higher threshold voltage in comparison with the drain to source supply voltage, mobility degradation of planar MOS transistors have inspired the researchers to explore for alternative possibilities [11], [12]. The double or multi gate structure which is most commonly known as FinFET (fin field-effect transistor), shown in fig. 1 [13], has successfully overcome the issues regarding short channel and emerged as the most promising alternative of conventional MOSFET. A FinFET has quasi planar structure which is very similar to planar MOSFET structure thus the fabrication process are analogous. A very thin silicon film is grown over the wafer and the gate wraps around the channel which facilitates the channel control from both front and back ends, and consequently, a better control is achieved over the threshold voltage and sub threshold leakage current [14], [15]. The channel of a FinFET is oriented vertically in contrast with the horizontal channel of MOSFET. Interestingly, the width of a channel is approximately twice the height of the channel plus the gate oxide thickness, which is known as width quantization; moreover, the effective width of a FinFET can be enlarged by raising the number of fins [16].

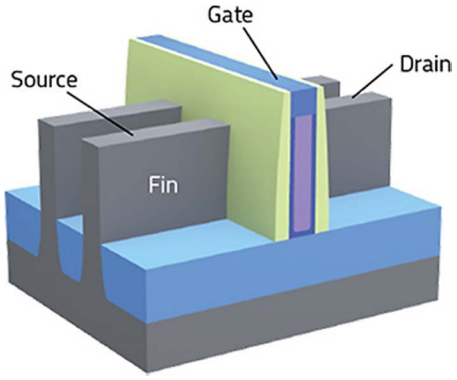


Fig. 1. FinFET architecture

In this paper, we have proposed a dynamic latch comparator using 32nm FinFET technology and excluding the preamplifier stage as it consumes significant amount of power. In section II, the circuit description of our proposed comparator together with propagation delay time and power consumption have been discussed. Simulation results and comparison with other research works are presented in section III. Lastly, section IV brings to an end with a brief conclusion.

II. Proposed Dynamic Latch Comparator

A. Circuit Description

In contrast with the amplifier chain type comparators, the dynamic latch comparators provide superior speed and low amount of power consumption attributes which particularly place them in a favorable position for designer's choice to implement flash A/D converters. Another distinct feature of latch type comparator is their immunity towards noise as once the decision is reached the output will hold that data indefinitely till next input.

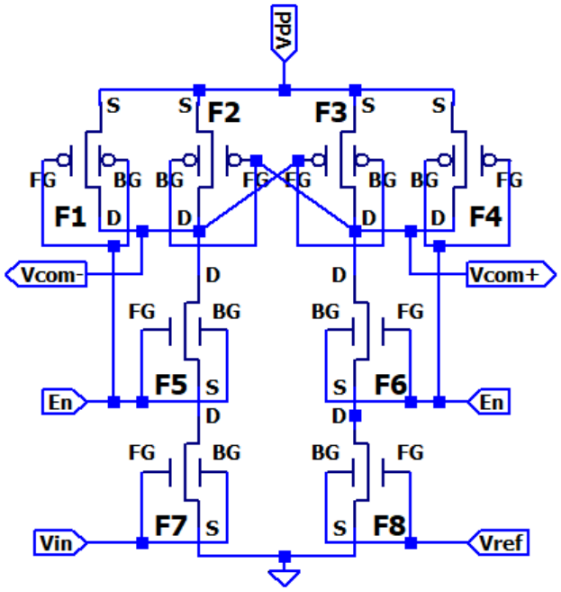


Fig. 2. Circuit design of the proposed dynamic latch comparator

In order to design the dynamic latch comparator, due to the small channel limitations of MOS transistor, FinFET has been introduced. The pFinFET and nFinFET have been designed by using the 32nm PTM (Predictive Technology Model) with LTspice. The schematic representation of our proposed dynamic latch comparator has been shown in fig. 2. The latch stage between F2 and F3 pFinFETs establish a cross coupled pair where the drain of one pFinFET is attached to the gate of another. The input and reference voltages are applied to the gate of F7 and F8 nFinFETs, respectively. The comparator performs its operation in two phases. The first phase is called the "Reset Phase" while the Enable (En) signal remains logic low. En is connected to the gate of F1, F4, F5 and F6. So, as the F5 and F6 nFinFETs are getting logic low, both of them enter into the cut off region, where as the F1 and F4 pFinFETs enter into the triode region or in simple words are turning on [17]. Now, through F4 and F1 both Vcom+ and Vcom- are charged to Vdd, and this voltage will be remained there as the discharging paths through F5 and F6 are disconnected. Input voltage (Vin) and reference voltage (Vref) might still be present, but during the reset phase their amplitudes are not compared. If F5 and F6 nFinFETs were not used then Vcom+ and Vcom- nodes would have never be remained charged at Vdd.

The next phase of the comparator is called the "Active Phase" where a logic high signal is applied through the En input. F5 and F6 nFinFETs will immediately enter into the triode region which means they will turn on. At the same time as both F1 and F4 pFinFETs are getting logic high to their gate, they will turn off as they would move into the cut off region. Initially, both Vcom+ and Vcom- would remain at Vdd. Now, depending on the voltage levels of Vin and Vref, the discharging path conductance of F7 and F8 would differ. If Vin is greater than Vref then the conductance of F7 would be higher than F8 thus will facilitate the discharge of Vcom- earlier than Vcom+. Once Vcom- is discharged to Vdd-Vtp level, where Vtp is the threshold voltage pFinFET, the gate of F3 pFinFET through the drain of F1 will receive logic low which will eventually turn on the F3 as the source of F3 is already connected to Vdd. So, Vdd volt will appear at Vcom+ node. Moreover, as the gate of F2 pFinFET is connected to the drain of F3, it will get logic high thus will remain into the cut off region. In this active phase the F2 pFinFET never gets the opportunity to turn on if Vin remains higher than Vref. Hence Vcom- will maintain the ground voltage. However, if the input signal is lower than the reference level then in that case the conductivity of F8 would be higher than F7 and Vcom+ would discharge swiftly than Vcom- and reach at ground voltage level. Now as the gate of F2 pFinFET will achieve logic low, it will turn on and the drain voltage would be Vdd which will be appeared on Vcom- node. Every time for the input signal to be compared with the reference signal, the dynamic latch comparator must have to be reset first and the subsequent process will be continued.

B. Propagation Time Delay

One of the most important parameters of logic circuit is propagation time delay (t_{PD}) by which a designed circuit's performance is declared as high speed or low speed. With the advancement of technology, it has always been demanded for higher switching speed thus resulting high frequency operation which could only be achieved by shorter propagation delay. Two basic measurements associated with t_{PD} are t_{PHL} and t_{PLH}. The time interval between the 50% point of input and 50% point of corresponding output that goes from logic high to low is t_{PHL} and similarly, time duration between 50% points of input and output while it goes form logic low to high is t_{PLH}.

$$t_{PD} = \frac{t_{PHL}+t_{PLH}}{2} \quad (1)$$

C. Power Consumption

Another major issue, on which the researchers have been working relentlessly decade after decade is the reduction of power consumption of a circuit. It is of the paramount importance for high frequency operation, otherwise each switching would consume huge amount of power that would eventually overheat the circuit, cost more, increase device size and deteriorate the circuit performance over time. There are two basic components as explained below on which the total power consumption of a circuit depends [18].

1) Static Power Consumption: During the operation, while all the inputs of a circuit are at a fixed voltage or logic level and the circuit is not in a charging state, due to the establishment of reverse bias junction between the substrate and the source/drain some leakage power will always be drained due to the flow of minority carriers is called the static power consumption.

2) Dynamic Power Consumption: When the transistors of a circuit switch from one logic state to the opposite, current flows from Vdd to GND for a brief period of time as both p-type and n-type transistors remain on simultaneously; moreover, some current is also passed to charge the internal nodes and output capacitive loads. Both of these factors

contribute to higher amount of power absorption which is called the dynamic power consumption. Majority of the power consumed by a circuit is due to the dynamic power for high speed devices.

Total Power Consumption = Static Power Consumption + Dynamic Power Consumption (2)

The average power consumed by a circuit is determine by the following equation

$$P_{avg} = \frac{E}{T} = \frac{1}{T}\int_0^T I_{dd}(t)V_{dd}(t)dt \quad (3)$$

III. Simulation Results and Discussion

The dynamic latch comparator presented in this paper is designed with LTspice simulation software. Moreover, the simulation data have been exported into MATLAB to plot all the waveforms with greater ease. Most importantly, the 32nm FinFET Predictive Technology Model (PTM) that we have used for our design was developed by Arizona State University which is available at their website [19].

In fig. 3, all the essential voltage amplitude and power consumption levels have been clearly demonstrated. The dc supply voltage (Vdd) is chosen as 0.8V for the 32nm FinFET process. The Enable (En) or clock signal is a periodic square wave which for a brief period of time of 0.1ns remains at 0V in order to provide the scope to both of the outputs, Vcom+ and Vcom-, to be charged at Vdd. This time could be varied depending on the application requirements. The input voltage (Vin) has been delicately set higher than the reference voltage (Vref) in one cycle and lower in another with respect to En in order to indicate the comparison for both cases. It is clearly evident from fig. 1 that Vcom+ is at logic high and Vcom- is logic low while Vin is greater than Vref and vice versa. From the average power consumption curve we can see the spikes of dynamic power consumption during momentary switching periods.

In Table I, the time delay and power consumption have been presented for different amplitude of dc supply voltages. It is very interesting to note that the t_{PHL} (high to low transition time of output) has increased with the increment of dc supply voltage. It is due to the higher capacitance of FinFET, obviously more time would be required while discharging from a higher stored charge level. On the other hand, t_{PLH} (low to high transition time of output) decreased with higher Vdd, as because it would take less time to charge the capacitance with higher voltage. Average power consumption definitely had to increase for higher supply voltage.

The comparison among our proposed system with other works has been included in Table II for convenience. It is clearly manifesting that the FinFET technology is superior in case of power consumption and propagation delay time than its CMOS counterpart.

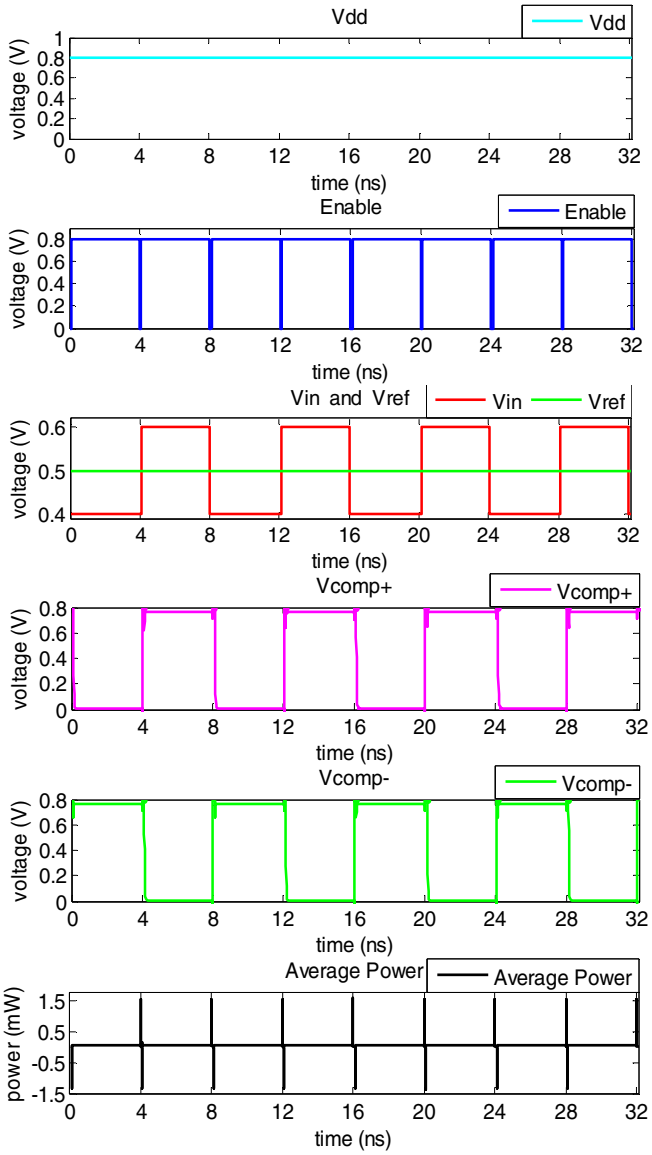


Fig. 3. Simulation results of the proposed dynamic latch comparator

TABLE I. Propagation Delay Time and Average Power Consumption with Different Supply Voltages

Supply voltage (V)	t_{PHL} (ps)	t_{PLH} (ps)	Propagation time delay, t_{PD} (ps)	Average Power (µW)
0.6	10.02	3.84	6.93	24.29
0.8	10.89	3.61	7.25	45.64
1	11.76	3.50	7.63	73.10
1.2	12.57	3.39	7.98	107.27

TABLE II. Performance Comparisons Among the Proposed and Other Dynamic Latch Comparators

Parameters	[20]	[21]	[22]	[23]	Proposed
Technology	CMOS	CMOS	CMOS	CMOS	FinFET
Channel Length	0.18µm	0.18µm	0.18µm	90nm	32nm
Supply Voltage	1.8V	1.8V	1.8V	1.2V	0.8V
Power Consumption	146.2 µW	460 µW	158.5 µW	225 µW	45.64 µW
Sampling Frequency	100 MHz	100 MHz	50 MHz	2 GHz	250 MHz
Propagation Delay	7ns	1.699ns	4.2ns	500ps	7.25ps

IV. Conclusion

In this research work, the design of a dynamic latch comparator, developed on 32nm FinFET PTM model, has been suggested and through simulations its performance has been evaluated. The proposed comparator circuit consumes less power and requires shorter delay time in comparison with other similar dynamic latch comparators as it is made with less number of transistors thus requires less chip area. FinFET based design results are so far remarkable and has opened the scope for researchers to explore the possibility and impacts on other essential circuit models.

References

[1] G. Yongheng, C. Wei, L .Tiejun, and W. Zongmin. “A novel 1GSPS low offset comparator for high speed ADC”, 5th International Joint Conference on INC, IMC and IDC, IEEE Computer Society, Seoul, pp.1251-1254, 25-27 August 2009.

[2] M.A.S. Bhuiyan, M.B.I. Reaz, J. Jalil, L. F. Rahman, "Design of a nanoswitch in 130 nm CMOS technology for 2.4 GHz wireless terminals", Bulletin of the Polish Academy of SciencesTechnical Sciences, vol. 62, pp. 399-406, 2014.

[3] F. Mohd-Yasin, M. Khaw, M. Reaz, "Radio frequency identification: Evolution of transponder circuit design", Microwave Journal, vol. 49, pp. 56-70, 2006.

[4] Razavi, Behzad, and Bruce A. Wooley. ”Design techniques for highspeed, high-resolution comparators.” Solid-State Circuits, IEEE Journal of 27.12 (1992): pp. 1916-1926.

[5] Chi-Hang Chan, Van Zhu, U-Fat Chio, Sai-Weng Sin, Seng-Pan U, R.P.Martins, "A Reconfigurable Low-Noise Dynamic Comparator with Offset Calibration in 90nm CMOS" In Preceding OfTEEE Asian Solid-State Circuits Conference, pp. 233-236, 14-16 Nov 2011.

[6] P. Uthaichana and E. Leelarasmee, "Low Power CMOS Dynamic Latch Comparators," IEEE, pp. 605-608, 2003.

[7] D’Amico, Stefano, et al. ”A 7.65-mW 5-bit 90-nm 1-Gs/s Folded Interpolated ADC Without Calibration.” Instrumentation and Measurement, IEEE Transactions on 63.2 (2014): 295-303

[8] B. Goll and H. Zimmermann, “ A 65nm CMOS comparator with modified latch to achieve 7GHz/1.3mW at 1.2V and 700MHz/47µW at 0.6V,” in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. papers, pp. 328-329, Feb. 2009.

[9] Hassanpourghadi Mohsen, Zamani Milad, Sharifkhani Mohammad. A lowpower low-offset dynamic comparator for analog to digital converters. Microelectron J 2014;45(2):256–62

[10] S. Babayan-Mashhadi, M. Sarvaghad-Moghaddam, "Analysis and Design of Dynamic Comparators in ultra-low supply voltages", 22 nd Iranian Conference on Electrical Engineering (ICEE) , May 2014.

[11] D. Blaauw, B. Zhai,“Energy Efficient Design for Subthreshold Supply Voltage Operation”, IEEE International Symposium on Circuits and Systems (ISCAS), 2006.

[12] H. Soeleman, K. Roy, B.Paul,“Sub-Domino logic: ultra-low power dynamic sub-threshold digital logic,” VLSI Design, pp.211-214, 2001

[13] Draeger, N. (2016, September 12), Tech Brief: FinFET Fundamentals. Available at : https://www.lamresearch.com/products/our- solutions/ transistor-solutions/

[14] Mayank Shrivastava, M. S. Baghini, D. K. Sharma, V. R. Rao “A novel bottom spacer FinFET structure for improved short-channel, power-delay, and thermal performance” IEEE Trans. On Electron Devices, Vol. 57, No. 6, june 2010, pp. 1287-1294.

[15] J. Kim, K. Roy, “Double Gate-MOSFET Subthreshold Circuit for Ultralow Power Applications,” IEEE Trans. Electron Devices, vol. 51 (9), pp.1468-1474, 2004.

[16] D. Bhattacharya, N. K. Jha, "FinFETs: From devices to architectures" in Advances in Electronics, Hindawi Publishing Corporation, pp. 21, 2014.

[17] Sedra, Adel S.; Smith, Kenneth C. (November 14, 2014). Microelectronic Circuits (7 ed.). Oxford, UK: Oxford University Press. p. 1488. ISBN 0199339139.

[18] CMOS Power Consumption and Cpd Calculation. Literature number SCAA035B. -Dallas, USA: Texas Instruments, June 1997. [Online] Available from: http://www.ti.com/lit/an/scaa035b/scaa035b.pdf

[19] Available at: http://ptm.asu.edu/

[20] Zhaohui Huang, Peixin Zhong, "An adaptive analog-to-digital converter based on low-power dynamic latch comparator," IEEE International conference, pp. 6, May 2005.

[21] Ili Shairah, Abdul Halim, Nurul Aisyah Nadiah Binti Zainal Abidin, A'zraa AfhzanAb Rahim, "Low Power CMOS Charge Sharing Dynamic LatchComparator using 0.18µm Technology", IEEE Regional Symposium on Micro and Nanoelectronics (RSM), pp. 156-160, Sep 2011.

[22] L.F. Rahman, M.B.I. Reaz, C.C. Yin, "Design of high speed and low offset dynamic latch comparator in 0.18 µm CMOS process", PLOS ONE, vol. 9, no. 10, pp. e108634:1-e108634:12, 2014.

[23] D. Schinkel, E. Mensink, E. Kiumperink, E. Tuijl, B. Nauta, “A Double-Tail Latch-Type Voltage Sense Amplifier with 18ps Setup+Hold Time,” IEEE Int. Solid-State Circuits Conference Dig. Tech. Papers, pp. 314-315, February 2007.

MOSFET-like Carbon nanotube Field Effect Transistor based Full adder design

Bappy Chandra Devnath and Satyendra N. Biswas*

Department of Electrical and Electronic Engineering
Ahsanullah University of Science & Technology, Dhaka, Bangladesh

*Corresponding Author: sbiswas@linuxmail.org / sbiswas.eee@aust.edu

First Author: b.devnath20@gmail.com

***Abstract*— CMOS transistor of the commercial chip is currently at 7nm technology node. Fabrication engineer is now facing some complexities to further scale down the size of the transistor. Quantum mechanical phenomenon is occurred and significantly worsens the performance as well as it changes the overall properties of the transistor. So, researchers have developed several types of devices in order to overcome the shortcomings of CMOS. One of the devices is the Carbon nanotube Field Effect Transistor (CNFET). It is a suitable device because it can be scaled down up to 2nm and it overcome some limitations. One of the major advantages of CNFET is that the carrier faces theoretically zero resistance when passing through the carbon nanotube because that tube acts as a channel analogous to E-MOS. In this research, a new adder circuit is designed using 32nm CNFET model. The proposed circuit is based on Pass Transistor Logic (PTL) in order to reduce the number of transistors and silicon area. A comparative study is conducted by simulation. The results concluded that the proposed circuit performs better than other conventional adders in terms of power consumption and time delay.**

Keywords*— *Carbon Nanotube, Chirality Angle, NWFET, CNFET, CPL Full Adder, PDP

I. INTRODUCTION

MOSFET is widely used in the integrated circuit since late 1960. In order to improve the performance such as higher switching speed, low power consumption and high latency of a digital chip, more transistors are to be fabricated very tiny space. So, it is necessary to reduce the size of the transistor to fit a large number of them. If the size of the transistor is significantly scaled down, nonlinear properties of the transistor become prominent. We can see some nonlinear behavior of MOSFET such as channel length modulation, subthreshold leakage, drain induced barrier lowering (DIBL), Junction capacitance, etc. If the channel length of MOSFET is less than 10nm (theoretically) the quantum mechanical effect is to appear (Tunneling). It is realized that we cannot scale down of MOSFET less than an atom but the problem will arise before reaching that level. So, researchers are working to find out a new type of device.

Many devices are proposed [4] but the two devices are in the front. One is Nanowire FET (NWFET) and another one is CNFET as shown in figure 1 [4]. CNFET is made out with the CN tube as shown in figure 2, which is invented after 1990. CN tube is a graphite sheet which is rolled in a certain angle known as chirality angel. On the basis of chirality angles, the tube acts as a semiconductor or conductor. So, by using doping principle one can make a FET. Two types of CNFET

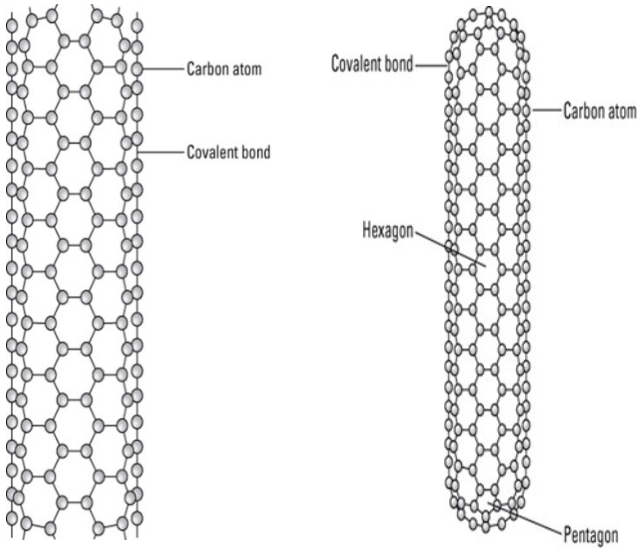


Figure 1: Carbon nanotube a) open, b) closed [google]

are existing: one is Schottky barrier CNFET and the other is Enhancement type CNFET [2]. Enhancement type CNFET is similar to the MOSFET. Different Logic circuits are designed using E-CNFET.

In this research, a new 1-bit full adder is designed using E-CNFET. Stanford university 32nm CNFET model is used [8]. Previous models of full adder circuits are discussed in section II, Section III describes the proposed model and simulation results are presented in section IV. Section V concludes the proposed research.

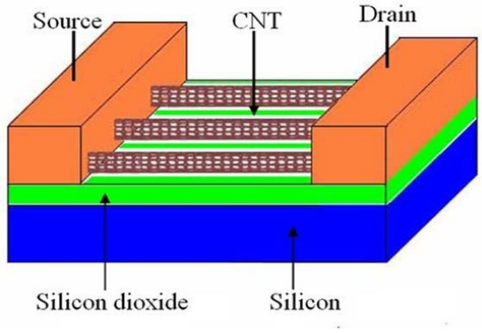


Figure 2: Carbon nanotube FET [google]

II. PREVIOUS WORKS ON FULL ADDER

CNFET is superior to MOSFET because of the ballistic transport of the carrier through the tube. Ballistic transport is basically the smooth traverse of electrons in a medium without encounters a scattering event. Theoretically, electron faces zero resistance during ballistic transport. As a result, the mobility of the carriers is theoretically infinite. For that reason, it is considered a prominent device in the near future.

In order to design an ALU, the adder is an important part. So many researchers proposed many adder circuits. Now we briefly discuss some published adder circuits

A. Majority function based full adder

Majority function based full adder means XOR-XNOR based full adder circuit. The capability of the majority function attracts the designer to use that function in their design [1]. Figure 3 shows the majority function based full adder [1].

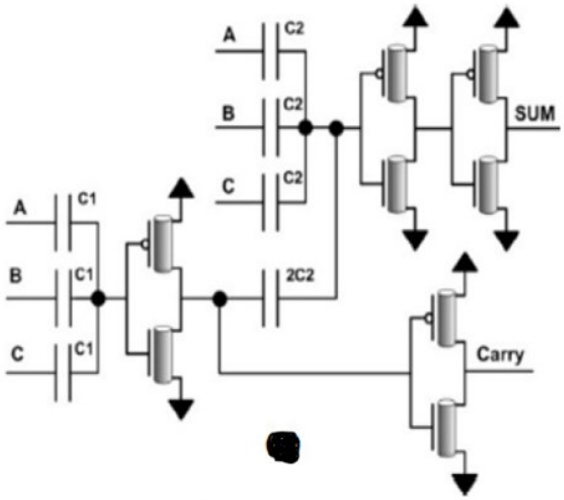


Figure 3: Majority Function based Full adder [1]

B. 24T Full Adder

This adder cell is proposed by A. Ghorbani and G. Ghorbani [1]. This adder is implemented with 24 CNFET and they claimed that the power of that adder is better than CMOS one. Figure 4 shows the 24T adder.

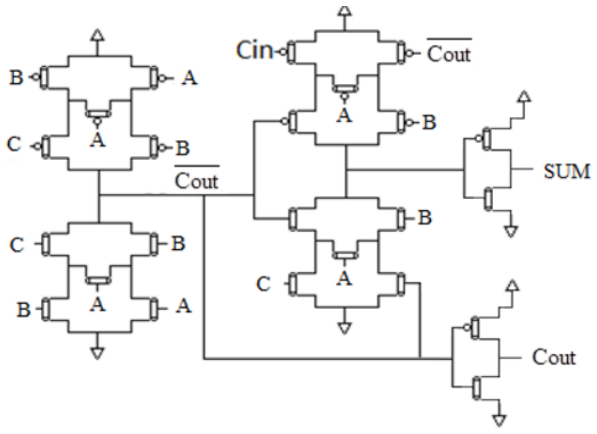


Figure 4: 24T Adder (A. Ghorbani and G. Ghorbani) [1]

C. TG Logic Based Full Adder

The advantage of TG logic-based circuit is the cell area is often very smaller [2]. The transistor count is smaller than CMOS logic. So, the TG logic-based system is power efficient compared with Conventional CMOS logic. Figure 5 shows the TG logic based Full adder [2].

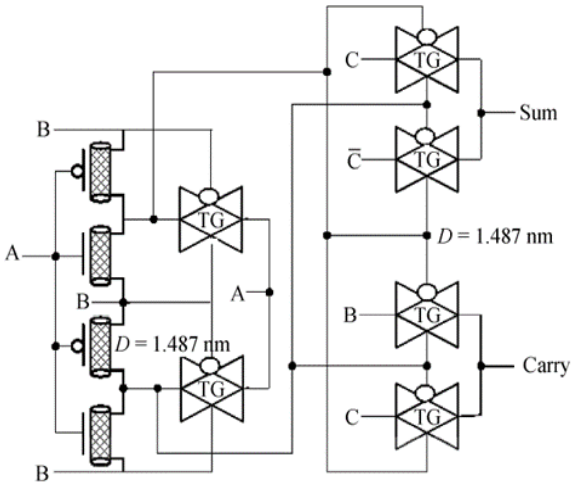


Figure 5: TG logic Based Full Adder [2]

D. Complementary logic based Full Adder

The advantage of CPL is that the complementary outputs are produced by in simple 4 transistor circuit [2]. The CPL full adder the sum and carry logic are structured separately on a more complex logic circuit [2]. In figure 6 the CPL full adder is showed.

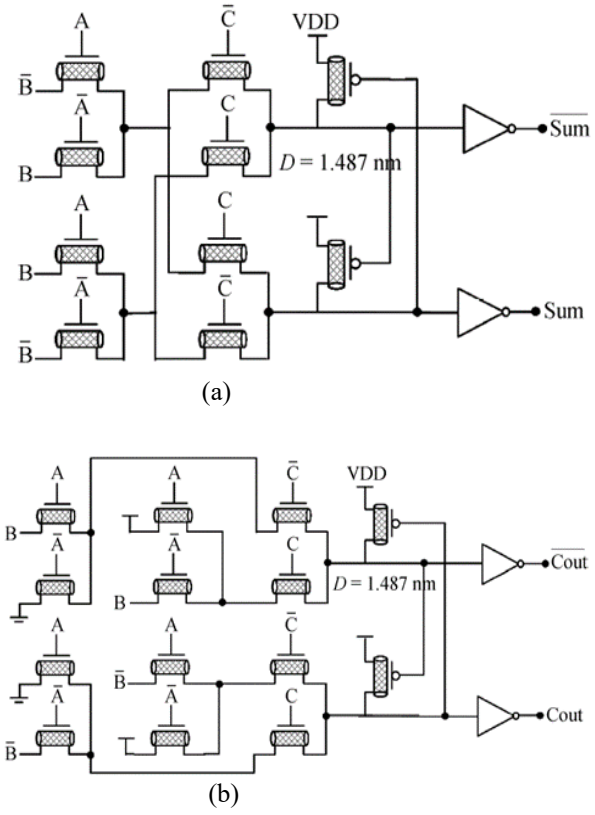


Figure 6: CPL full adder a) sum unit, b) carry unit [2]

III. PROPOSED FULL ADDER

The proposed adder is designed using PTL logic style. The goal is to design an efficient full adder which consumes less power, low silicon area, and better high-speed operation. So, we choose PTL, because it is a power efficient logic and the transistor count is also less. HSPICE is used for simulating our adder circuit. 10 transistors are needed for implementing the proposed adder. 4T XOR configuration is used for produce sum module and the carry module is generated using only 2 transistors. Figure 7 shows the schematic of our proposed adder.

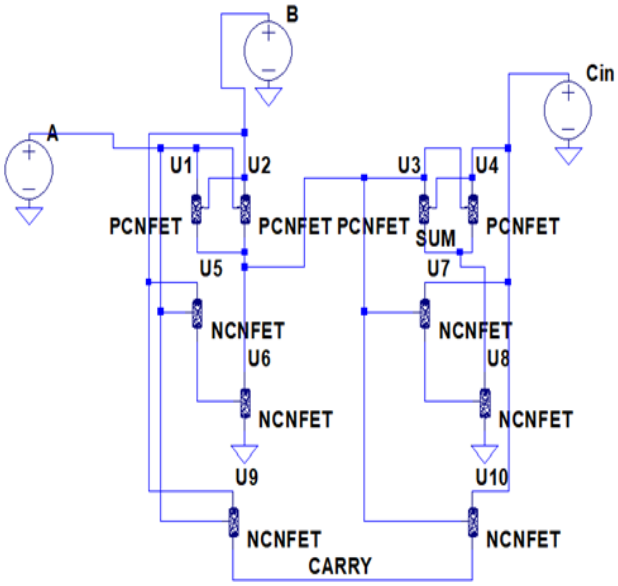


Figure 7: Proposed 10T Adder Using CNFET

TABLE I. PARAMETER LIST OF OUR PROPOSED CIRCUIT

Parameter Name	Value
Channel length	32nm
Channel width (PCNFET)	32nm
Number of the tube (PCNFET)	100
Channel width (NCNFET)	4nm
Number of the tube (carry NCNFET)	3
n1 and n2 (Chirality Vector)	(19,0)

IV. SIMULATION RESULTS

Table 1, displays the parameters, used in the simulation experiments in this research. The output of our adder circuit is given in figure 8. Some distortion is seen in the output. Here only one cycle is shown. The operating frequency of our circuit is 25MHz and the operating voltage is 0.9v.

To compare our findings with some published data [5-7]. The bar graphs of figure 9, 10 and 11 are given for comparing the performance among different types of adders. Figure 9 shows the Power, figure 10 displays the Time Delay and figure 11 depicts the PDP for adder circuits. From the calculated Power Consumption, Average Delay and the

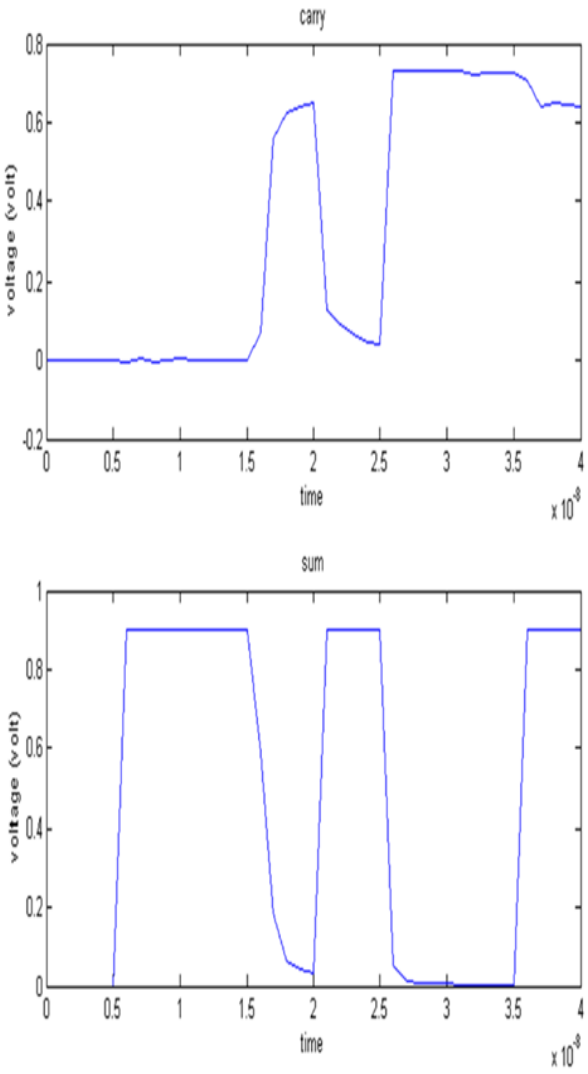


Figure 8: Output response of proposed 10T adder circuit

Power-Delay Product of the proposed adder, it is seen that the proposed circuit outperforms the other conventional models.

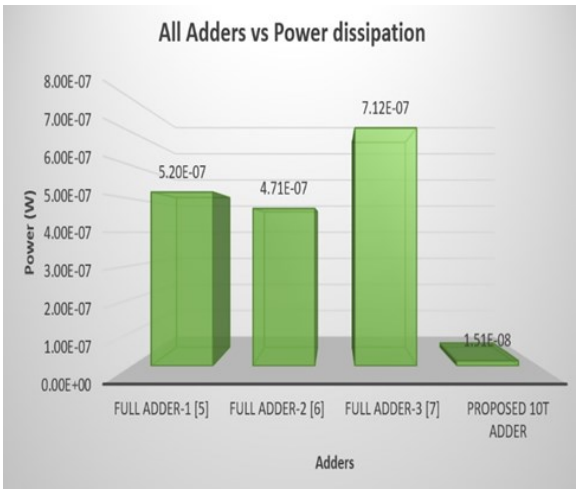


Figure 9: Bar Graph of Power dissipation of all Adders

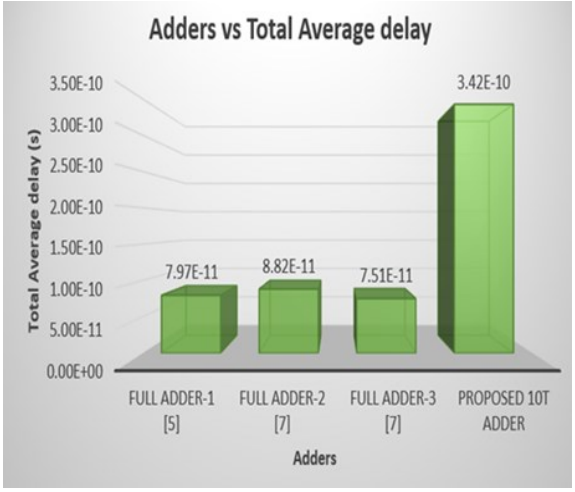


Figure 10: Bar graph of Average delay of all Adders

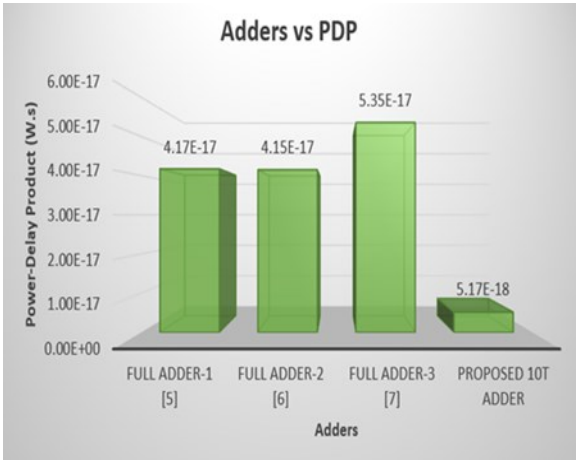


Figure 11: Bar graph of PDP of all Adders

V. CONCLUSION

A new adder is designed using a new type of device which is prominent now a day. We use the Stanford university model for obtaining a better result. From the comparative results, it can be concluded that the proposed circuit is superior in terms of power but slightly higher time delay. However, the proposed circuit is compact because it only needs 10 transistors. In terms of PDP, this circuit outperforms all other adder circuits.

VI. REFERENCE

[1]. A. Ghorbani and G. Ghorbani *"Energy Efficient Full Adder Cell Design by using Carbon nanotube Field Effect transistor in 32nanometer technology" International Journal of VLSI design & Communication Systems (VLSICS),* Vol.5. No.5, October 2014.

[2]. P.A Gowari Sankar and K. Udhayakumar *"MOSFET-like CNTFET based logic gate library for low-power application: a comparative study" Journal of semiconductor*, Vol. 35, No. 7, July 2014.

[3]. Babitha lincy R., Sharmila P., and Jenisha J. *"Design of Arithmetic Logic Unit using 32-nm CNFET Technology" IJCTA*, 9(9), 2016, pp. 3751-3758, International Science Press.

[4]. Reshma M. *"Carbon nanotube FET Device Fabrication, Design, Modeling, and Simulation of CNTFET For Reconfigurable Logic Gate Design" Asia Pacific Journal of Research*, ISSN: 2320-5504, E-ISSN-2347-4793, vol.1, Issue XVII, September 2014.

[5]. Navi K, Momeni A, Sharifi F, Keshavarzian P. *"Two Nobel ultra-high-speed carbon nanotube Full-Adder cells." IEICE Electronics Express*, 2009; 6, 19:1395-1401.

[6]. Navi K, Sharifi Rad R, Moaiyeri M H, Momeni A. *"A Low Voltage and Energy-efficient Full Adder Cell Based on Carbon Nanotube Technology." Nano-Micro Letters*, 2010; 2,2:114-120.

[7]. Navi K, Rashtian M, Khatir A, Keshavarzian P, Hashemipour O. *"High-Speed Capacitor-Inverter Based Carbon Nanotube Full Adder." Nanoscale Ress Letter*, 2010; 5:859-862.

[8]. Stanford University Nano Electronics Lab website.

[9]. HSPICE user Manual.

A Novel Approach to detect the Compound Level of Fertilizer and Pesticide in the Soil

Abu Shufian, Riadul Islam, and Md. Mominur Rahman
Department of Electrical and Electronic Engineering
American International University-Bangladesh, Dhaka-1229, Bangladesh
shufian2aiub@gmail.com, riadul999@gmail.com, and mominur18aiub@gmail.com

***Abstract*—Fertilizers (Triple superphosphate) and pesticide (fungicide and insecticide) are one of the major products used to increase the productivity of crops. As time passes the amount of pesticide in crops is not maintained fixed. Pesticide does react with water and the level of pesticide decreases day by day. Hence, it is important to know the present level of pesticide in soil and crops. TSP and pesticide (fungicide and insecticide) are used in this project as a sample of fertilizer and pesticide, cause of its high demand in Bangladesh. To detect the amount of TSP and pesticide in the oil sample, the conductivity property of TSP and pesticide (fungicide and insecticide) are used. Three separate solutions of TSP and pesticide (fungicide and insecticide) is made with water then AC source is applied to the solution instead of DC to avoid electrolysis of the given solution. Through the op-amp and two comparators, the final result is accomplished. Those two comparators, comparator 1 and comparator 2 were used to find the presence of 30%, 50%,70% TSP and Pesticide respectively. As result suggest, the resistance of the solution decreases with the incremental frequency. The compatible frequency for AC source is selected as 40 KHz. The reference voltage of comparator 1 is 1.4v and comparator 2 is 5v.**

Keywords—co-axial cylinder electrode, triple super-phosphate, fungicide, insecticide, phosphoric acid, electrolysis, comparator.

I. INTRODUCTION

From the start of cultivation farmers all around the world are using different technologies to increase the productivity of their crops. Therefore, in old days farmer used organic pesticides. However, with time, science has given farmers so many tools for an easier and effective irrigation system. The pesticide is one of those creations of science. Farmer of our country used fertilizer in a different amount for the different type of crops. Without considering the future outcome, they use the same amount of fertilizer. Sometimes this uncalculated use of fertilizer disrupts soil fertility. As a result, foods health quality gets decreased. Though some method can be used to test the existing rate of fertility in soil however this method is not only costly but also difficult to operate. Although fertilizer level detection could be done by a simple electrical circuit at a low cost using pesticides resistivity property. It was observed that if water is added to the pesticide solution the resistivity of the solution is decreased. This decrease depends on the number of pesticides used. So using that a set of data were taken for the different amount of fertilizer. From those data, a certain voltage level was set in the comparator to show the output through a LED for a different level of fertilizer. If the voltage level of a sample is greater than the set voltage, then the LED will grow.

II. LITERATURE REVIEW

The fertilizer and pesticide detection process are not very commonly used in our country because of existing costly and difficult detection method. In the early days, for pesticide detection, one has to rely on a laboratory test. In the laboratory, a number of tests are conducted on a sample to identify existing materials on that particular sample. Nowadays scientists do chromatography and spectroscopy to identify chemical material on a sample [1]. The problem with this chemical test is it has to be done and individual with textual and practical knowledge. Therefore, these tests cannot be conducted by a farmer. Though there is some chemical detected method of pesticide whereas a few electrical approaches are noticed in this case. In the case of the chemical pesticide detection method, scientists preferred chromatography and spectroscopy as those are the most advanced and reliable detection method. There are some research works have been done to detection the pesticide level by an electrical method. In recent days, researchers have made a polymer coated signal IDT sensor to detect pesticide level researcher has used it in practical case but the problem with this sensor is there cost [2]. Moreover, these sensors are not still available for practical work. It is still in the developing stage. Another approach of testing soil fertility can be done by PH detection. But it does not give an idea about the pesticide level [3].

Using pesticide according to soil's need is not common in Bangladesh because of lack of pesticide detection. The scarcity of pesticide detector is because there has not been much work done in this sector. Therefore, for making a pesticide detector with demonstrating pesticide level required reviewing a large number of literature. As there was not any work directly related to this field so a large amount of modification was done to the available literature work in order to use it in our detection purpose. The basic theory was adopted from formalin detection process which worked on comparing resistance level of the sample solution and referred value of a different level of pesticide [4]. Furthermore, Electronic Nose for pesticide detection was reviewed to get an idea about present research work [5]. Effect of water PH on the chemical stability of pesticide was reviewed to get a more accurate output of our work [6].

III. METHODOLOGY

Conductivity is the ability to conduct current through any substance conductivity is directly connected to the resistivity of a substance. The low resistivity of a substance indicates a substance as highly conductive indicates the opposite. The resistivity of a substance can vary from time to time depending on many facts. For example, Resistivity of a copper wire varies with temperature increase and decrease [7].

The whole fertilizer, pesticide level detection process is working depending on the conductivity property of materials. If a pesticide is mixed with water the resistivity of the solution is decreased. This decrease depends on the amount of pesticide mixed in water. So in this project different voltage

level has been determined for a different level of pesticide. In the main circuit, a non-inverting amplifier is used to take voltage from an individual sample. Here a couple of comparators is used compare the voltage level of the sample solution and voltage level set for the different level of pesticide. Two LEDs are used to show the output of the corresponding comparator. This way using this method the presence of 30%, 50%, 70%, 100% pesticide in a sample can be determined. As two comparators are used in a circuit, as a result, the presence of any two levels of pesticide can be determined at a time.

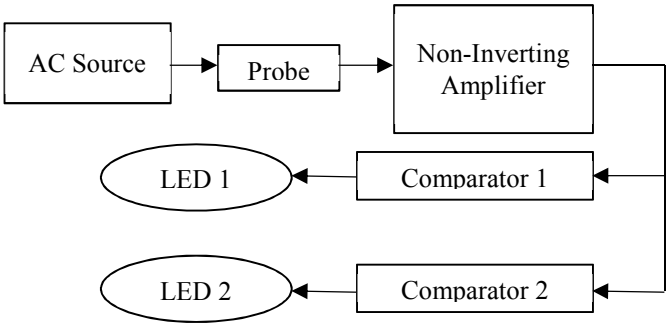


Fig. 1: Block diagram of pesticide detection

A. Working Methode

A set of data was taken for determining a different level of voltage for the different level of pesticide. Those voltage levels were set in the comparator for comparing sample voltage and default voltage set for the different level of pesticide. In comparator 1 a certain voltage level was set to detect the presence at any certain level of pesticide. In fig 2, Op-amp was used to measure the solution voltage gain using two probe method. In comparator 1, a certain voltage level was set to detect the presence at any certain level of pesticide. And in comparator 2, another voltage level was set for another level of pesticide. The comparator was used to show the result by comparing between the defaults voltages and sample voltage. Two LED was used to indicate the output result where one LED 1 shows one level of pesticide and another LED 2 shows the presence of another level of pesticide.

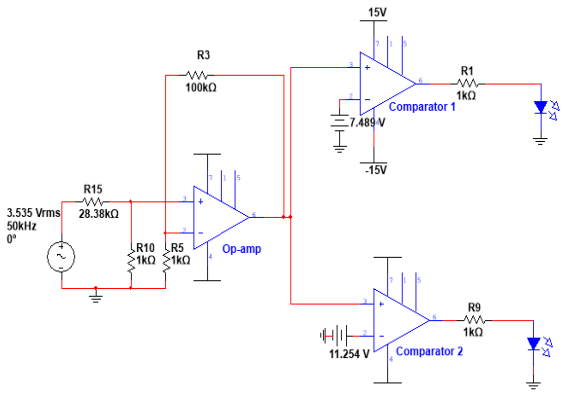


Fig. 2: Fertilizer and pesticide level detection circuit.

To make the solution, at first, TSP is mixed with a fixed amount of water. The amount of water is used here is 500ml. According to TSP's atomic molar mass 30%, 50%, 70%, 100% solution is made. A similar procedure is followed for pesticide (fungicide). And for insecticide 1 drop, 2 drops, 3 drops of sample are used for this detection process. The reaction of TSP and pesticide with water is given below:

$$Ca(H_2OPO_4) + H_2O \longrightarrow 2H_3PO_4 + CaO$$

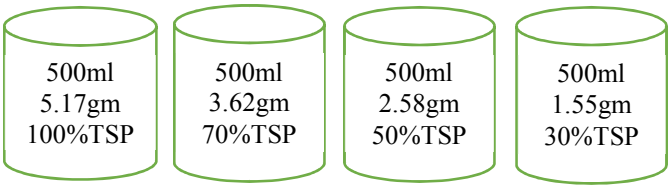


Fig 3: Reaction of TSP with water



Fig. 4: Reaction of pesticide (insecticide) with water

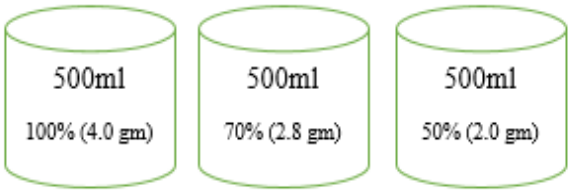


Fig. 5: Reaction of pesticide (fungicide) with water

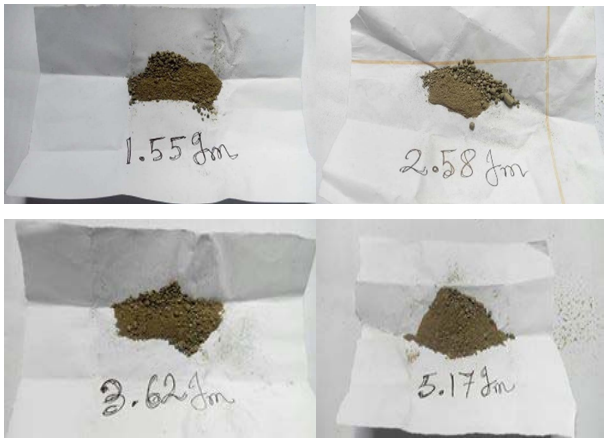


Fig. 6: Different amount of TSP according to their atomic masses.

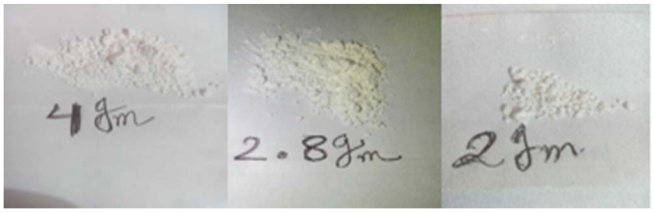


Fig. 7: Different amount of pesticide(fungicide) according to their atomic masses.



Fig. 8: Different amount of pesticide(fungicide) according to their atomic masses.

IV. Experiment Setup

A user-friendly and cost-effective circuit was used in this project for the purpose of detection. Basic electrical active devices like op-amp, comparator was used to finish the project. The AC source was applied to the solution. According to the solution resistance, the non-inverting op-amp will give an amplified output voltage. Then this output voltage was shorted to two comparators. Comparator-1 and Comparator-2 were used to detect 30%, 50%,70%, 100% TSP and pesticide respectively. Two DC voltages were assigned to the negative pin of each comparator. When the amplified AC voltage is more than the DC reference voltage the LED will glow. From the fig. 9, it can be inspected easily. In the main circuit, the op amp take gets a voltage gain from any kind of sample of pesticide using two probe method. From the op amp this voltage gain goes to the comparator 1 to compare the voltage gain with the default voltage set previously. If the sampled voltage is less than the set voltage for the presence of pesticide, then the LED 1 will not glow. That indicate the sample voltage level is lower than that of the presence of any level pesticide. On the contrary, if the LED does glow it indicated that the sample voltage is higher than the corresponding level and there is pesticide in the sample. This way four voltage levels are set for four pesticide level detection. The LED indicated the level of TSP or Pesticide present in a sample soil. In case of comparator 2 if the voltage gain of the sample is greater than the set voltage level than the LED 2 would glow indicating the pesticide level of corresponding circuit. As we are detecting different level of fertilizers and pesticides so different voltage levels were set in the comparators. If the voltage level of a sample is greater than the set voltage LED will glow. Two LED are used to show the output of corresponding comparator. Two comparators are used in a circuit, show the result by comparing between the default voltage and set voltage as a result presence of any two level of pesticide can be determine at a time. So we might have to test sample with all circuit to get accurate result.

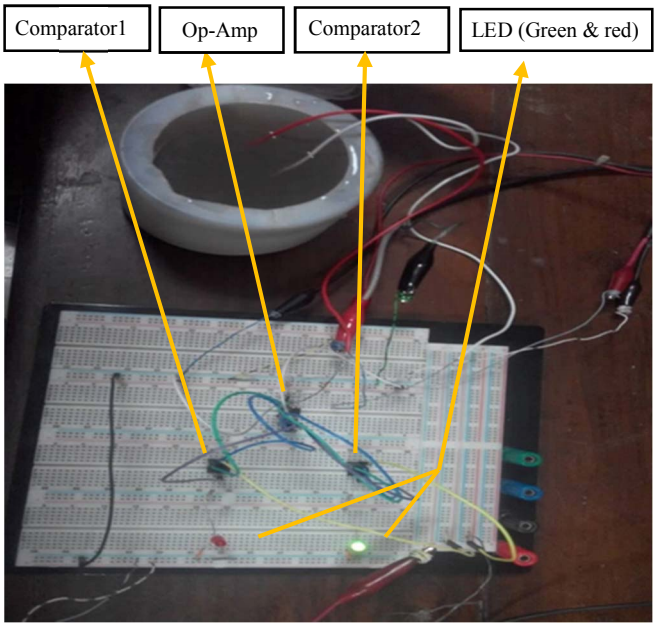


Fig. 9(a): Hardware implementation of the fertilizer and pesticide (detection circuit)

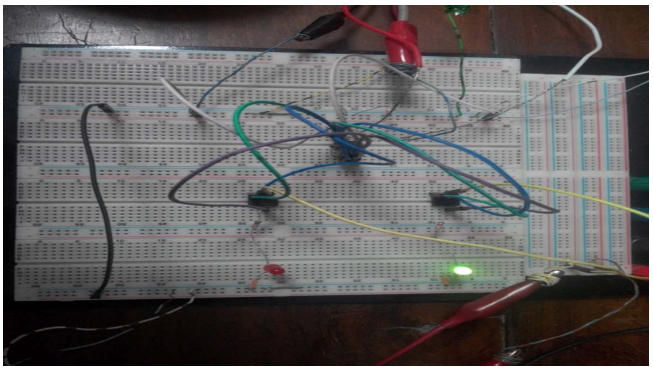


Fig. 9(b): Breadboard Setup (detection circuit)

Table 1: Components used to design the detection circuit

Components	Quantity
Op-amp (LM741)	4 pieces
Comparator (LM741)	16 pieces
Resistor	24 pieces
LED	8 pieces
Breadboard	4 pieces
TSP	1 kg
Urea	1 kg
Male to male jumper wire	40 pieces

Co-axial cylinder probe was used to detect the pesticide. Coaxial cylinder method is more accurate than two wire method.



Fig. 10: Co-axial cylinder probe (Lateral view).



Fig. 11: Co-axial cylinder probe (Top view).

V. Simulation Result

Before implementation of hardware part, simulation was done by National Instrumentation Multisim. The simulation shows that Led 1 is flickering and current is 2.513 mA and other 2 led is off so showing 0 A.

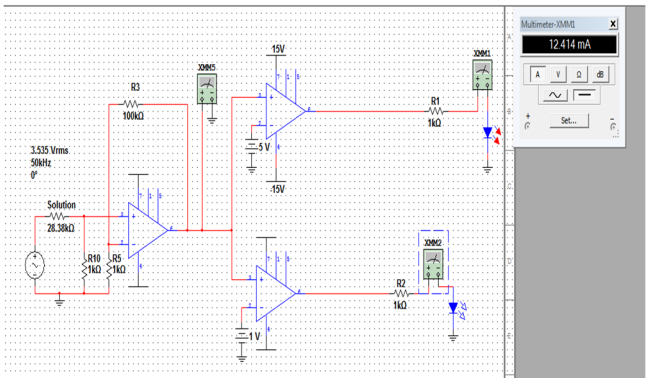


Fig. 12: Simulation for 30% level of the solution (TSP).

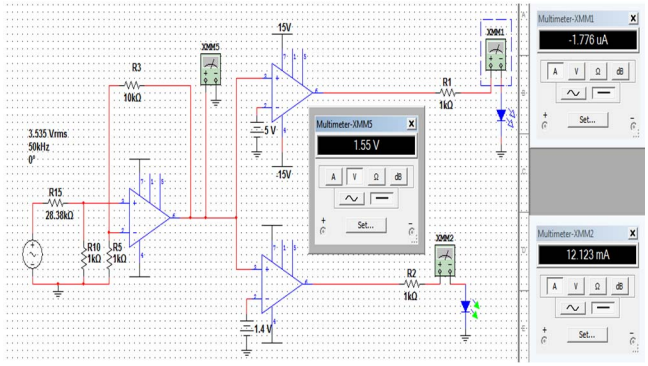


Fig. 13: Simulation for 50% level of the solution (TSP).

Fig. 12 and Fig. 13 are fertilizer solution (30% & 50%) simulation. Fertilizer solution has some resistance. For that solution resistance op-amp gives an output voltage. This output data is send to both comparators pin (3). Then the comparator which reference voltage is close to output voltage is changed by varying the power supply. And when the reference voltage matched with the output voltage of op-amp the corresponding comparator LED light turns ON.

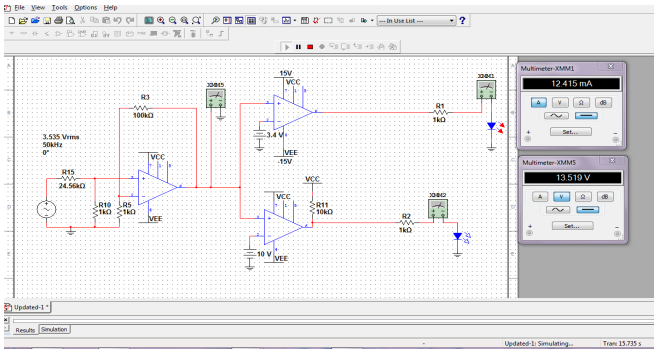


Fig. 14: Simulation for 1 drop (1ml) insecticide of the solution.

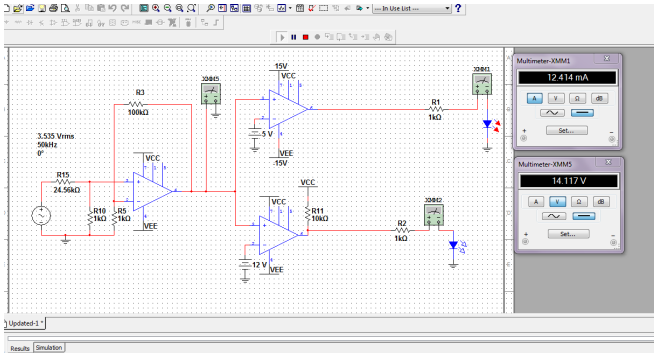


Fig. 15: Simulation for 2 drop (2ml) insecticide of the solution.

Fig. 14 and Fig. 15 are insecticide solution (1 drop & 2 drop) simulation. Insecticide solution has some resistance. For that solution resistance op-amp gives an output voltage. This output data is send to both comparators pin (3). Then the comparator which reference voltage is close to output voltage is changed by varying the power supply. And when the reference voltage matched with the output voltage of op-amp the corresponding comparator LED light turns ON.

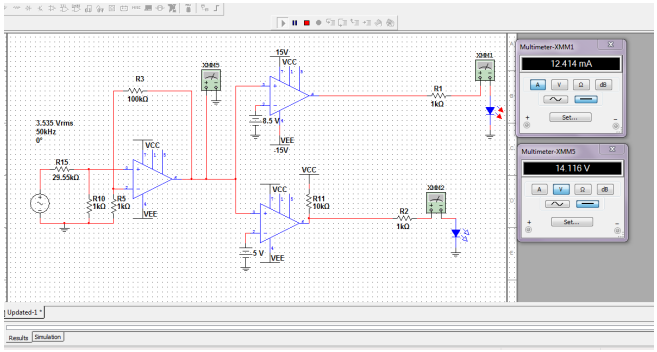


Fig. 16: Simulation for 50% (2gm) level of the solution (Fungicide).

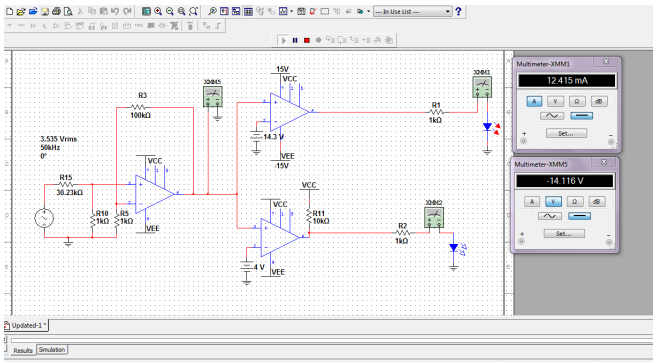


Fig. 17: Simulation for 100% (4gm) level of the solution (Fungicide).

Fig. 16 and Fig. 17 are fungicide solution (2gm & 4gm) simulation. These solution has some resistance (29.55 k-Ω & 30.23 k-Ω) . For that solution resistance op-amp gives an output voltage. This output data is send to both comparators pin (3). Then the comparator which reference voltage is close to output voltage is changed by varying the power supply. And when the reference voltage matched with the output voltage of op-amp the corresponding comparator LED light turns ON.

Table 2: Result of reference voltages

Vin	5V
The output voltage of the amplifier	2.5V
1st Comparator reference voltage (30%)	1.4V
2nd Comparator reference voltage (50%)	5V

In the case of a sample of 30% of a pesticide, the output voltage of op-amp was around 2.5V. As the reference voltage of 1st comparator was 1.4V which was less the output voltage of op-amp so the 1st LED was turned on. This 1st LED represents the presence of 30% of TSP pesticide in the soil. And the reference voltage of 2nd comparator was 5V which was for 50% of pesticide. It was higher than the voltage of op-amp which did not turn the 2nd LED. That means the presence of TSP pesticide in the soil is 30%. In the same way, we can also find the presence of a different level of TSP pesticides in the soil like 50%, 70%,100%.

Table 3: Referred voltages for a different level of solutions of TSP.

Level of solutions	Voltages (v)
30%	1.4 V
50%	5 V
70%	7 V
100%	11 V

The table given above is about the referred voltage of the different level of pesticide. One very important fact is that these referred voltages were measured from a solution of pesticide and water. If the sample was taken from field mixed with soil then the referred voltage seems to increase by 0.2~0.5 in each level. To get a better result it is observed that taking a sample of 5gm would be sufficient. The change in voltage level occurs due to the presence of other chemical compounds in the soil.

Table 4: Referred voltages for a different level of solutions of fungicide

Level of solutions (0.5 Liter water)	Reference Voltages (v)
1drop (1ml)	3.4V
2 drop (2ml)	5 V
3 drop (3ml)	10 V

Table 5: Referred voltages for a different level of solutions of insecticide

Level of solutions (0.5 Litter water)	Reference Voltages (v)
50%(2.0gm)	8.5V
70%(2.8gm)	9.5V
100%(4.0gm)	14.3 V

VI. Limitation & Future Enhancements

Every system has its limitations. This detection process has some limitations too such as it is only done for loamy soil. It can detect only two levels of TSP and pesticide at a time. In this process for each two level detection, one circuit is used. Presence of other chemical might affect the result.

As there has not been much work done in this domain in our country, so this would encourage new engineers to work in this domain. To take this work in the next stage what detection of all the level of pesticide can be done in one circuit. It would be challenging as using a couple of the number of the comparator and giving result through it would require using logic gate which leads to a challenge as logic gate do not work on AC voltage so a bridge rectifier has to be used in the process to convert the AC to DC. So using a more comparator would allow farmers to detect the different level of pesticide with one circuit.

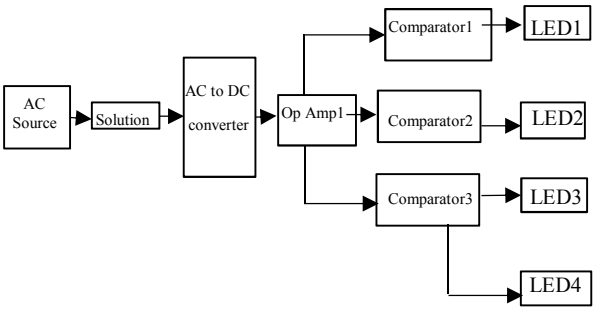


Fig. 18: Future proposed model (block diagram) which can detect 4 different levels of pesticide at a time.

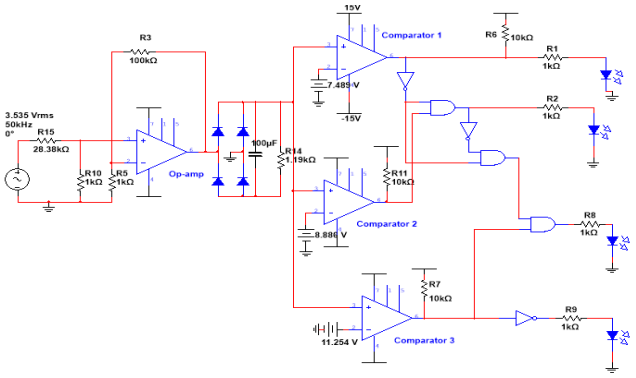


Fig. 19: Future proposed model circuit which can detect 4 different levels of pesticide at a time.

VII. Conclusion

This project was accomplished by detecting the amount of fertilizer presents in the given soil sample and pesticide in the water sample.However it not only indicates the presence of pesticide in sample but also it indicates the level of pesticide in the given sample. Conductivity property of TSP was used for the purpose of detection. This method of detection is very helpful for farmers and agricultural instructors as they can use this tool at a reasonable cost and use it easily. The wide availability of this project components is a major advantage in the case of manufacturing.

References

[1] P. Phil Taylor, "Testing Foods for Pesticides Using LC–MS/MS and SWATH Acquisition", The Column, 2017. [Online]. Available: http://www.chromatographyonline.com/testing-foods-pesticides-using-lc-msms-and-swath-acquisition. [Accessed: 04- Feb- 2019].

[2] Michael J. McCarthy, "Analysis of the Detection of Organophosphate Pesticides in Aqueous Solutions Using Polymer-Coated Single IDT Sensors", Master's Theses (2009), pp. 2-7. 2014.

[3] "pH meters | Hanna Instruments - Official UK Store", Hannainst.com, 2019. [Online]. Available: https://hannainst.com/. [Accessed: 04- Feb- 2019].

[4] H. M. Deer and R. Beard, "EFFECT OF WATER PH ON THE CHEMICAL STABILITY OF PESTICIDES", Extension.usu.edu, 2001. [Online].Available: https://extension.usu.edu/waterquality/files-ou/Agriculture-and-Water-Quality/Pest/AG_Pesticides_14.pdf. [Accessed: 04- Feb- 2019].

[5] Rashed Md. M. Hasan, A. N. Dewanjee and S. N. Shemul, “Design and Construction of a Formalin Detector using the Conductivity Property”, Electrical & Communication Engineering & Renewable Energy (ECERE), 2014, pp.071, 2014.

[6] F. Leccese, M. Cagnetti, S. Giarnetti, A. Pecora, L. Maiolo, E. De Francesco, R. Durovic, T. Dordevic, A. Tomaševic, V. Bursic, S. Tuti, I. Luisetto, P. Gabriele, “Analysis of Pesticides: a First Study Toward a Smart Analysis”, International Journal of Engineering, Science and Innovative Technology-IJESIT, Vol. 5, Issue 5, pp. 9-16, Nov. 2016.

[7] "Conductivity, Salinity & Total Dissolved Solids – Environmental Measurement System", Environmental Measurement Systems, 2019. [Online]. Available: https://www.fondriest.com/environmental-measurements/measurements/. [Accessed: 04- Feb- 2019].

Performance Analysis of Electrowetting on Dielectric (EWOD) Lens with Minimum Response Time and Optimized Damping Oscillation

Sumaiya Rahman
Department of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
sumaiyasifat18@gmail.com

Mst. Fateha Samad
Department of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
fatehaeteruet@gmail.com

***Abstract*—This paper represents the performance analysis of liquid lens formed on the basis of electowetting principle. Electrowetting on Dielectric (EWOD) lens facilitates the conversion of lens from convex to concave without manual movement. The lens is designed with optimized dielectric layers and minimum damping oscillations. To characterize the performance of the lens, response time is considered. The simulation is conducted using the dielectric layer of polyvinylidene fluoride (PVDF) and hydrophobic layer of Teflon AF 1600, and the droplet of distilled water. The simulated optimized response time is approximately 28 ms with minimum damping oscillations for the high input voltage of 150 V and low viscosity of insulating fluid (0.01 Pa-s).**

Keywords—Electrowetting on dielectric (EWOD), response time, dielectric layer, damping oscillation, hydrophobic layer.

I. INTRODUCTION

The electrowetting on dielectric (EWOD) refers to a device where the movement of polar droplet is occurred electrically. This technique is applicable in modern technology because of no need of machanical movement [1]. Besides it also provides a huge range of tunability as well as faster frame rates and negligible power consumption [1]. Therefore, it is a perfect technique for the tunable focal lenses and miniature cameras [2]. Initially it was used only for industrial purposes but recently it applies for optical display, lab-on-a-chip (LOC) [3], optical switches with high rejection ratio, microscopy, remote sensing and even for biological applications including polymerase chain reaction, DNA replications [1]. Miniature cameras based on this technology can easily change from narrow to wide angle just by altering the electrical input.

Since decades high-speed focusing technology has been aimed for the above mentioned applications. However, the focusing speed of conventional optical systems is limited by the slow response time and requirement of moving parts manually [4]. To overcome this, the phenomenon of electrowetting was first invented by Lippman in 1875 [5]. This invention was an alternative to conventional lens with improved features of no moving parts as well as faster response time. Thus EWOD allows one bit of further flexibility in lens design.

Wide ranges of research have been conducted on EWOD lens. For instance, Ouyang et al. [6] showed an application of electrowetting principle where the shape and location of the droplet was adjusted which consecutively tunes the focal length and position of the lens. As possessing qualities such as tunability of focal length upto 20% along with low actuation voltage and low power consumption, it can be used as a lens. This idea clarifies the legitimacy of using EWOD devices as lens. Now-a-days, the research effort for EWOD lens is focused on the usage of lower viscosity's insulating fluid, optimized response time, and minimum damping oscillation. Moreover, Berge presented an experimental setup for liquid lens technology where more emphasis was put on the selection of suitable insulating fluid for centering of liquid lens [7]. The importance of the selection of proper insulating fluid with lower possible viscosity for stabilizing optical axis was shown here. Additionally, Park et al. [8] demonstrated a practical model of planner liquid lens where the range of input voltage was 0-200 V and in accordance with this, the confined Si-oil changed the droplet's shape. Here the tunability of focal length from 12 mm to 22 mm with a 3 mm diameter of liquid lens was tested. In another work, Supekeret [9] used parylene HT and Teflon as the dielectric, and found the response time ranging from 115-500 ms with DC driving voltage of 0-28 V. The effect of DC driving voltage function on the response time of EWOD lenses was observed. Similarly, the response time for the almost same setup was found approximately 36.5 ms where parylene C, LiCl, Si-oil were used [10]. Thus, it is clear that the modification of the materials can change the result.

In this work, a EWOD lens is designed and simulated. A new dielectric material with different layer configuration of lens is used to improve the system performance. Though Nb_2O_5 or CEP, both have higher dielectric constant, they are not convenient for biological applications because of their toxic property. Thus, polyvinylidene fluoride (PVDF) is chosen as a dielectric. The EWOD lens made by PVDF can also provide lower response time.

Though, in previous work, a multi-layer dielectric coating with PVDF was used on EWOD device to reduce the actuation voltage [11], a PVDF layer along with a hydrophobic layer (Teflon AF 1600) is used here for a EWOD lens to change the droplet's actual shape as well as lens's focal length and magnification. Hence the response time is minimized and the oscillation is optimized. The paper is organized as follows: Section II shortly describes the theory of EWOD lenses and presents the proposed design approach. Section III represents the simulation results and discussion with a comparative study of the existing works, and finally, Section IV concludes the remarks.

II. Theory of Liquid Lens and Design

A. Basic Principle of EWOD Lens

An EWOD device consists of a substrate layer as base, electrode's array, a dielectric & a hydrophobic layer [12]. A conductive liquid is used as a droplet of precise volume and density on the electrode's surface. This surface is surrounded by either air or an immiscible media (e.g. Si-oil) to prevent the volatilization of the droplet. The droplet indicates a lens characteristics because it owns a rotational symmetry around its principal axis. Theoretically, the relationship between the focal length (f) and contact angle (θw) by [14]:

$$cos\theta_w = cos\theta_0 + \frac{\epsilon_0 \epsilon_{rd} \epsilon_{rh} V^2}{2\sigma(t_{d1} \epsilon_{rd} + t_{d2} \epsilon_{rh})} \quad (1)$$

Where, θ_0 = initial contact angle, θ_w = actuated contact angle, ϵ_{rd} = dielectric constant of dielectric layer, t_{d1} = thickness of dielectric layer, ϵ_{rh} = dielectric constant of hydrophobic layer, t_{d2} = thickness of hydrophobic layer, σ = surface tension at the interface between two fluids, and V = applied voltage. In addition, the relationship between the focal length (f) and contact angle (θ_w) of an EWOD lens is given by [13]:

$$f^3 = \frac{3V}{\pi(1 - cos\theta_w(2 - \cos^2\theta_w - cos\theta_w)(n_0 - n_g))} \quad (2)$$

Here, V = volume of the droplet, n_0 = refractive index of oil, and n_g = refractive index of droplet. The relation shows that the contact angle decreases with the increase of applied actuation voltage and as the contact angle decreases, focal length increases consecutively [15]. Moreover, increased focal length implies more magnification i.e., zoomed in view whereas decreased focal length implies zoomed out view [16].

In terms of adaptive liquid lens, response time is an important factor. It includes both rise time and falling/settling time. Basically response time is the time required for a generic system to react to a given input. The following equation shows a relationship of the response time during droplet actuation.

$$V(t) = V_0\left(1 - e^{-\frac{t}{\tau}}\right) \quad (3)$$

Where, t = response time, τ = input rise time. Besides, in physical system, stored energy is dissipated through damping. So, it is required to optimize the damping oscillation as well as response time to achieve an optimized performance of a liquid lens.

B. Design Procedure

The proposed model is designed using two different fluids, such as polar and non-polar. Though both fluids contain similar density, however the polar liquid acts as a liquid lens. Here, distilled water is used as polar liquid and Si-oil is used as non-polar liquid. A single droplet of the polar liquid is placed inside of a glass tube, which is coated with an Indium Tin Oxide (ITO) electrode, a PVDF dielectric, and a 1wt % Teflon AF 1600 hydrophobic coating. In this work, the interior wall of the glass tube is made with a 200 nm thick layer of ITO which is an optically transparent and works as an electrode. Since the main focus is on the effect of dielectrics of EWOD device, so more emphasis is given on the properties of the dielectric layers and their thicknesses. The thickness of dielectric layer and hydrophobic layer are selected as 3 μm and 1 μm respectively. Additionally, non-polar liquid is filled on the top of the droplet to prevent the volatilization. Then, a DC driving voltage is applied between the droplet and the electrode for actuating the liquid lens. In order to design and simulate the proposed liquid lens, FEM software package of COMSOL Multiphysics 5.3 is used. The numbers of features for creating geometric primitives are provided in this software package. The schematic diagram of the design including all the parameters of liquid lens is shown in Fig. 1.

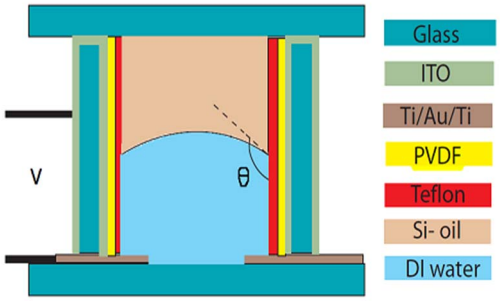


Fig.1. Schematic diagram of a EWOD liquid lens.

The solid structure of the design is obtained by taking the geometrical properties of 1.5 mm square and Bezier polygon (Fig. 2). Bezier polygon are basically a set of segment of polygon which has two rows divides as row 1 and row 2 (where, R = 1.55 mm and z = 0.76 mm). All properties of this polygon are chosen under the linear phase. Then the lens is filled by a polar liquid (distilled water) and a non-polar liquid (Si-oil) with the same density of 1000 kg/m3 at the temperature from 0 degree to 10 degree. The material properties for both polar and non-polar liquids are given as follows (Table I):



Fig.2. Geometrical structure of the liquid lens.

TABLE I. Material Properties of Liquid Lens Used in The Model.

Material	Properties	
	Density	***Dynamic Viscosity***
Conducting fluid (distilled water)	1000 [Kg/m^3]	1.57×10^{-3} [Pa-s]
Insulating fluid (Si-oil)	1000 [Kg/m^3]	muoil[a]

[a]muoil is varied from 0.01, 0.03, and 0.05[Pa-s].

Thus, for designing the model, firstly, the values of the parameters are defined in accordance with the geometry. The essential variable in this simulation is the contact angle, which varies directly with the applied voltage. The apparent contact

angle of the distilled water droplet on the Teflon coating is approximately 152 degree without any applied voltage. This angle is determined by the balance of forces at the contact points. In this work, the surface tension at the interface between two fluids is 0.049 N/m.

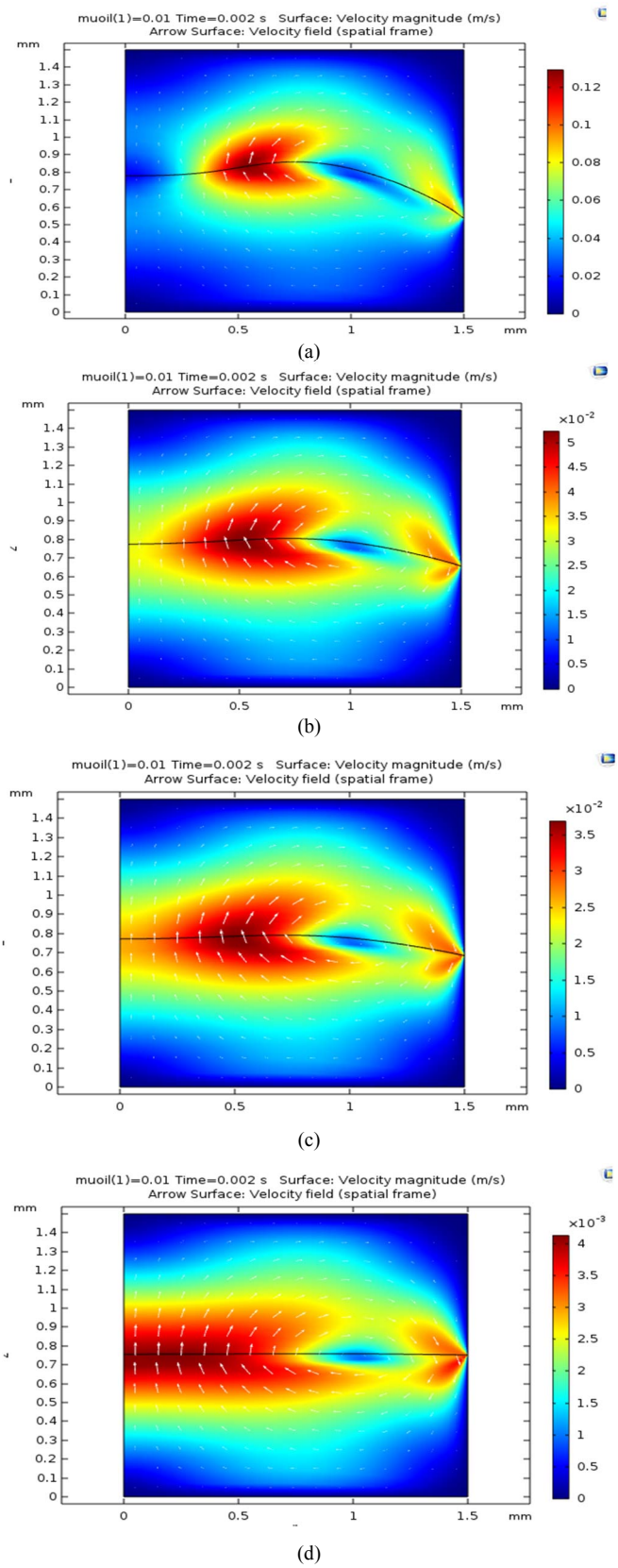
III. Simulation Result and Discussion

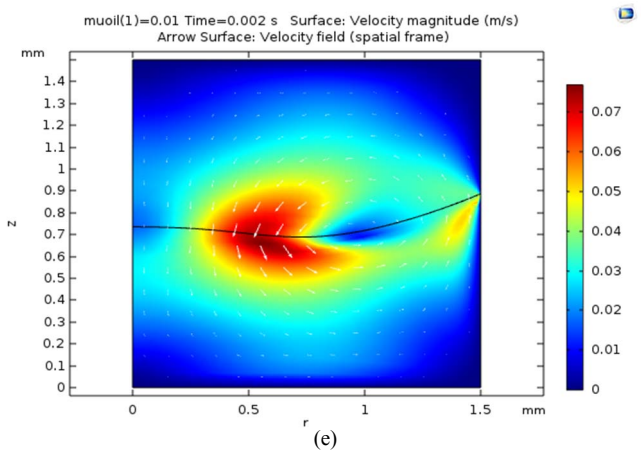
In this work, the FEM based COMSOL software is used in order to understand the temporal dynamics of our proposed lens. The simulation process is performed in two steps: (i) by adding a parametric sweep to the insulating fluid's viscosity and (ii) then defining the ranging of oscillation by adding a proper timing interval. It is desired to optimize the viscosity of the insulating fluid to achieve a faster response time. Thus, the range of viscosity is kept in between $10e^{-3}$, $30e^{-3}$, and $50e^{-3}$, whereas the time is taken in between 0, $1e^{-3}$, and $50e^{-2}$. The input voltage is switched from 0 V to 150 V. Therefore, the variation of the contact angles are achieved which is shown in Table II. It is clear that if the applied voltage increases, then the contact angle decreases, which justifies the theory. At higher voltage, the contact angle of the lens is significantly reduced. For example, at 120 V, the achieved contact angle is 91 degree or 1.59 rad.

TABLE II. VARIATION OF CONTACT ANGLES

Voltage [V]	0 V	30 V	60 V	90 V	100 V	120 V	150V
Contact angle [rad]	2.65	2.55	2.3	1.98	1.86	1.59	1.09

The total pressure withstood is found at a constant range as the water droplet in a Si-oil medium are confined into a chamber and driven together against a known pneumatic pressure until any advance forward. However, the velocity does not follow this consistency. The velocity of a fluid effectively describes everything about the motion of a fluid. At the initial condition, the internal force keeps the shape of the droplet fixed. When 0-60 V of the voltage is applied, then the velocity of fluid is observed as 0-0.12 ms^{-1}, indicating the convex shape of the droplet. Therefore, the droplet is started to change the shape and at the applied voltage of 90 V, the velocity of polar liquid is decreased which indicates a flat shape within a congested medium. The droplet exhibits a more flat shape with the reduction of velocity at 100 V. The absolute flat shape is obtained at 120 V of input potential. Furthermore, the highest velocity of 0.07 ms^{-1} is obtained at the applied voltage of 150 V. The velocity increases toward the center of the sealed chamber under the conditions of the laminar flow. This indicates the actuation of the droplet from flat to concave shape. All the numeric figures are obtained from COMSOL simulation. The fluid velocity as well as the change of curvature at different voltages is shown in Fig. 3.

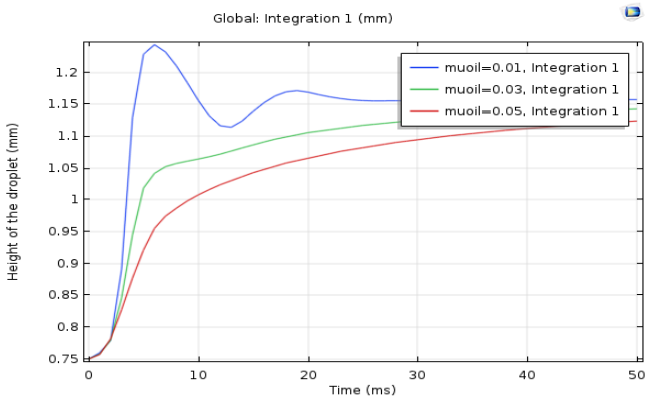




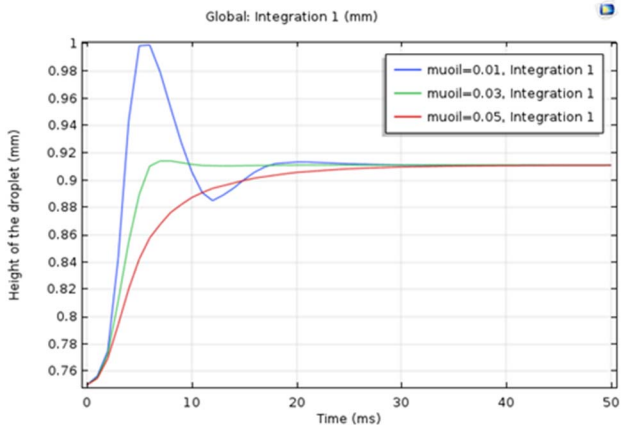
(e)

Fig. 3. The fluid velocity as well as the change of curvature at different voltages (a) 0 V, (b) 90 V, (c) 100 V, (d) 120 V, and (e) 150 V.

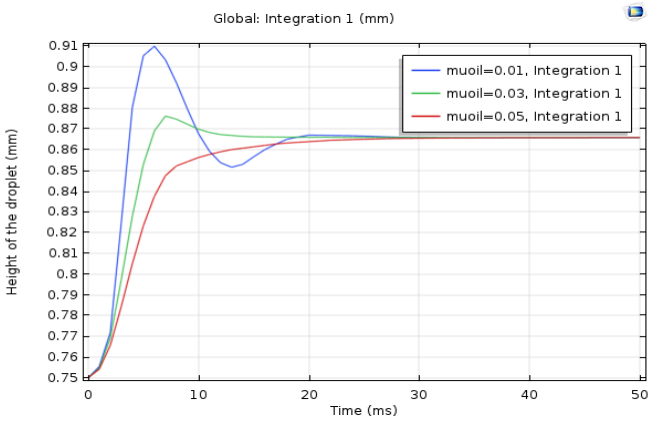
In addition, the plot of the response time at different actuation voltages is given in Fig. 4. It is clearly noticed that at minimum viscosity of 0.01 Pa-s, the achieved faster response time is approximately 36 ms. When the voltage is increased, the achieved response time is decreased. At the voltage of 150 V, the response time is obtained as 28 ms. The measurement of the response time is prepared with the help of Microsoft excel. Moreover, the damping oscillation at a particular voltage is be gathered by the output graphs obtained by the simulation software.



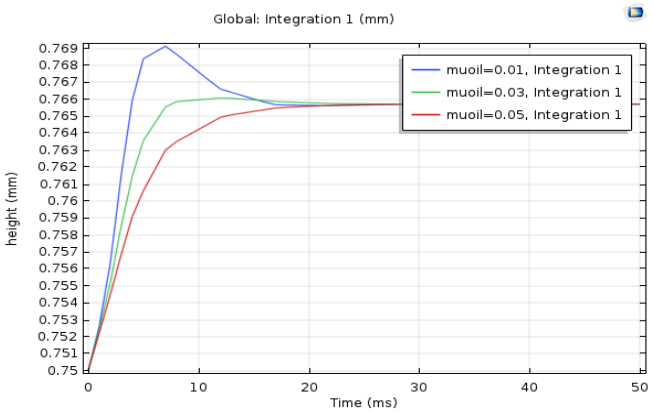
(a)



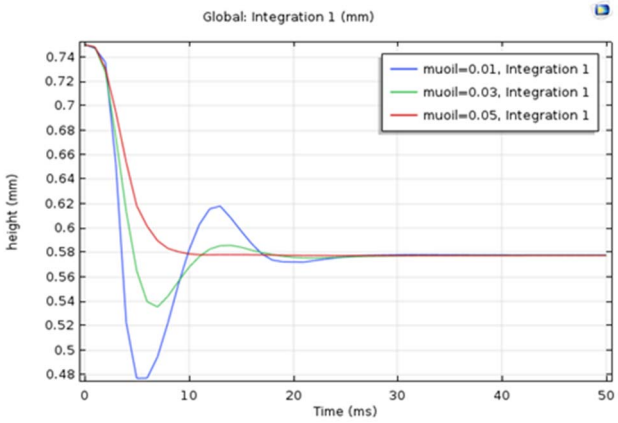
(b)



(c)



(d)



(e)

Fig. 4 : The response time at different actuation voltages of (a) 0 V, (b) 90 V, (d) 100 V, (d) 120 V, and (e) 150 V.

These simulation curves represent the change of the droplet height with respect to time from convex at 0 V (initial) to flat at 120 V to concave at 150 V with rising and settling time as well damping oscillations. Initially, the droplet exhibits a convex shape into the sealed chamber with a high damping loss. At 90-100 V, the droplet indicates a flat shape where the damping loss is partly minimized. Absolute flat shaped lens at 120 V is obtained with minimum damping oscillation to get distortion free imaging. And finally, at the highest applicable voltage of 150 V, the concave shape with a satisfactory amount of damping oscillation is achieved. Thus, the lens transformation from convex to flat to concave without any manual movement is attained using this principle.

This proposed design has better performance criteria in terms of response time. As in [9], the obtained response time was 500 ms at the maximum voltage of 28 V where the double layered (parylene HT and PTFE) EWOD lens was implemented. Moreover, the proposed design provides better performance than [10] too. The use of high dielectric material of PVDF provides a better response time of 28 ms, whereas the response time was 36.5 ms in [10]. Besides, our model has a good tolerance of voltage, it can withstand more than 150 V whereas the operation voltage limits was only 120 V in [10]. Thus the requirement of the lower possible response time is obtained by this proposed design.

IV. CONCLUSION

The principle of EWOD lens has been showed with the aid of simulation software COMSOL Multiphysics 5.3. System response time can be tuned from overdamped to underdamed with the reduced viscosity of Si-oil which indicates the change of the shape of liquid lens from concave to flat to convex without any mechanical parts. Moreover, our proposed model possesses a very good tolerance limit which allows it to perform at high actuation voltage with minimum response time and damping oscillation. Since the design of our EWOD lens is a simulated work, an experimental evaluation is needed to verify the simulation results and demonstrate the operation of the proposed EWOD lens. Although the fabrication procedure of the EWOD lens might be taken an additional cost, however this phenomenon can be very effective in the real time biological applications.

REFERENCES

[1] B. Berge and J. Peseux, "Variable focal lens controlled by an external voltage: An application of electrowetting," The European Physical Journal, vol. 3, pp. 159–163, October, 2000.

[2] R. D. Montoya, K. Underwood, S. Terrab, A. M. Watson, V. M. Bright, and J. T. Gopinath, "Large extinction ratio optical electrowetting shutter," Optics express, vol. 24, pp. 9660 , April 25,2016.

[3] J. Heikenfeld, N. Smith, M. Dhindsa, K. Zhou, M. Kilaru, L. Hou, J. Zhang, E. Kreit, and B. Raj, "Recent Progress in Arrayed Electrowetting Optics," Optics and photonics news, vol. 20 , January, 2009.

[4] G. Lippmann, "Relations entre les phénomenesélectriques et capillaires," Doctoral dissertation, Gauthier-Villars, 1875.

[5] H. Oku and M. Ishikawa, "High-speed liquid lens with 2 ms response and 80.3 nm root-mean-square wavefront error," Applied Physics Letters, 94, Japan, May 2009.

[6] F. Ouyang, J. Wu, M. Kang, R. Yue, and L. Liu, "Planar Variable-Focus Liquid Lens Based on Electrowetting on Dielectric," IEEE International Conference on Nano/Micro Engineered and Molecular Systems, September, 2012.

[7] B. Berge, "liquid lens technology: principle on electrowetting based lenses and application to imaging," 18th IEEE International Conference on Micro Electro Mechanical Systems, USA, February, 2005.

[8] J. Park, C. X. Liu, and J. W. Choi, "A Planar Liquid Lens Design Based on Electrowetting," IEEE SENSORS, vol. 8, pp. 4389-4396, 2007.

[9] O. D.Supekar, M. Zohrabi, J. T. Gopinath, and V. M. Bright, "Enhanced Response Time of Electrowetting Lenses with Shaped nput Voltage Functions," Biomedical Optical Express, vol. 8, issues 12, pp. 5412, 2017.

[10] T. Satyanarayana, G. S. A. K. Reddy, and V. S. P. Rajesh" Design and Simulation of Micro Electro Wetting Liquid Lens for Miniature Cameras," International Journal of Science and Research (IJSR), vol 2, pp 180-185, February, 2013.

[11] M. F. Samad and A. Z. Kouzani, "Reducing electrowetting-on-dielectric actuation voltage using a novel electrode shape and a multi-layer dielectric coating," Microsystem Technologies, accepted 20 July 2016.

[12] M. F. Samad, A. Z. Kouzani, and M. Samad, "Design and analysis of a low actuation voltage electrowetting-on-dielectric device," in 2013 ICME International Conference on Complex Medical Engineering, pp. 513-517, 2013.

[13] B. Jin, H. Ren and W. K. Choi," Dielectric liquid lens with chevron-patterned electrode," in Optics Express, vol. 25, pp. 26, December, 2017.

[14] Y. Park, S. Seu, P. Gruenberg and J. H. Lee, "Self-Centering Effect of a Thickness-Gradient Dielectric of an Electrowetting Liquid Lens," IEEE Photonics Technology Letters, vol. 25, pp. 6, March 15, 2013.

[15] J. H. Chen, C. A. Chen and W. H. Hsieh," Electrowetting Technique for Lens Fabrication,"IEEE International Conference on Mcchatronics,Taipei, Taiwan, July 10-12,2005.

[16] K. Noda, N. B. Khiem, Y. Tskei, and T. Takahata, "Focal length measurement of a varifocal liquid lens with capacitance detection," Applied Physics B, vol. 115, pp. 69–76, 2014.

Impact of Increased Quantum Confinement on Gate Capacitance and Threshold Voltage of P-Channel Junctionless Transistor

Md. Khairul Alam Titu[1], Sajid Hasan[1], Md. Mohsinur Rahman Adnan[1*]
[1] Department of Electrical and Electronic Engineering, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh.
*e-mail: adnanmohsin27@gmail.com

***Abstract*—In this work, by employing self-consistent solver we study Capacitance-Voltage (C-V) characteristics and the threshold voltage variation of p-type Rectangular Gate Junctionless Field Effect Transistor (RG-JLFET) and Double Gate Junctionless Field Effect Transistor (DG-JLFET). The self-consistent solver solves Schrodinger-Poisson equations with appropriate boundary conditions and takes wave function penetration and other quantum mechanical effects into account. We compare the effect of device parameter variation (such as channel thickness, oxide thickness and doping concentration) on threshold voltage control for both transistor structures. Physical explanations of the found results are also provided.**

***Index Terms*—Junctionless Field Effect Transistor, Threshold Voltage, Gate Capacitance, C-V curves, Self-Consistent Schrodinger- Poisson Solver, Multi-gate Transistors, Rectangular Gate, Double Gate.**

I. Introduction

The Junctionless Field Effect Transistor (JLFET) is a transistor without any physical p-n junction in the body region. The transistors have a heavily doped Silicon channel which in our case is p-type doping. The channel is covered by a thin Silicon Dioxide layer. In the outer most layers there is a Poly-silicon layer which is n-doped. This n-doped Poly-silicon acts as the gate of the transistor. We can see the cross-sectional view of the studied structures in Figure 1 & 2.

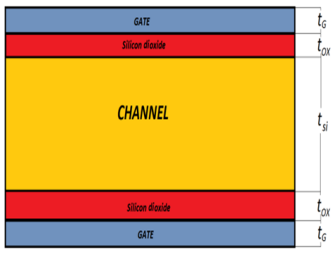


Fig. 1: A 1D cross-sectional view of Double Gate (DG)-JLFET. The cross-sectional view is perpendicular to the transport direction. For the devices, the channel thickness is 10 nm and the SiO_2 thickness is 2 nm.

The reason behind working on JLFET is that it has a huge advantage compared to regular MOSFETs against Short Channel Effects (SCEs). Also, the Multi-gate structures provide better performance against Short Channel Effects (SCEs)

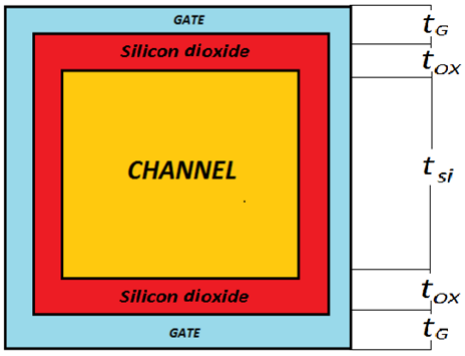


Fig. 2: A 2D cross-sectional view of Rectangular Gate (RG)-JLFET. The cross-sectional view is perpendicular to the transport direction. For the devices, the channel thickness is 10 nm and the SiO_2 thickness is 2 nm.

compared to single gate structure. Form [1] and [2], we know that the SCEs of a device is reduced if the gate length of the device is 6-10 times larger than its natural length (λ). Moreover, the natural length (λ) is related to its effective number of gates (n) in an inverse manner. If number of gates (n) is increased natural length (λ) will be decreased [1]. Then again with same Silicon thickness, the number of gates (n) is inversely proportional to the variation in threshold voltage ($\Delta V_{\rm th}$) [3]. In recent literature, there are already some works for p-doped JLFET on finding the threshold voltage using analytical models [4]. But most of them were done without considering any Quantum Mechanical (QM) effect. Previously, a comparative study on threshold voltage and Capacitance-Voltage (C-V) characteristics incorporating QM effects was done for the n-channel Double Gate (DG)-JLFET and Rectangular Gate (RG)-JLFET [5]. Because the electrostatic properties of n-MOSFET and p-MOSFET were found to be behaving differently [6], so it is necessary to do QM incorporated study for p-channel JLFET and compare the results with the study in [5]. In present work, a self-consistent solver incorporating QM effect like wave function penetration was used to evaluate the Capacitance-Voltage (C-V) characteristics for both DG-JLFET and RG-JLFET considering p-type doping. The threshold voltages have been obtained from the

Capacitance-Voltage (C-V) characteristics for both structures. And by varying the Silicon thickness (t_{si}), oxide thickness (t_{ox}) and doping concentration (N_a) of these structures, the variation in threshold voltage i.e. $\Delta V_{\rm th}$ have been obtained. The structure that had lower variability in threshold voltage ($\Delta V_{\rm th}$) was found out. Lower variability in threshold voltage means better control over channel potential. The device which had the lower $\Delta V_{\rm th}$ will have higher number of gates (n) and for that reason, it will also have lower natural length (λ) which eventually leads to better performance against Short Channel Effects (SCEs). In study [5] it was found that n-channel RG-JLFET was better in terms of SCE control over n-doped DG-JLFET. It is the target of this study to find out if the change in dopant type (p-doping) results in similar or different variation for increased confinement of JLFET device.

II. SIMULATOR MODEL

A coupled Schrodinger-Poisson equation was solved in order to compare the properties of both structures by self-consistent analysis [7]. To include the wave function penetration effects at the Si/SiO_2 interface open boundary condition had been used. In the case of DG-JLFET, the gates are placed on two sides of the channel. As a result, 1D Schrodinger-Poisson equation was solved for DG-JLFET. But for RG-JLFET, the gates are all around the channel (four sides) and because of that 2D Schrodinger-Poisson equation was solved for RG-JLFET.

A. Self-Consistent Analysis:

The Poisson's equation, as we all know, is,

$$\nabla(\varepsilon \nabla V) = -\rho \tag{1}$$

Here V is electrostatic potential, ρ = q[N_d-N_a+p-n] is charge density and ε is the permittivity of the substrate. N_d, N_a, p and n are donor density, acceptor density, hole concentration and electron concentration respectively. For the depletion region of p-type, ρ = q[-N_a + p] and if it is not depleted then, ρ = qp. Here, q represents charge of a hole.

The effective mass approximation can be found by the general form of Schrodingers equation and the equation is,

$$\left(\frac{-\hbar^2}{2m^*_{ds}}\nabla^2 + qV(r)\right)\psi(r) = E_j\psi_j(r) \tag{2}$$

Here, m^*_{ds} is the effective mass of holes. E_j and ψ_j are respectively the maximum energy and corresponding wave function of the j-th valance sub-band. Using (2) we determine the normalized ψ_j. Then in the j-th sub-band, the hole concentration is given by,

$$p_j = |\psi_j|^2 \int_{-\infty}^{E_j} g_j(E)[1 - f_{FD}]dE \tag{3}$$

Here, g_j(E) is the j-th sub-band Density-of-States (DOS) for valance bands and f_{FD}(E) is the Fermi-Dirac distribution function.

B. Finding the Capacitance and the Threshold Voltage:

The self-consistent analysis is converged and the coupled Schrodinger-Poisson is solved for the overall charge density. The equations are

In case of DG-JLFET :

$$Q = q\left[\sum\nolimits_j \int_0^{t_d} p_j dx - N_a t_d\right] \tag{4}$$

In case of RG-JLFET:

$$Q = q\left[\sum\nolimits_j \iint_0^{A_d} p_j(x,y)dxdy - N_a A_d\right] \tag{5}$$

Here, N_a is acceptor concentration, t_d is the depletion width and A_d is the depletion area. We calculate the gate capacitance per unit area (DG-JLFET) and per unit length (RG-JLFET) and it is expressed as,

$$C = \frac{dQ}{dV_{gate}} \tag{6}$$

To determine the threshold voltage we have to draw a suitable tangent at the partial depletion region of the C-V curve and then extend it toward the gate voltage (V_{gate}) axis [8]. At a certain point, the tangent intercepts the gate voltage axis and this is the voltage where charge contribution from free holes starts to dominate (figure 3). And this point is used to determine the threshold voltage of both structures [8].

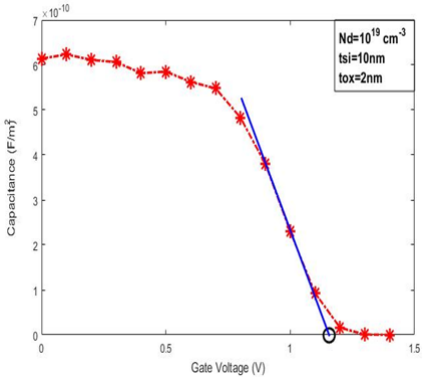
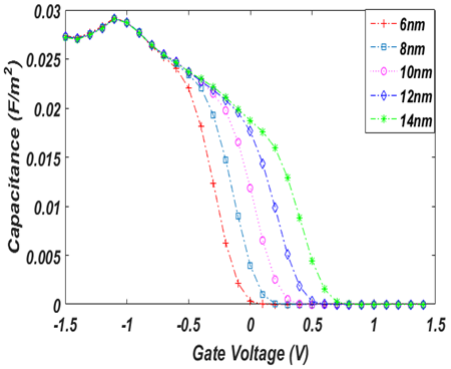


Fig. 3: Threshold voltage (V_{th}) extraction from C-V curve.

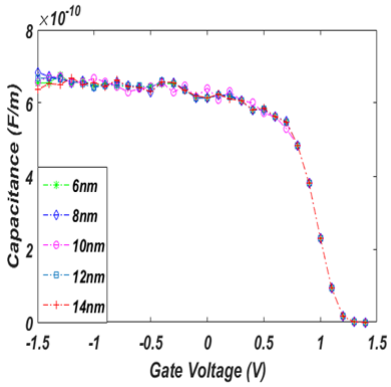
III. RESULT AND DISCUSSION

From our developed simulator, we have established C-V characteristics curves for both RG-JLFET and DG-JLFET. Figure 4 shows how the C-V characteristics of both structures are being changed by varying channel thickness (t_{si}). All three regions of operation of both structures have been shown in the figure. We get the full depletion region where the gate capacitance (C) is zero. When the mobile holes begin to gather then the capacitance (C) starts to rise. When it crosses a minimum value, then a channel is created. The partial depletion region can be found when the C is rising steadily to its maximum. At the Flat Band voltage V_{FB}, C reaches its maximum value for DG-JLFET $\left(C_{max} = \frac{2\varepsilon_{ox}}{t_{ox}}\right)$.

From the figure, we see that the maximum value does not depend on the Silicon thickness after a certain point. But in case of RG-JLFET, the capacitance rises steadily to maximum after full depletion. At V_{FB}, C reaches its maximum value $\left(C_{max} = \frac{\varepsilon_{ox} Perimeter_{gate}}{t_{ox}} = 4\varepsilon_{ox}\frac{t_{si}}{t_{ox}}\right)$ and in the Accumulation region it remains there (or very slowly increases). From the figure we can see that the capacitance (C) in DG is in 10^{-2} range but for RG-JLFET it gets to 10^{-10} as the term is multiplied by t_{si} where t_{si} is in the 10^{-9} range. We can see from the fig. 4(b) the maximum value of C changes for RG-JLFET as we vary the Silicon thickness, but this variation is very small compared to n-doped case from study [5].



(a)



(b)

Fig. 4: C-V characteristic varying Silicon thickness (t_{si}) of (a) DG-JLFET and (b) RG-JLFET.

From fig. 5 for DG-JLFET, we see that as the t_{si} increases the threshold voltage also increases. There will be more acceptor charge in the channel with increased thickness. To create a channel at the center, more and more positive voltage is required at the gate to attract that extra charge to the surface as p-channel dopant acceptors are negative. But for the RG-JLFET, fig. 5 shows much smaller V_{th} variation than DG-JLFET as t_{si} varies. For RG-JLFET, the effect of varying channel thickness on threshold voltage is insignificant due to better control from gates on all side i.e. increased confinement.

Figure 6(a) shows the C-V characteristic of DG-JLFET, varying oxide thickness. From the figure, we can see that there are different maximum value for different oxide thickness for

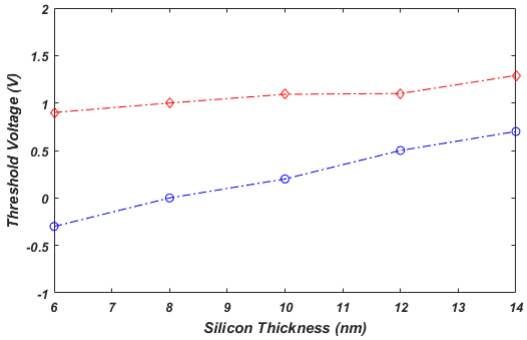
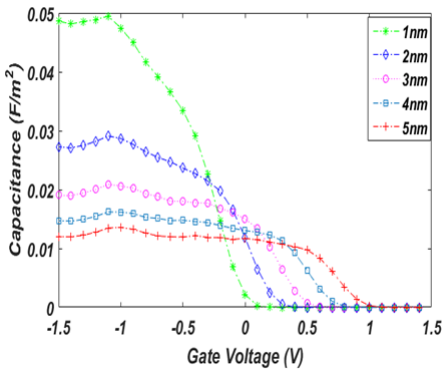
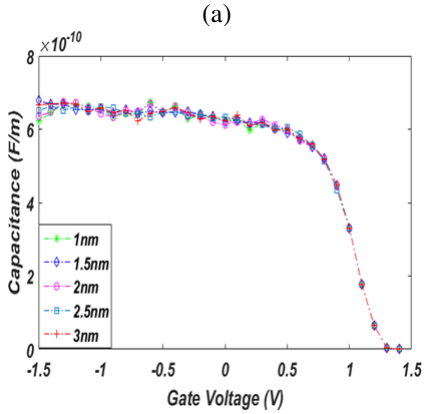


Fig. 5: The threshold voltage variation ($\Delta V_{\rm th}$) vs Silicon thickness for RG-JLFET (Red ◇) and DG-JLFET (Blue ○).



(a)



(b)

Fig. 6: C-V characteristic varying oxide thickness (t_{ox}) of (a) DG-JLFET and (b) RG-JLFET.

DG-JLFET. From figure 6(b) a similar scenario can be seen for RG-JLFET. The final value varied with varying oxide thickness. From figure 7, it can be seen that the V_{th} becomes more positive as the t_{ox} increases. This happens because the voltage drop across the oxide increases with the increase of t_{ox}. More positive V_{gate} is needed to create a channel. Here, in fig. 7, RG-JLFET again shows much better V_{th} control than DG-JLFET as t_{ox} varies.

But from fig. 8, it is noticeable that the final values of gate capacitance for both DG-JLFET and RG-JLFET remained almost same while N_a was varied.

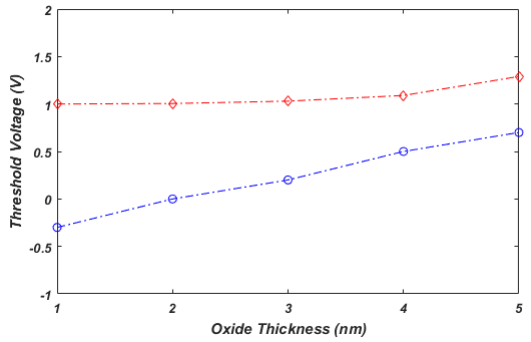
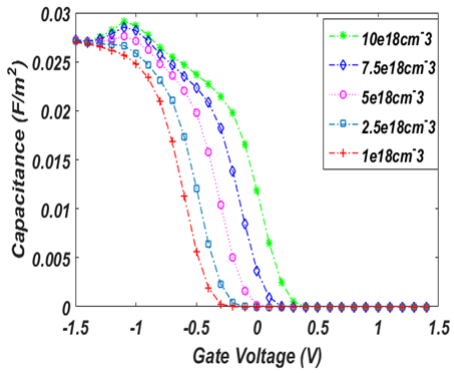
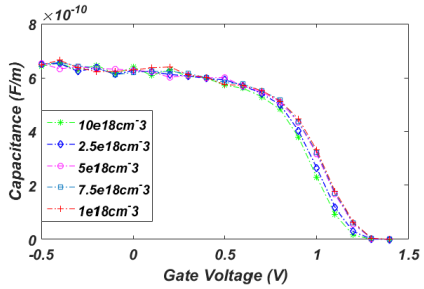


Fig. 7: The threshold voltage variation ($\Delta V_{\rm th}$) vs oxide thickness for RG-JLFET (Red ◇) and DG-JLFET (Blue ○).



(a)



(b)

Fig. 8: C-V characteristic varying doping concentration(N_A) of (a) DG-JLFET and (b) RG-JLFET.

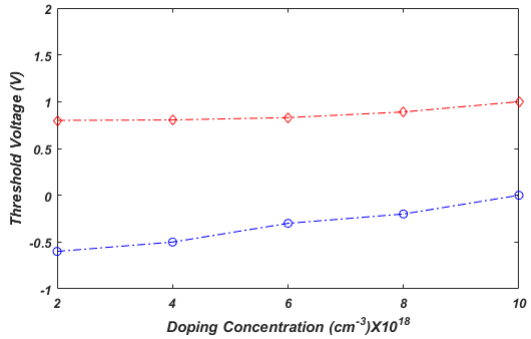


Fig. 9: The threshold voltage variation ($\Delta V_{\rm th}$) vs doping concentration for RG-JLFET (Red ◇) and DG-JLFET (Blue ○).

For each case, the V_{th} is calculated by the procedure which was stated in section II. But in fig. 9, it is clear that with the increase of N_a, V_{th} also increases. In this case, more positive voltage is needed at the gate to create a channel as there is more negative acceptor charge in the channel to be attracted so that the free holes can conduct in the center of the bulk. This is why V_{th} increases with the increase of N_a. Again RG-JLFET shows much better V_{th} control than DG-JLFET as N_a varies. Figure 5, 7 and 9 show that the variation in threshold voltage (ΔV_{th}) is greater in DG-JLFET than RG-JLFET with varying t_{si}, t_{ox}, and N_a. This is a confirmation of V_{th} being more sensitive in DG-JLFET than RG-JLFET with the change in t_{si}, t_{ox} and N_a. It means that DG-JLFET will have lesser control over channel potential than RG-JLFET. And the decrease in ΔV_{th} also proves that the natural length (λ) is lower and number of gates (n) is higher (which is obvious) in RG-JLFET then DG-JLFET which proves that RG-JLFET will be more efficient against Short Channel Effects (SCEs). All these observations are in line with the observations presented in study [5].

However, from figure 5, 7 and 9, it is evident that p-type RGJLFET has higher positive V_{th} then p-type DGJLFET. The reason is, to accumulate the free holes from the source and create a channel between drain and source at the body center, DG-JLFET has a one dimensional control (two-sided gate) whereas RG-JLFET has two dimensional control (four-sided gate). This is why to obtain the same amount of hole at the channel center, DG-JLFET needs more negative voltage at gate than RG-JLFET.

So higher quantum confinement means better control both in terms of accumulating holes at channel center and attracting acceptor dopant charges away from center towards the surface.

IV. Conclusion

For both JLFET devices structure threshold voltage (V_{th}) was calculated and compared by varying different parameters as Silicon thickness (t_{si}), oxide thickness (t_{ox}) and doping concentration (N_a). In that search, we found that with an increase of the given parameters, the V_{th} increases. In case of RG-JLFET, we show that threshold voltage variation ($\Delta V_{\rm th}$) is less than that of DG-JLFET. From this finding, we can conclude that RG-JLFET will perform better against Short Channel Effects (SCEs) and will have a better control over channel potential compared to DG-JLFET. Also we found that having better confinement resulted in a reduced need of negative gate voltage to achieve threshold condition for RG-JLFET compared to DG-JLFET. Thus, quantum confinement made controlling the occurrence of threshold condition easier for p-JLFET, similar to n-JLFET.

References

[1] Yan R., A. Ourmazd, Lee K., Scaling the Si MOSFET: From Bulk to SOI to Bulk. *IEEE Transactions on Electron Devices*, Vol. 39, No. 7, pp. 1704-10, Jul 1992.

[2] Suzuki K., Tanaka T., Tosaka Y., Horie H., Arimoto Y, Scaling theory for double-gate SOI MOSFETs., *IEEE Transactions on Electron Devices*, Vol. 40, Issue. 12, pp.2326- 2329, Dec 1993.
[3] Lee C., Yun S., Yu C., Park J., J. Colinge, Device design guidelines for nano-scale MuGFETs, *Solid State Electronics*, Vol. 51, No. 2, pp. 505-510, March 2007.
[4] M.M.R.Adnan and Q.D.M.Khosru, "A Simple Analytical Model of Threshold Voltage for P-Channel Double Gate Junctionless Transistor", Proceedings of IEEE International Conference on Electron Devices and Solid-State Circuits, 2018.
[5] N. Tasneem, M.M.R. Adnan, M.S.B. Hafiz and Q.D.M. Khosru, " Comparative study of quantum mechanical capacitance voltage characteristics and threshold voltage of two different structure of junction less nano wire transistor", Proceedings of IEEE Region 10 Conference, TENCON, 2016.
[6] A. Haque and M.Z. Kauser, "A Comparison of Wave Function Penetration effects on Gate Capacitance in Deep submicron n-and p-MOSFET.", *IEEE Transaction on Electron Devices*, Vol. 49, No. 9, September 2002.
[7] Stern F., Self Consistent results for n-type inversion layers, Vol.5 Issue 12, pp 4891-99, June 1972.
[8] S. U. Z. Khan, M. S. Hossain, F. U. Rahman, R. Zaman, M. O. Hossen and Q. D. M. Khosru, Impact of high-k gate dielectric and other physical parameters on the electrostatics and threshold voltage of long channel gate-all-around nanowire transistor., International Journal of Numerical Model. , Vol. 28, No. 4, Sep 2014.

Design and Analysis of Dual-Core *D*-Shape Photonic Crystal Fiber Biosensor

Mohammad Al Mahfuz
Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
almahfuz.ete.ruet@gmail.com

Md. Nazmus Sakib
Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
sakib.ruet12ete@gmail.com

Mohammad Hasibul Hasan Hasib
Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
hasib.eee.ruet@gmail.com

Akash Bosu
Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
akash.ruet13@gmail.com

Md. Roniuzzaman
Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
roni4033@gmail.com

***Abstract*— In this paper, a simple D-shaped photonic crystal fiber (PCF) plasmonic biosensor is proposed and analyzed. The sensor performance is investigated by applying a mode solver based finite element method (FEM). The sensing layer and plasmonic metal (gold) layer both are deposited on the external upper plane surface for comfortable and easy fabrication. The sensor exhibits the maximum sensitivity of 8000 nm/RIU and 522 RIU^{-1}, and resolution of 1.25×10^{-5} RIU and 1.92×10^{-5} RIU using the wavelength and amplitude interrogation method, respectively, in the detection range between 1.45 and 1.48. Owing to the simple design and highly sensitive characteristics, the proposed biosensor can be precisely applicable for biomolecular analytes detection.**

Keywords— Surface plasmon resonance, Finite element method, D-shape Biosensor, Sensitivity.

I. Introduction

The devices that are associated with the detection of several analyte concentrations are well known as the biosensor. In optical sensing applications, surface plasmon resonance (SPR) bio-sensors are highly recommended. These sensors are most preferable and have been extensively applied to applications such as pharmaceutical diagnosis, antigen-antibody interaction, bio-imaging, food quality measurement, environmental observation, and so on [1]. In earlier SPR techniques, prism coupling based Kretschmann configuration with thin metal layer was introduced as the SPR sensors [2]. Because of bulkiness, excessive cost of optical and mechanical components, remote sensing limitations, and miniaturization roughness these types of SPR sensors are not preferable. To minimize these problems, in 1993 a fiber-optic based SPR sensor was proposed by Jorgenson [3]. The working principle of the PCF-SPR sensor described by the mutual interaction of surface electron and evanescent field. At a specific wavelength, the real part of the effective refractive index (n_{eff}) of fundamental core mode and surface plasmon polariton (SPP) mode becomes similar. Due to the effect of core-cladding evanescent field, emission of the electron occurs from the metal surface. As a result, the surface plasmon wave is created. Gold is preferred as the most attractive plasmonic material in SPR sensing [4]. Silver is another plasmonic metal [5] which faced oxidation problem. To measure the performance, it should be considered that the sensor shows more easy fabrication and highly sensitive characteristics. Huang proposed a *D*-shape SPR-PCF showing wavelength sensitivity of 6000 nm/RIU [6]. Therefore, excessive surface polishing increases fabrication difficulties. For external sensing purpose, Chakma *et al.* presented a sensor which showed the maximum amplitude and wavelength sensitivity of 318 RIU^{-1} and 9000 nm/RIU [7]. However, here the amplitude sensitivity is not so high in this sensor. Dual-core *D*-shape photonic crystal fiber has many excellent characteristics like flexible to design, easy to demonstrate, and light can be smoothly coupling with the metal surface.

In this paper, we proposed a highly sensitive SPR-PCF biosensor based on the gold coating on the upper planar surface. By applying the finite element method (FEM), numerical performances are investigated concerning wavelength sensitivity, resolution, amplitude sensitivity, the figure of merit (FOM), linearity, respectively. Varying structural parameters the sensing performances are also optimized.

II. Structure Design and Theoretical Basis

The schematic cross-sectional view of the proposed *D*-shaped PCF sensor is illustrated in Fig. 1. The PCF lattice designed with three hexagonal rings consisting of all circular air-holes. Two air holes are removed from the first ring to form a dual core. Most of the prior dual-core SPR sensors are demonstrated with circular PCF. In our design, we have efficiently applied the dual-core concept for *D*-shape PCF structure. Four air holes (two from the first ring and two from the second ring) are scaled down for the minimum gap between the sensing medium and core. Therefore, surface plasmons can interact with the metal surface easily.

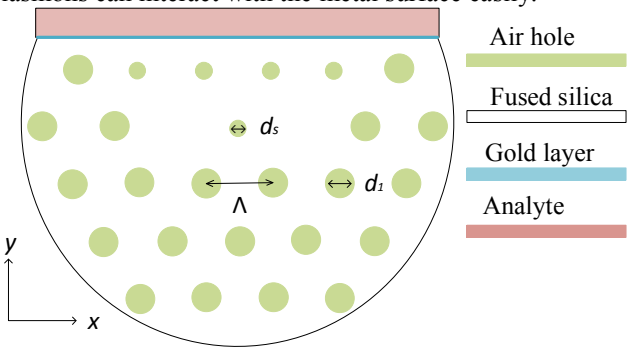


Fig. 1: Two dimentional view of the *D*-shaped PCF sensor with Λ = 1.9 µm, d_s = 0.25Λ, d_1 = 0.5Λ, and t_g=40 nm.

To form the *D*-shape, the top portion of the hexagonal PCF air-holes are removed. In this design, two adjacent air-holes center to center distance are symbolized as pitch (Λ), the diameter of smaller air holes is denoted as d_s, and d_1 denotes the diameter of larger air holes. Here, the parameters are fixed as, Λ=1.9 µm, d_s = 0.25Λ, and d_1 = 0.5Λ. The gold layer with the thickness (t_g=40 nm) is performed as the plasmonic material which placed at the plane surface of the sensor using chemical vapour deposition (CVD) method. Fused silica is selected as the background material. The Sellmier equation is used to calculate the complex refractive index (RI) of fused silica which is given by [7]. The dielectric constant of gold is obtained by the Drude-Lorentz model [7]. An analyte layer thickness of 1.5 µm was introduced during the simulation. The mesh parameters for the structure are introduced below. The structure contains total triangular elements of 17750 and mesh area of 97.76 µm2. Total vertex elements of 113 and edge elements of 1486. The average element quality of 0.9626 and the minimum element quality of 0.7694.

The modal analysis is carried out in *XY*-plane while the light is propagated in the *Z*-direction. *Y*-polarized propagation mode shows comparatively better light propagation intensity upon the gold surface. Hence, the evanescent incoming light can easily penetrate through the gold layer. So, we have considered *y*-component fundamental core mode for overall simulation purpose. The electric field, as well as resonance property, is given in fig. 2. Here, analyte RI is of 1.48. The real part of the RI of the *y*-polarized mode and SPP mode coincides at a wavelength of 1.30 µm. This wavelength is called the resonance wavelength. At that condition, we observe a sharp resonance peak that indicates the maximum energy transfers from fiber core to the plane surface.

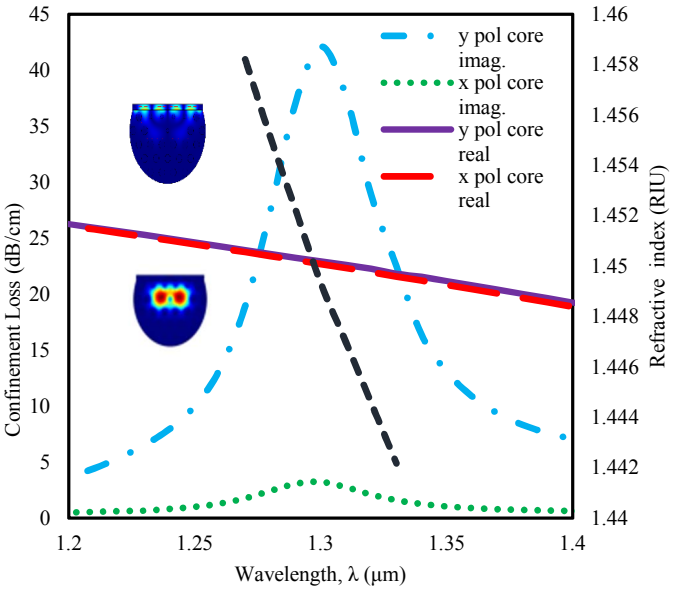


Fig. 2: Electric field distribution and dispersion relation of core mode and SPP mode with n_a = 1.48 and t_g = 40 nm.

Before assessing the characteristics of the PCF from a simulation, we checked fine and finer mesh numerical methods for to be sure about the precision which is shown in Fig. 3. From this figure, it is observable that the real effective refractive index for fine mesh element and finer mesh element are almost the same for a specific analyte RI. Therefore, we have considered finer mesh elements in our simulation to map the smaller air holes accurately.

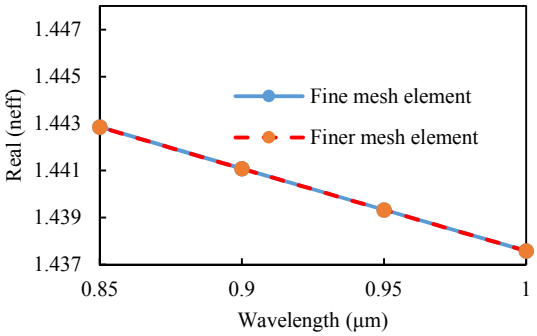


Fig. 3: Simulation for a different numerical method with fine and finer mesh elements.

III. SIMULATION AND PERFORMANCE ANALYSIS

A. Analyte variation and sensitivity

Confinement loss plays a significant role in sensing performance calculation which is obtained from the following equation [5]:

$$\alpha(dB/cm) = 8.686 \times k_o \, \mathrm{Im}(n_{eff}) \times 10^4 \qquad (1)$$

where *Im* (n_{eff}) is the imaginary part of the core mode index, k_0=2π/λ is wave number where λ is operating wavelength The variation of analyte RI profoundly influences the real part of the effective refractive index (n_{eff}).

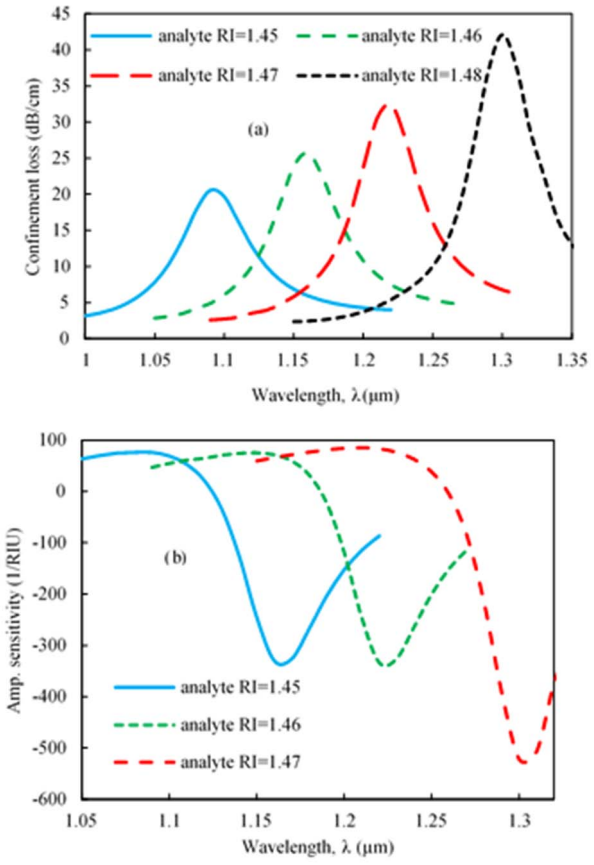


Fig. 4: (a) Analyte RI variation and loss from 1.45 to 1.48 and (b) amplitude sensitivity with Λ = 1.9 µm, d_s = 0.25Λ, d_1 = 0.5Λ, and t_g = 40 nm.

An unknown analyte can be easily detected with the small variation of RI with shifting of resonance point towards the higher wavelengths Fig. 4(a) shows resonance peak wavelength shift due to the variation of analyte RI from 1.45 to 1.48. For changing the value of n_a from 1.45 to 1.46, 1.46 to 1.47, and 1.47 to 1.48, the resonance wavelength is shifted from 1.08-1.15 µm, 1.15-1.23 µm, and 1.23-1.3 µm, respectively. Minimum confinement loss value of 20 dB/cm was found at 1.08 µm for analyte RI value of 1.45, and highest peak loss was observed at 1.30 µm is 45 dB/cm with the analyte RI of 1.48.

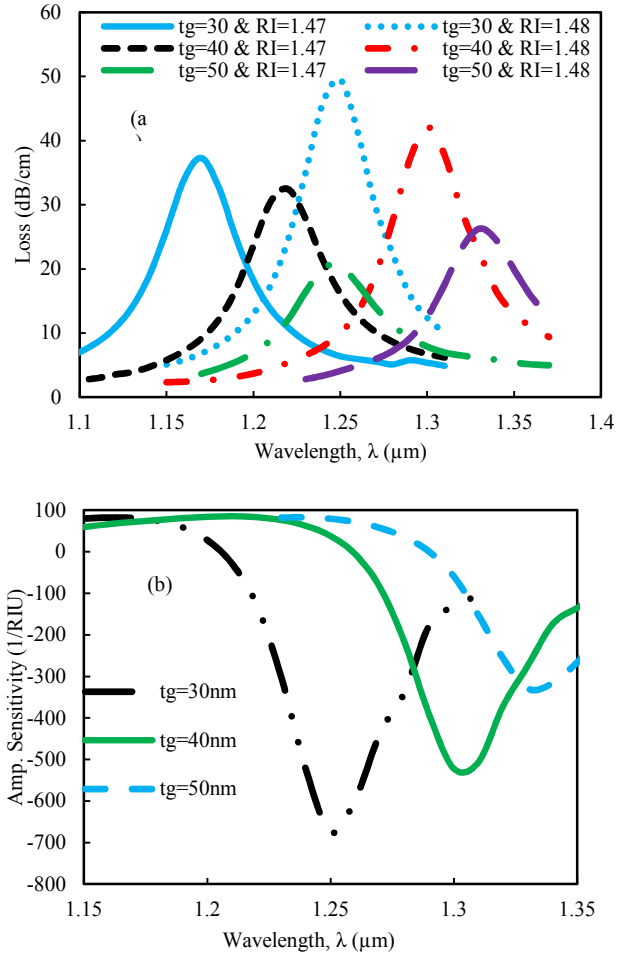


Fig. 5: (a) Gold layer thicknesses variation and loss for 1.47 and 1.48, and (b) Amplitude sensitivity with $\Lambda = 1.9$ µm, $d_s = 0.25\Lambda$, and $d_l = 0.5\Lambda$.

The PCF-based SPR sensor performance is measured in terms of sensitivity. The sensitivity of the proposed biosensor is obtained by using the wavelength, and amplitude interrogation method. The wavelength sensitivity can be calculated by the wavelength interrogation method using the following equation [5]:

$$S_\lambda(\text{nm/RIU}) = \Delta\lambda_{peak}/\Delta n_a \quad (2)$$

where $\Delta\lambda_{peak}$ denotes the wavelength peak variation, and Δn_a denotes the analyte RI change. The proposed sensor has a difference between peak and wavelength shift, $\Delta\lambda_{peak}$ of 70 nm, 80 nm, and 70 nm for RI variation Δn_a of 0.01, with changing of analytes RIs from 1.45-1.46, 1.46-1.47, and 1.47-1.48, respectively. Therefore, the theoretical wavelength sensitivities are computed of 7000, 8000, and 7000 nm/RIU, respectively. Another crucial parameter is the resolution of the sensor which shows how the sensor can detect a little change in the refractive index of an analyte. The following equation calculates the resolution [4]:

$$R\ (\text{RIU}) = \Delta n_a \times \Delta\lambda_{min}/\ \Delta\lambda_{peak} \quad (3)$$

Assuming $\Delta n_a = 0.01$ with $\Delta\lambda_{min} = 0.1$ nm, and $\Delta\lambda_{peak} = 80$ nm, the highest resolution of the *D*-shape sensor is found about 1.25×10^{-5} RIU. As a result, the proposed biosensor can be potentially detected analyte RIs variation of 10^{-5} order. To obtain the sensitivity, the wavelength interrogation method uses spectral manipulation which makes detection process complex but another process, amplitude variation can remarkably mitigate this problem by measuring sensitivity for a particular wavelength. The following formula defines the amplitude sensitivity [5]:

$$S_A = -\frac{1}{\alpha(\lambda, n_a)}\frac{\partial\alpha(\lambda, n_a)}{\partial n_a} \quad (4)$$

where $\alpha(\lambda, n_a)$ denotes the propagation loss and $\partial\alpha(\lambda, n_a)$ denotes the loss difference. Fig. 4(b) indicates the amplitude sensitivity curve for different analyte RIs. A maximum sensitivity of 522 RIU^{-1} is found at 1.30 µm wavelength with an n_a of 1.47. Besides, the amplitude sensitivity of about 330 and 333 RIU^{-1} are also found corresponding to the analyte RI of 1.45 and 1.46. Here, assuming 1% change of transmitted intensity, a maximum sensor resolution of 1.92×10^{-5} RIU is detected by the proposed sensor. The sensitivity of this proposed sensor can be comparable with the previously proposed work [5-8].

B. Gold Layer Variation and Sensitivity

Gold layer has a significant impact on sensing performance. The gold layer thickness (t_g) is tuned from 30 nm to 50 nm. Corresponding loss curve with analyte RI of 1.47 and 1.48 are given in Fig. 5 (a). We observed that increasing value of gold layer thickness increases loss depth. We also observed increment of gold layer thickness results a shifting of peak loss towards higher wavelength. The amplitude sensitivity of 700, 522, and 350 RIU^{-1} are obtained for the t_g of 30 nm, 40 nm, and 50 nm, respectively as illustrated in Fig. 5 (b). The physical reason for shifting of loss peak to the higher wavelength and decreasing of loss peak and amplitude sensitivity is of the reason for damping characteristics of gold. The frequency of the incoming light and the penetration depth are mutually proportional. Hence, the comparatively thin gold layer is the reason for better performance, and a thicker layer of gold decreases the sensitivity.

C. Linearity

A suitable sensor shows high linearity characteristics. Fig. 6 shows the linearity curve with analyte RI variation from 1.45 to 1.48. The following equation for linear line fitting can be written as y = 6.9x-8.916, where y represents the resonance wavelength, and x represents the analyte RI. The linear fitting characteristics of the proposed sensor shows the R-square value of about 0.9971 which indicates excellent linearity.

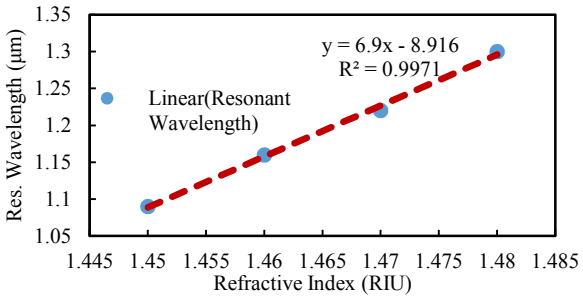


Fig. 6**:** Representation of line fitting properties of the sensor with varying n_a from 1.45 to 1.48.

Because of excellent line fitting properties, the suggested sensor can be a potential candidate for the liquid RI detection. The proposed *D*-shape PCF sensor can be practically implemented with the standard stack-and-draw fabrication process [15].

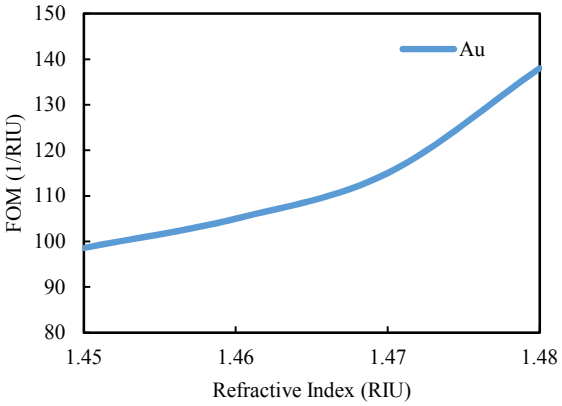


Fig. 7: FOM characteristics representation.

D. Figure of Merit

For the improved design, the signal to noise ratio (SNR) is an important parameter. The larger SNR can be the reason for smaller standard deviation with the variation of spectra. Moreover, the improvement of the SNR and spectral width will be a reason for a better detection range for the sensor, which can be formed with the term figure of merit (FOM). The FOM is obtained from [9]. During the simulation, the FOM of the designed sensor is obtained about 140 with a gold coating layer as shown in Fig. 7. Comparison with existing work has shown in Table 1.

TABLE 1: PERFORMANCE COMPARISON WITH THE EXISTING SENSOR.

Ref.	S_W (nm/RIU)	W_R (RIU)	S_A (RIU^{-1})	A_R (RIU)
Ref. [5]	4200	2.38×10^{-5}	300	3.33×10^{-5}
Ref. [7]	9000	----	318	1.11×10^{-5}
Ref. [8]	7700	1.29×10^{-5}	150	6.66×10^{-5}
Proposed Sensor	8000	1.25×10^{-5}	522	1.92×10^{-5}

IV. CONCLUSION

We proposed and numerically investigated the performances of a simple dual core *D*-shaped PCF sensor with the gold coating on its plane surface. The sensor with the gold layer is free from oxidation. The proposed PCF sensor requires less polishing effort, which potentially reduces fabrication difficulties. Sensing performance shows maximum wavelength interrogation sensitivity of 8000 nm/RIU and resolution of 1.25×10^{-5} RIU. Moreover, maximum amplitude sensitivity 522 RIU^{-1} and resolution of 1.92×10^{-5} RIU in the sensing range from 1.45-1.48. We believe that the proposed PCF can be considered as an ideal sensor for analytes refractive index detection because of its improved sensitivity and high linearity.

REFERENCES

[1] Y. Zhao, Z.Q. Deng, and J. Li,"Photonic crystal fiber based surface plasmon resonance chemical sensors."*Sens. Actuators B Chem.*, vol. 202, no.3, pp.557–567, 2014.

[2] E. Kretschmann, "Radiative decay of non-radiative surface plasmons excited by light," in *Z Naturforsch*, vol. 23, no. 12, pp. 2135-2136, 1968.

[3] R. Jorgenson and S. Yee, "A fiber-optic chemical sensor based on the surface plasmon resonance," in *Sensor Actuat. Biol. Chem.*, vol. 12, 1993, pp. 213–220.

[4] M. A. Mahfuz, Md. A. Mollah, M. R. Momota, A. K. Paul, A. Masud, S. Akter, and Md. R. Hasan, "Highly Sensitive Photonic Crystal Fiber Plasmonic Sensor: Design and Analysis", *Optical Materials*, vol 90, 2019, pp. 315-321.

[5] M. R. Momota and M. R. Hasan, "Hollow-core silver coated photonic crystal fiber plasmonic sensor," in *Optical Materials.*, vol. 76, 2018, pp. 287–294.

[6] T. Huang, "Highly sensitive spr sensor based on d-shaped photonic crystal fiber coated with indium tin oxide at near-infrared wavelength," in *Plasmonics*, June 2016.

[7] S. Chakma, M. A. Khalelk, B. K. Paul, K. Ahmed, M. R. Hasan, and A. N. Bahar, "Gold-coated photonic crystal fiber biosensor based on surface plasmon resonance:design and analysis," in *Sensing and biosensing Research*, vol. 18, 2018, pp. 7–12.

[8] R. K. Gangwar and V.K. Singh, "Highly Sensitive Surface Plasmon Resonance Based D-Shaped Photonic Crystal Fiber Refractive Index Sensor." *Springer Science and Business Media*, 25 Sep. 2016.

[9] F. Haider, R. A. Aoni, R. Ahmed, and A. E. Miroshnichenko, "Highly amplitude-sensitive photonic crystal fiber based plasmonic sensor", *Journal of the Optical Society of America B,* vol 35, no 11, November 2018, pp. 2816-2821.

An Efficient Soft Error Tolerant Approach to Enhance Reliability of TCAM

Rumana Afrin
Department of Computer Science & Engineering
Khulna University of Engineering & Techonology
Khulna-9203, Bangladesh
r.afrin07@gmail.com

Muhammad Sheikh Sadi
Department of Computer Science & Engineering
Khulna University of Engineering & Techonology
Khulna-9203, Bangladesh
sadi@cse.kuet.ac.bd

Jan Jürjens
Department of Computer Science
University of Koblenz-Landau, and
Fraunhofer-Institute for Software and Systems Engineering ISST
Koblenz, Germany
juerjens@uni-koblenz.de

***Abstract*— Ternary Content Addressable Memory (TCAM) is a kind of extraordinary memory which is used to search operations in routers, firewalls, network address translators, and entire content in a single clock cycle. It is a special kind of memory which is used in highly reliable network applications. These systems require high reliability since a single error may interrupt any computation which may cause total system failure. However, TCAM is facing serious reliability challenges due to soft errors. The probability of soft errors in TCAM is increasing progressively due to alpha-particle, cosmic rays, high-vitality particles, and ionizing impacts of barometrical neutron. In order to minimize the risks of soft errors in TCAM, a new idea is proposed in this paper where the search key is constructed in a tricky way and the contents of the TCAM are ordered in a sequential pattern. The proposed method provides soft error tolerance with a low data overhead. Experimental study shows how the proposed method outperforms the existing method to tolerate soft error.**

Keywords— Soft Errors; Ternary Content Addressable Memory (TCAM); Memory Reliability; Error Correction Code; Don't Care Bits.

I. INTRODUCTION

Due to continuous downscaling of CMOS circuit, transistors are becoming very sensitive towards soft errors [1]. A little amount of charge can change the state level of a transistor, as a result of causing soft error [2]. Soft errors are matters of concern for electronic hostile systems and/or the systems which require high reliability [3], [4], [5]. TCAM which is one of the most essential components for networking devices and data transforming tool [6], are of high risks due to soft errors. TCAM is a vital tool for router, because TCAM helps to lookup entire routers, packet classification, and packet forwarding. It instantaneously compares the input data with the entire data saved in TCAM [7]. TCAM uses a special kind of data bit, named don't care bit (*), which has a binary value of 1 and 0 so that the data do not need to match exactly and it helps to speed up the networking process. If a smaller sized transistor is used in a high-speed memory of the networking device, it will be more susceptible to soft errors [8]. So, it is necessary to filter, encode, decode, and deliver data between sender and receiver without any soft error. Therefore, a soft error tolerant TCAM is extremely important as networking devices are used for dynamic communication setup [9]. However, TCAM dimensions and design practices are vital to face the challenges of soft errors. Due to increase of memory sizes, the chances of soft errors are also increasing [10].

The conventional TCAMs suffer from effective error correction techniques. Circuit or architectural solutions, which are generally aimed at increasing the critical charge, require modifications of the TCAM architecture; hence, these types of solutions incur a large amount of process cost. Several techniques involve the utilization of additional DRAM cells to tolerate soft errors. However, apart from the large cost overheads and the increased complexity of DRAM fabrication, the solution retains the possibility that incorrect data may lead to false miss or match results between two refresh intervals [11]. Error Ccorrection code (ECC), which is commonly used for RAMs [12], is another type of solution. However, at least 3 times larger memory bits are required than the original memory bits of the TCAMs to correct only 1-bit error per word when ECC is applied [9]. Therefore, further research are needed to prevent soft errors against multiple bit upset in TCAM.

In this paper, we have proposed a new methodology to improve the efficiency of TCAM. The proposed methodology is able to provide more reliable TCAMs for sequential data words. It will also help to speed up network device performance, with a low cost and information overhead, and circuit complexity.

Section II entails related work in this area. In Section III, the proposed methodology is illustrated. Test contemplates are introduced in Section IV. At last, conclusion is encompassed into Section V.

II. RELATED WORK

Several hardware and software methodologies are used to avoid soft error. Eftaxiopoulos, Axelos and Pekmestzi proposed a new methodology in [11] where, a special storage component was used to prevent soft errors. Here, two techniques were combined; one is a guard gate technique and another one is the usage of a pass transistor as a delay element. This methodology worked only for Single Node Upset (SNU) in case of Content Addressable Memory (CAM) and 75% soft error tolerance for TCAM. Moreover, this method needs two extra transistors to provide soft error tolerance for TCAM.

A Soft-error tolerant TCAM is proposed in [13] where an arbitrarily selected search key was used to estimate; and parallel TCAMs were used to detect single-bit flip error (here, the search key matches the last word). This method can only detect single bit upset and to correct this upset bit, it used a

backpack-SRAM. It also used an extra TCAM which is comparatively expensive.

A Bloom Filter based method is proposed in [14] to detect soft errors in TCAMs with a low power and data overhead. It can detect and correct single bit upset. The main disadvantage of this method is false positive. It happens when the Bloom Filter detect the presence of a data even though there was no data in the Bloom Filter.

Onawa, Matsunaga and Hanyur proposed a transistor /magnetic-tunnel-junction (MTJ) hybrid structure based asynchronous TCAM method in [15] which used 20 transistors and 4 MTJ device to detect single bit upset. This hardware-based method is much costlier than most other approaches.

A probabilistic approach to determine scrubbing interval is proposed in [16] where Compound Poisson (CP) process is used to count clustered random events. This method can detect and correct both single and multiple bit upsets, but CP model failing possibility is large due to independent assumptions in the experimental data.

A software-based error detection and correction technique was proposed in [17] which used three models to serve the purpose of capturing the underlying phenomenon of a sensor network and then guide the error detection to discover any anomaly observation. This method used two TCAMs which is highly expensive.

Therefore, we can notice that almost none of the existing techniques can work effectively in tolerating soft errors for TCAMs.

III. THE PROPOSED METHODOLOGY TO TOLERATE SOFT ERRORS

To prevent single bit and multiple bit soft errors, a new method is proposed in this paper. The proposed methodology does not need any extra TCAM or ECC-SRAM to detect and correct soft error. The several steps of the proposed methodology are illustrated as follows.

A. TCAM Encoding and Decoding

Suppose data bits are saved in TCAM with n words' length. Then total number of data words is 2^n. Now total data words are divided into $\frac{2^n}{4}$ groups sequentially. Each group contains 4 words and each word contains n data bits. In each word, nth and (n-1)th data bits are different for a group. Alternatively, 1 to n-2 data bits' are same for all words within a group. For generating search key in each group, the paper uses the following steps.

- First, all similar 1 to n-2 data bits are saved. Then, it ensures that in each group, the last 2 data bits are same for all groups' word. In case of all groups' first word, n-1th and nth data bit are "00" respectively. Similarly, in the second word, n-1th and nth data bit are "01"; in the third word, n-1th and nth data bit are "10"; and finally in the fourth word, n-1 and nth data bit are "11" respectively.
- Now for generating search key with don't care bit, any data word can be chosen within a group. Then the search key for each group is $S_1S_2S_3 \ldots S_{n-1}S_n$.
- $D_1,\ D_2,\ D_3 \ldots . D_{n-2}$ are similar data bits of each group and $D_{n-1},\ D_n$ are last two data bits which are different for each word within a group.
- Finally, it generates a search key including the don't care bits which are $S_1S_2S_3 \ldots S_{n-1} *$ and $S_1S_2S_3 \ldots *S_n$.

If first word's last two data bits "00" are chosen for search key, then the search key with don't care bit will be $S_1S_2S_30*$ and $S_1S_2S_3 \ldots *0$ respectively. For detecting and correcting soft errors, these two search keys need to fulfill the following conditions:

- For the first data word of each group, if $S_1S_2S_3 \ldots . 0*$ and $S_1S_2S_3 \ldots *0$ match with the data word, no error will occur. And if any error occurs in last data bit of first word then only $S_1S_2S_3 \ldots . 0*$ search key will match with the first data word. If the second last bit is erroneous, then it will match only with $S_1S_2S_3 \ldots *0$ search key. If the last and second last data bits are both erroneous then it will not match with any search key.
- For the second data word of each group, if $S_1S_2S_3 \ldots . 0*$ and $S_1S_2S_3 \ldots *0$ search keys match with the data word, the last data bit is erroneous bit. If the second last data bit is erroneous, it will not match with any search key. If both the last and second last bit are erroneous, it will match only with this $S_1S_2S_3 \ldots *0$ search key. Otherwise, stored data word is correct.
- For the third data word of each group, if $S_1S_2S_3 \ldots 0*$ and $S_1S_2S_3 \ldots *0$ both match with the data word, second last data bit is erroneous. If the first data bit is erroneous, it will not match with any search key. If both the last and second last bits are erroneous then it will match only with $S_1S_2S_3 \ldots . 0*$ search key. Otherwise, stored data word is correct.
- For the fourth data word of each group, if both $S_1S_2S_3 \ldots 0*$ and $S_1S_2S_3 \ldots *0$ search keys match with the data word, then the last and the second last data bit are erroneous. If the first data bit is erroneous then it will match only with $S_1S_2S_3 \ldots *0$ search key. If the second last bit is erroneous then it will match only with $S_1S_2S_3 \ldots 0*$ search key. Otherwise, stored data word is correct.

Similarly, we can choose 2nd, 3rd or 4th words' last two data bits "01", "10" or "11" respectively for generating the search keys. Then the search key with don't care bit for 2nd word will be $S_1S_2S_3 \ldots 0*$ and $S_1S_2S_3 \ldots *1$. For 3rd data word, the search keys will be $S_1S_2S_3 \ldots 1*$ and $S_1S_2S_3 \ldots *0$. And for 4th data word, the search keys will be $S_1S_2S_3 \ldots 1*$ and $S_1S_2S_3 \ldots *1$. To detect and correct soft errors, these two search keys need to fulfill the conditions which are shown in TABLE I.

TABLE I shows the conditions for detection and correction of soft error for various search keys. Green words represent correct data words which are saved in TCAM, and red words represent erroneous data words which are saved in TCAM. A checkmark (✓) shows correct match with the search keys and a cross-mark (**×**) shows false match among search keys and data words.

The proposed methodology to detect and correct soft errors is illustrated in Fig. 1, The first N sequential data words are saved in memory and then these are divided into four groups. Each group of the data words are then divided into two further groups. One is similar bits' data group and another one is dissimilar data bits' group. Then search keys, along with don't care bits, for all groups are generated.

TABLE I. CONDITIONS FOR DETECETING AND CORRECTING DATA WORDS BY USING VARIOUS SEARCH KEYS

Search Keys	Correct Data Bits	Error in Last Data	Error in the Second Last Data	Error in Both Data
first group	$D_1D_2D_3..00$	$D_1D_2D_3..01$	$D_1D_2D_3..10$	$D_1D_2D_3..11$
$S_1S_2S_3..0*$	✓	×		
$S_1S_2S_3..*0$	✓		×	
$S_1S_2S_3..1*$			×	×
$S_1S_2S_3..*1$		×		×
second group	$D_1D_2D_3..01$	$D_1D_2D_3..00$	$D_1D_2D_3..11$	$D_1D_2D_3..10$
$S_1S_2S_3..0*$	✓	×		
$S_1S_2S_3..*0$		×		×
$S_1S_2S_3..1*$			×	×
$S_1S_2S_3..*1$	✓		×	
third group	$D_1D_2D_3..10$	$D_1D_2D_3..11$	$D_1D_2D_3..00$	$D_1D_2D_3..01$
$S_1S_2S_3..0*$			×	×
$S_1S_2S_3..*0$	✓		×	
$S_1S_2S_3..1*$	✓	×		
$S_1S_2S_3..*1$		×		×
fourth group	$D_1D_2D_3..11$	$D_1D_2D_3..10$	$D_1D_2D_3..01$	$D_1D_2D_3..00$
$S_1S_2S_3..0*$			×	×
$S_1S_2S_3..*0$		×		×
$S_1S_2S_3..1*$	✓	×		
$S_1S_2S_3..*1$	✓		×	

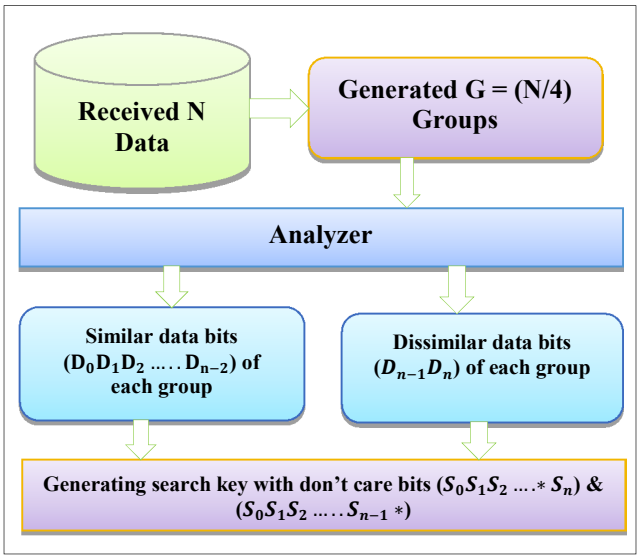


Fig. 1. A diagram of generating search keys with don't care bit.caption.

As shown in Fig. 2, the proposed method compares the search keys with the saved similar data bits of each group. If it matches, similar saved data bits are correct and if it does not match then an error is occurred in the saved data bits.

As a result, similar data bits are corrected by using search key and the different data bits are corrected by using search keys along with don't care bits. If search keys and data bits fulfill the conditions, the saved data bits are correct; otherwise an error is detected. In the next step, erroneous bits are corrected by using NOT operation. The correct data word is generated and saved in memory.

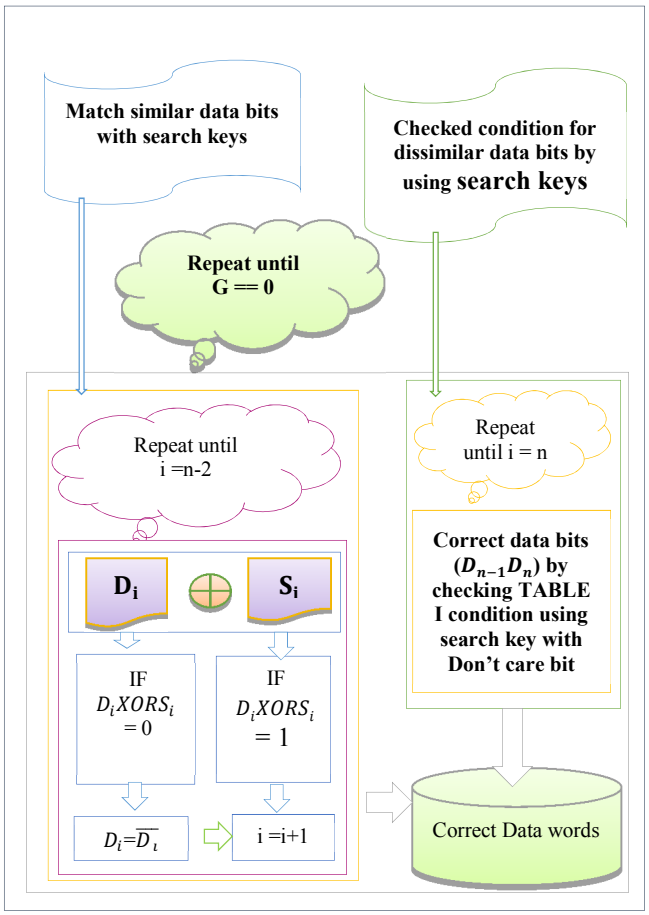


Fig. 2. The Flow diagram of the Proposed Methodology.

B. A Typical Example

Let us give an example of the proposed methodology. Suppose data words are saved in a TCAM with a word length of n=4. Then the number of data words which are saved in TCAM are $2^n=2^4 = 16$. In Table II, the saved data words in a TCAM are presented. Total data words are divided into $\frac{16}{4}$ =4 groups. For the first group, similar data bits are "00", for the second group, similar data bits are "01", for the third group, similar data bits are "10" and finally for the fourth group, similar data bits are "11". Within a group, for each word, the last two data bits do not match with each other. It is different for every word in a group. For generating search keys with don't care bits, we can choose any data word's last two bits. Here, we choose "11" as last two bits of search keys and the generated search keys for the first group are 001* and 00*1, for the second group search keys are 011* and 01*1, for the third group search keys are 101* and 10*1, and for the fourth group search keys are 111* and 11*1. Hence, if a single error or multiple soft errors occur in the data word, we can detect and correct these error(s) by using the conditions as shown in TABLE I.

In the next step, errors are injected in the saved data words. Few erroneous data words are shown in Table III. By using search keys 001* and 00*1, we can detect and correct erroneous data words within the data words that are saved in the first group.

TABLE II. DATA WORDS SAVED IN TCAM WITH A WORD LEANGTH OF 4

Data Word	Similar Data Bits in each Group	Dissimilar Data Bits in each Group	Search Keys with Don't care Bit
0000	00	00	001*
0001		01	
0010		10	00*1
0011		11	
0100	01	00	011*
0101		01	
0110		10	01*1
0111		11	
1000	10	00	101*
1001		01	
1010		10	10*1
1011		11	
1100	11	00	111*
1101		01	
1110		10	11*1
1111		11	

TABLE III. TABLE CORRECTION OF ERRONEOUS DATA WORDS BY SEARCH KEYS

Data Words with erroneous bits	Similar Data Bits in each Group	Dissimilar Data Bits in each Group	Search Keys with Don't care Bit	Correct Data Words
1111	00	11	001*	**0000**
0001		01		0001
0010		10	00*1	0010
1011		11		**0011**
0100	01	00	011*	0100
0101		01		0101
0110		10	01*1	0110
0111		11		0111
1000	10	00	101*	1000
1010		10		**1001**
1010		10	10*1	1010
1011		11		1011
1100	11	00	111*	1100
1101		01		1101
1111		11	11*1	**1110**
1111		11		1111

- In the first group, first two bits of the first and the fourth data words do not match with the first two bits of the search keys. Therefore, an error is detected in here. We correct the data bits "11" to "00" and "10" to "00". After that, the conditions of Table I is applied. Here, both search keys match with the first data word. So according to the conditions of Table I, the last two data bits of both second and fourth data words are erroneous bits and we correct these bits "11" to "00" by using inverse operation.
- Similarly, we checked other groups. In the third group, we got a mismatch. By using the conditions of Table I, we observed that both search keys are matched with the second data word. So, the last two bits of second data word are erroneous, and we correct those bits "10" to "01" by using inverse operation.
- In the fourth group, the third data word do not match with the search key. Here, the last data bit is an erroneous bit. It matched only with search key $S_1S_2S_3\ldots1*$. So, we correct the data bits "1" to "0" by using inverse operation.

In this way, we could detect and correct single or multiple bits' soft errors in the data words.

IV. EXPERIMENTAL ANALYSIS

The experimental analysis is done in three stages. The first one is experimental setup, the second is fault injection, and the third is results analysis.

A. *Exerimental Setup*

The following tools are used to implement the proposed methodology.

- Intel(R) Core ™ 2 Duo CPU T6570 @ 2.10 GHz
- 4 GB RAM
- C++ Language
- Code Block Equations

B. *Fault Injection*

The fault injection method which is *applied to* estimate the error detection and correction abilities of the proposed method is illustrated in Algorithm 1

Algorithm 1: Fault injection in TCAM

/* Input: D, n, where D is words saved in a TCAM, n each word length*/

1. Generating TCAM 1 from saved TCAM
2. Declare variable i, j and k
3. Initialize variable i, to zero
4. Initialize variable j, to zero
5. Repeat the step five until i less than or equal to D
6. Repeat the step six until j less than or equal to n
7. **IF** D [i][j] equal to zero **THEN**
8. Reset D [i][j] equal to one
9. **ELSE**
10. Reset D [i][j] equal to zero
11. **END IF**
12. Set k equal to j
13. Repeat the step thirteen until **k** equal ton
14. **IF** D [i][k+j] equal to zero **THEN**
15. Reset D[i][k+j] equal to one
16. **ELSE**
17. Reset D[i][k+j] equal to zero
18. **END IF**
19. Return TCAM 1

C. *Results Analysis*

Here 4*16= 64 bits' data are taken for analyzing the capability of the proposed method. After that, the error correction rate for the proposed method and KX-TCAM [7] are calculated. The error correction coverage for the proposed

method and KX-TCAM is shown in TABLE IV. Here, the first column represents the number of soft errors in words. From Table IV, we can observe that the proposed method has better error correction coverage than KX-TCAM.

TABLE IV. COMPARISON OF CORRECTION COVERAGE BETWEEN THE PROPOSED METHOD AND KX-TCAM

Corrupt bits in words	KX-TCAM	Proposed Method For sequential data words	Proposed Method For random data words
1	100%	100%	100%
2	100%	100%	100%
3	100%	100%	100%
4	50%	100%	100%
5	20%	100%	80%
6	0%	100%	66.7%
7	0%	100%	57.14%
8	0%	100%	50%

For sequential data words, the proposed method can detect and correct 100% single and multiple erroneous data bits. Alternatively, for randomly stored data words, the proposed method's error detection and correction rate is lesser than 100%. However, the proposed method has better coverage than the existing method because the existing method worked well only for some combination but not for all. Moreover, randomly selected search keys cannot detect and correct single or multiple bits of soft error. In Fig. 3, the comparison of the error correction rate between KX-TCAM and the proposed method is shown.

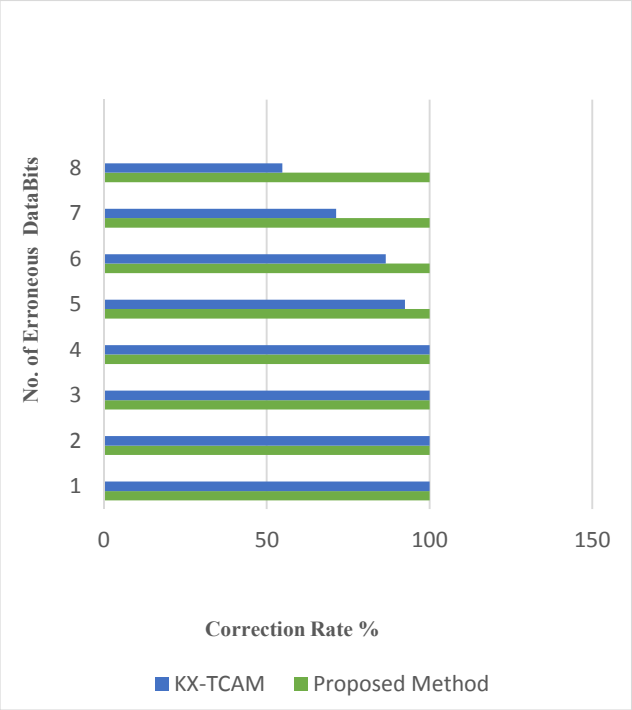


Fig. 3. Comparison of error correction rate between the Proposed Method and KX-TCAM

The comparison of coding efficiency between the Proposed Method and KX-TCAM is presented in TABLE V. Here, we can observe that for 64 bits' data word, KX-TCAM required 64 extra bits for detecting and correcting erroneous

TABLE V. COMPARISON OF CODING EFFICIENCY BETWEEN THE PROPOSED METHOD AND KX-TCAM

Error correction method	Data bits	Redundant bits	Correct error bits	Code word	Search key for MCU	Code Overhead (%)	Coding efficiency (%)
KX-TCAM	64	64	4	128	6	100	50
Proposed Method	64	40	64	104	2	62.5	38.46

bits whereas the proposed methodology required 40 extra bits for detecting 64 erroneous bits. KX-TCAM used 6 different search keys to detect and correct multiple cells whereas the proposed method used 2 different search keys. The fault tolerance capability of the proposed method is higher than the existing method. The code overhead of the proposed method is 62.5% whereas the existing method's code overhead is 100%. Fig. 4, illustrates the comparison of code overhead and coding efficiency between KX-TCAM and the proposed method.

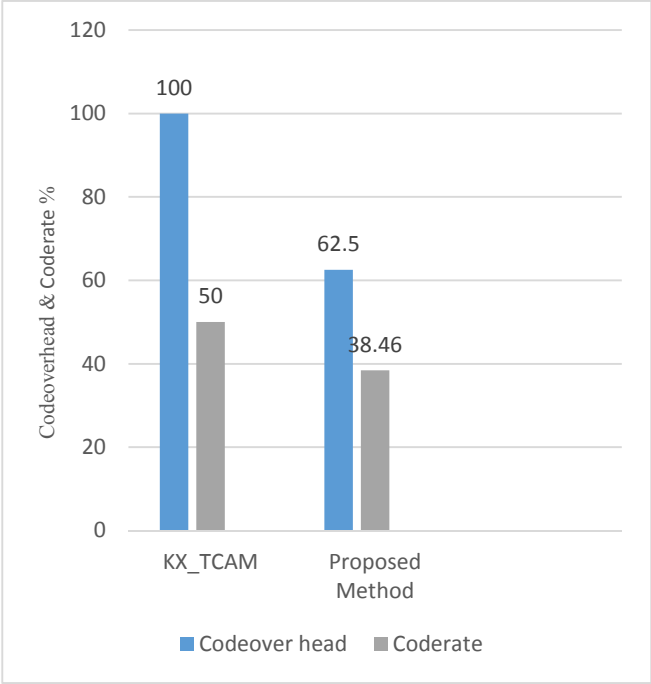


Fig. 4. Comparison of codeoverhead and coderate between KX_TCAM and Proposed method

Fig. 5 shows the comparison of memory requirement between KX-TCAM and the propose method. KX-TCAM used more storage space in memory than the proposed method.

Fig. 6 illustrates the comparison of execution time between KX-TCAM and the proposed method. The proposed method has a high-speed error detection and correction capability which helps to speed up the network.

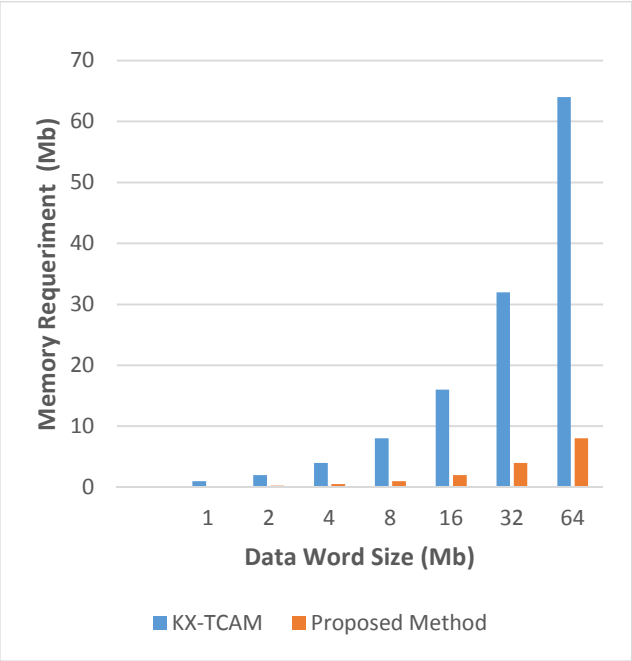


Fig. 5. Comparison of memory requeriment between KX_TCAM and Proposed method.

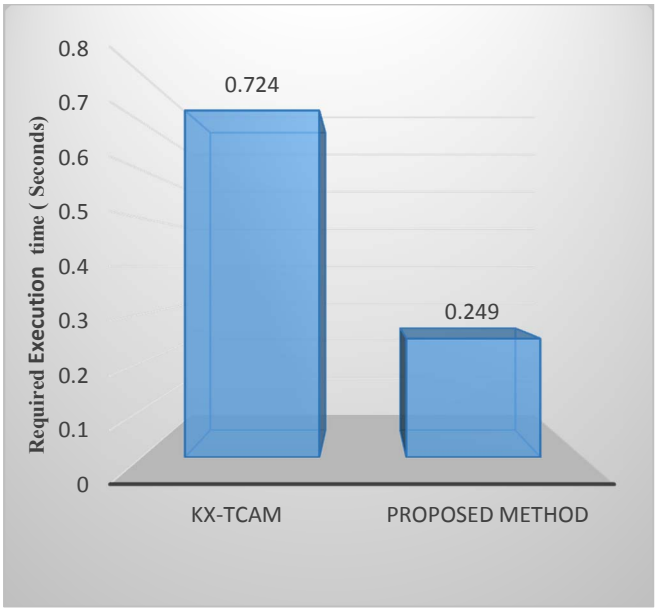


Fig. 6. Comparison of execution time requeriment between KX_TCAM and Proposed method.

V. CONCLUSIONS

This paper has proposed a novel idea to tolerate soft errors in TCAM. The proposed method uses a limited amount of memory space which helps to reduce the cost and time complexity than other existing methods. Moreover, its detection and correction rate is higher for both single bit error and multiple bit errors for a sequentially stored data words in TCAM. It also uses a simple code which can make the searching and forwarding speed of the network device faster with a lower time overhead. Although the proposed method works for sequentially stored data words only and this method do not work for randomly stored data words, the experimental analysis validates the effectiveness of the proposed method over existing dominant method.

Further scopes in order to extend this research is to develop a methodology which can detect and correct soft errors from both sequentially and randomly stored data words in TCAM.

REFERENCES

[1] X. Cao, L. Xiao, J. Li, R. Zhang, S. Liu and J. Wang, "A Layout-Based Soft Error Vulnerability Estimation Approach for Combinational Circuits Considering Single Event Multiple Transients (SEMTs)," in *IEEE* Transactions on Computer-Aided Design of Integrated Circuits and Systems. doi: 10.1109/TCAD.2018.2834425

[2] T. F. Tay and C. Chang, "A Non-iterative Multiple Residue Digit Error Detection and Correction Algorithm in RRNS," IEEE Transactions on Computers, vol. 65, no. 2, pp. 396-408, Feb. 2016.

[3] F. Vincenzo De and B. Chris, "A survey of linguistic structures for application-level fault tolerance," *ACM Computing Surveys,* vol. 40, pp. 6-43, 2008.

[4] G. P. Saggese, N. J. Wang, Z. T. Kalbarczyk, S. J. Patel, and R. K. Iyer, "An experimental study of soft errors in microprocessors," *Micro, IEEE,* vol. 25, pp. 30-39, 2005.

[5] P. N. Sanda, J. W. Kellington, P. Kudva, R. Kalla, R. B. McBeth, J. Ackaret, R. Lockwood, J. Schumann, and C. R. Jones, "Soft-error resilience of the IBM POWER6 processor," *IBM Journal of Research and Development,* vol. 52, pp. 275-284, 2008.

[6] D. E. Taylor, "Survey and taxonomy of packet classification techniques,"ACM Computing Surveys, vol. 37, no. 3, pp. 238-275, Sept. 2005.

[7] Infall Syafalni, Tsutomu Sasao, and Xiaoqing Wen, "A Method to Detect Bit Flips in a Soft-Error Resilient TCAM," TCAM," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 37, no. 6, pp. 1185 – 1196, Aug. 2017.

[8] W. Leung, F. Hsu, and M. E. Jones, "The ideal SoC memory: 1TSRAM," 13th ASIC/SOC Conference, pp 32-36, Sept. 2000.

[9] Sriram C. Krishnan, Rina Panigrahy, andSunil Parthasarathy, "Error-Correcting Codes for Ternary ContentAddressable MemoriesIEEE Transactions on Computer, vol. 58, no. 2, pp. 275 – 279, Feb. 2009.

[10] R. Reed, M. Carts, P. Marshall, C. Marshall, O. Musseau, P. McNulty, D. Roth, S. Buchner, J. Melinger, and T. Corbiere, "Heavy ion and proton-induced single event multiple upset," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 6, pp. 2224–2229, Dec. 1997.

[11] Nikolaos Eftaxiopoulos, Nicholas Axelos, Kiamal Pekmestzi, "Low leakage radiation tolerant CAM/TCAM cell," IEEEConferences 21st International On-Line Testing Symposium (IOLTS), pp.206 - 211, Aug. 2015.

[12] J. A. Fifield and C. H. Stapperet, "High-Speed On-Chip ECC forSynergistic Fault-Tolerant Memory chips", *IEEE Journal of Solid-StateCircuits,* vol.26 pp.1449-1452, Oct 1991.

[13] I. Syafalni, T. Sasao, X. Wen, S. Holst, and K. Miyase, "Soft-error tolerantTCAMs for high-reliability packet classification," IEEE APCCAS Conferance, pp. 471-474, Feb. 2015.

[14] S. Pontarelli, M. Ottavi, A. Evans, and S.-J. Wen, "Error detection in ternary CAMs using bloom filters," IEEE DATE Conference, pp. 1474-1479, March 2013.

[15] Naoya Onizawa ; Shoun Matsunaga ; Takahiro Hanyu, "A Compact Soft-Error Tolerant Asynchronous TCAM Based on a Transistor/Magnetic-Tunnel-Junction Hybrid Dual-Rail Word Structure," IEEE Conferance20th IEEE International Symposium on Asynchronous Circuits and Systems, pp. 1 - 8,June 2014.

[16] S. Baeg, S. Wen, and R. Wong, "Minimizing soft errors in TCAM devices: A probabilistic approach to determining scrubbing intervals," IEEE TCAS, vol. 57, no. 4, pp. 814-822, April 2010.

[17] M. Z. Shafiq, C. Meiners, Z. Qin, K. Shen, and A. X. Liu, "TCAMChecker: A software approach to the error detection and correlation of TCAM-based networking system," Journal of Network and System Management, vol. 21, no. 3, pp. 335-352, Sept. 2013.

An investigation on the surface states evolution of polycrystalline silicon nanowires for gas sensing application

M. M. A Hakim, Afrida Arshad, Sharmin Shoma and Urmila Saha
Department of Electrical & Electronic Engineering, East West University, Dhaka-1212, Bangladesh
*E-mail: dmmah@ewubd.edu

Abstract- **We perform a feasibility study on the application of thin film polycrystalline silicon (polysilicon) nanowires for gas sensing application upon exposure into different gas ambient. The nano-sensors are developed in a relatively cheap reactive ion etching system using a low-density plasma based etch and mature 3 micron lithography. Systematic measurements in air, vacuum, hydrogen and isopropyl alcohol demonstrate promising gas sensitivity and impressive response/recovery times in room temperature. The modification of surface states in nanowires upon exposure into different gas ambient is confirmed through 3D numerical simulation. This investigation demonstrates the potential of polysilicon nanowires for detecting molecules in gas phase thereby reveals that the polysilicon nanowire sensor is promising for a simple, low cost, manufacturable route for gas-sensor fabrication.**

***Index Terms*— Thin films, nano-sensors, gas sensing, biosensors**.

I. INTRODUCTION

Smart nanosensors with the capability of detecting toxic gases from the environment and multiple biomarkers from physiological fluids have significant impact on the society. NO_x, O_2, NH_3, SO_2, CO, CO_2, and hydrocarbon detection has been the major concern for environmental protection [1]. The medical use of gas sensors is also a rapidly growing field [2-3]. The different compounds which have been identified in human breath, most of them are potential indicators (markers) of more than one type of diseases [2-5]: NO has been widely studied as a bio-marker for oxidative stress [4]; exhaled hydrocarbons of low molecular mass, such as ethane, n-pentane; ethylene, isoprene are markers for blood cholesterol levels [5].

Conventional gas sensing usually employs IR detection, UV spectroscopy, mass spectroscopy, or gas chromatography which are conclusive and highly reliable methods for detecting parts per billion (ppb) levels [6]. Even though these methods are precise, their physical size, cost and operation complexity prevent them from entering into everyday life use. The electrochemical metal oxide gas sensors although reduces the cost and complexity, most of the sensor devices developed so far requires high temperature operation [7]. Over the last two decades, enormous efforts actually have been made to lower the operating temperature of metal oxide gas sensors by adopting metal nanoparticles on metal oxide surface [8]. These efforts were successful in lowering the sensors operating temperature to a value close to 100°C, for some gases but operating temperature is still reasonably high [8].

Silicon nanowire-based gas sensors are currently in interests due to its promise for fast response in room temperature [9] which may ensure low power operation for incorporation into personal items like mobile phones that typical gas sensor does not offer. Within the framework that is readily amenable into standard CMOS technology, Si nanowire-based gas sensors would also allow signal processing and interconnects to be implemented on the same substrate in order to form smart sensor system which can simplify potentially complex and costly gas detection process [10]. However, silicon nanowire based gas sensors reported so far mostly exploited bottom up or top down processed single crystal silicon nanowires [11-13]. Bottom-up grown nanowires are not location controlled and hence, construction of functional device arrays challenging [11]. These shortcomings are eliminated in conventional top down approaches but have disadvantage of elevated budget associated with these advanced lithography techniques and costly SOI wafers [12-13].

Spacer etch process of thin film has been recently reported as a route towards low-cost nanowire fabrication [4-16]. Mature lithography and industry standard deposition/etch are the attractive features of this technology. This process can also enable low cost disposable diagnostics and portable, wearable or even implantable sensors due to the use of glass or plastic substrates. However, the sensor application of polysilicon nanowires is concentrated on the detection of analytes in aqueous environment, mainly within the context of biosensing (e.g., to detect biological species like DNA, proteins etc.) [14-16]. The surface states modulation of nanowires upon exposure of different gas ambient and hence, the detection of molecules in gas phase using polycrystalline silicon nanowires and are rare in the literature.

This article first time performs a feasibility study on the application of thin film polycrystalline silicon (polysilicon) nanowires (NW) for gas sensing application. Systematic measurements of nanowire conductance in air, vacuum, hydrogen and isopropyl alcohol demonstrate the potential of polysilicon nanowires for gas sensing. In addition, experimental investigation is augmented with numerical simulation to get insight into the surface states modulation of polysilicon nanowires in different gas ambient.

II. SENSOR FABRICATION AND CHARACTERISATION

The sensor fabrication and characterization were done in the Southampton Nanofabrication Center, University of Southampton, UK. 500 nm nitride and 100 nm oxide layers were deposited on (100)-oriented n-type silicon wafers which had a doping density of $10^{16}/cm^3$. Mature lithography (linewidth > 3 μm) was used to dry etch 100 nm oxide layer to create pillars. A low-pressure chemical vapor deposition (LPCVD) method was then used to deposit a 100 nm undoped amorphous silicon (α-Si)

film at 560°C and subsequently the film was doped by boron implantation (dose of $1.5\times10^{13}/cm^2$, energy of 25 keV). A special etch process with 12 sccm SF_6 and 12 sccm O_2 flow at a pressure of 30 mT was used to dry etch the α-Si film in an Oxford Instruments capacitively coupled RIE system. Such a spacer etch process leaves amorphous silicon nanowires around the oxide pillar sidewalls, which were later crystallized using thermal anneal. The physical structure of the nanowires (width, shape and surface morphology) was investigated by plan-view and cross-sectional scanning electron microscopy (SEM). Electrical characterisation of the nano-sensors was performed using an Agilent B1500A Semiconductor Parameter Analyser. The gas sensing measurements were performed utilizing the sealed chamber of a Lakeshore EMTTP4 probe station which was connected to a H_2 gas cylinder with a pressure control valve. Fig. 1 shows an optical microscopic image of the nanowire sensor.

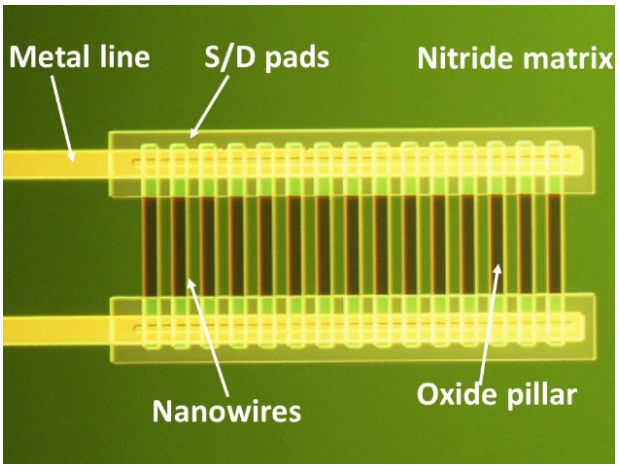


Fig. 1: Optical microscopic image of the polysilicon nanowire sensor.

III. 3D SIMULATION OF POLYSILICON NANOWIRES

Numerical simulations of nanowires were performed in the SILVACO TCAD platform installed in the VLSI lab of the East West University [17]. A p-type silicon nanowire 100 nm thickness and doping density was 6×10^{16} cm^{-3} (p-type) was created on a 500 nm Silicon Nitride (Si_3N_4) which was deposited on a 500 nm n-type Si substrate with substrate doping density $1\times10^{16}/cm^3$. A 20 nm thick aluminum material used as back gate and in the silicon nanowire, two heavily doped regions ($10^{20}/cm^3$, p-type) in the two sides of the channel were implemented to ensure ohmic contacts on the source/drain regions. There was 10 nm oxide on the polysilicon NW which was embedded in the sea of nitride layer. This specific type of NW and associated platform was chosen for 3D simulation to resemble our fabricated NWs which was fabricated targeting selective chemical functionalization on the NW only. During simulation, 2D coupled Poisson's drift-diffusion solver was used to model carrier transport in Polysilicon NW [17]. The Shockley-Read-Hall model for carrier emission and absorption was used to reflect the recombination phenomena within the device [17]. Moreover, the bandgap narrowing (BGN) effects in the heavily doped source/drain regions were included by Slotboom and de Graaf's analytical BGN model [17]. The polysilicon is modeled using a continuum of trap states consist of both donor-like and acceptor-like states distribution across the energy band gap [18]. Fig. 2 shows the 2D cutline view of the realized 3D structure in SILVACO platform.

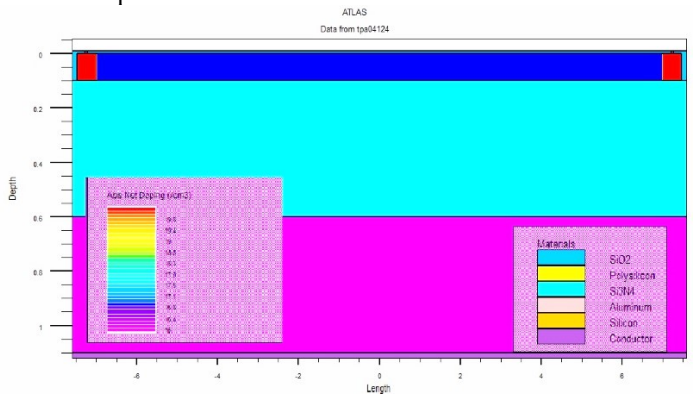


Fig. 2: 2D cutline view of realized polycrystalline silicon nanowire.

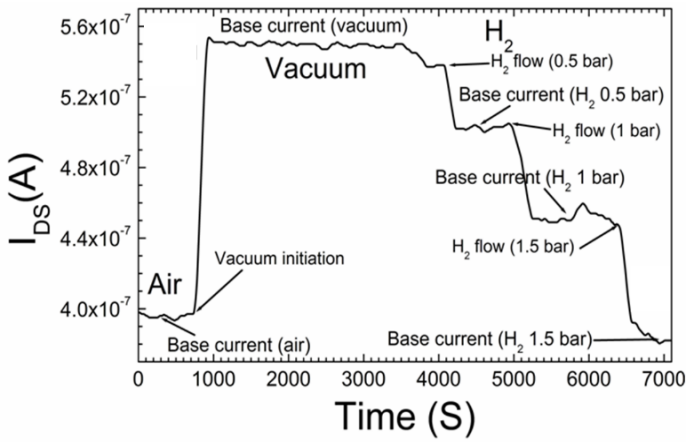


Fig. 3: The measured current response as a function of time for 58 µm long nanowires, sequentially in air, vacuum and different pressures of H_2 (5%) gases.

IV. RESULTS & DISCUSSION

To investigate the gas sensing capability of polysilicon nanowires, conductance measurements were made sequentially in different gases. Figure 3 shows the measured current response as a function of time for 58 µm long nanowires, sequentially in air, vacuum and different pressures of H_2 (5%) gases. The nanowires were supplied with bias voltages of V_{DS} = 5V and V_{GS}= -10 V. A systematic change of the nano-sensor current is observed when ambient gases are changed. The change of ambient from air to vacuum increases the nanowire current by 45 % and the typical response time (time required to reach 90% of the equilibrium value) is 101 seconds. A decrease in the NW current is observed when nanowires are exposed into H_2 gases. Numerically 13%, 10% and 16% decrease from the base current level are observed for H_2 gas pressures 0.5, 1.0 and 1.5 bar. The response times are 100, 127 and 120 seconds for gas pressures 0.5, 1.0 and 1.5 bar respectively. The nanowire response in Isopropyl Alcohol (IPA) gas is shown in figure 4 where the IPA vapour ambient is created by applying a drop of the solvent (5μl) to a tissue above a hole in a measurement box. The same nano-sensor chip that was used for H_2 gas exposure test has been used in this experiment. A significant response is observed where the NW current is increased by 450 % from the low base current

around 417 nA when the NW is exposed to IPA vapour. The response time of NW to IPA vapour is found to be 15 seconds. When the IPA vapour source is removed, the current returns to the base current level in air with a recovery time around 32 seconds.

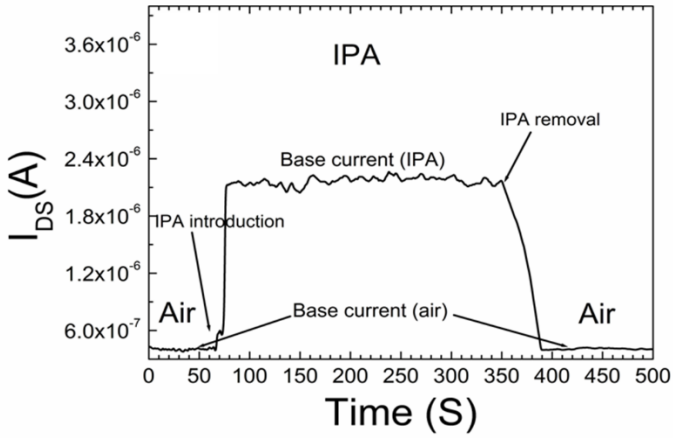


Fig. 4: The response of 58µm long polysilicon nanowire in Isopropyl Alcohol (IPA) gas.

The transfer characteristics of NW in air, vacuum, IPA and in H_2 gases at different pressures are shown in figure 5. During this measurement the substrate is used as gate and the 500 nm nitride layer is used as gate insulator. When the nano-sensor chip was exposed into different gas ambient, interesting changes in the sub-threshold characteristics can be observed. The sub-threshold slope of the nanowires remains unchanged after exposure into air, vacuum, H_2 gases which has a value around 2.6 V/dec. However, the change of ambient gas results in the shift of the sub-threshold characteristic. We have quantified this shift using the constant current threshold method at a current level of 4×10^{-8} A. A positive threshold voltage shift (ΔV_{TH}) of 2.47V is found when the ambient is changed from air to vacuum. However, the ambient change from vacuum to H_2 gas exhibits negative ΔV_{TH} shift with values of -1.14 V, -2.09 V and -3.16 V respectively for H_2 gases with pressures of 0.5, 1.0 and 1.5 bar. The introduction of the IPA gas results in the significant positive ΔV_{TH} of 10.48V with respect to air. It can also be observed that the ambient change associated nanowire's conductance change is strongly dependent on the substrate voltage (V_{GS}). The responses presented in figure 3 and 4 are measured for V_{GS} = -10 V which is actually in the saturation region of the characteristics and hence, there is plenty of room for response improvement by appropriate bias so that operating point falls inside the linear region of the transfer characteristics.

The nanowire conductance change (%) for different ambient as a function of substrate voltage (V_{GS}) is extracted through the numerical differentiation of figure 5 and presented in figure 6. Obviously, a phenomenological response of polysisilicon nanowire sensor can be obtained through appropriate bias which ensures the operating point to fall within the linear region of the sub-threshold regime. When the ambient is changed from air to vacuum, air to IPA and vacuum to H_2 gases with different pressures, promising conductance changes of 1560%, 3.6 × 10^5%, 70%, 90% and 97% can be obtained. These results are the testament of the gas sensing capability of polysilicon nanowire sensors at room temperature.

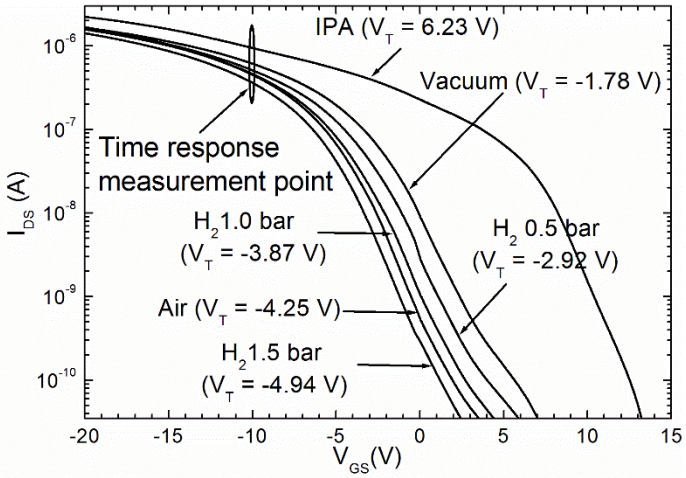


Fig. 5: The transfer characteristics of 58µm long polysilicon nanowire in air, vacuum, IPA and in H_2 gases at different pressures.

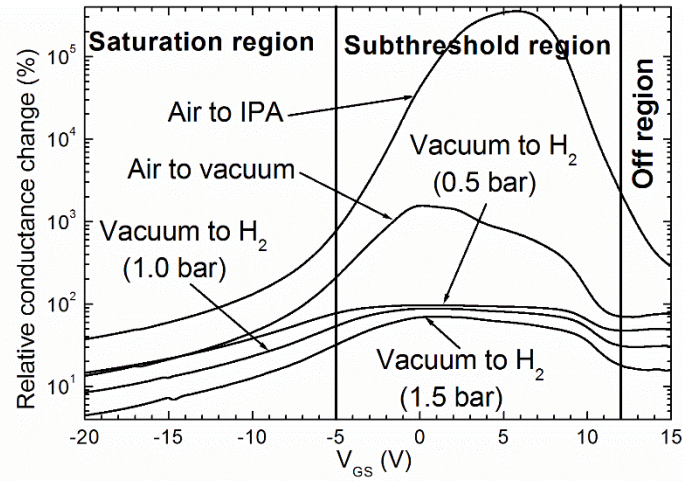


Fig. 6: The nanowire conductance change (%) for different ambient as a function of substrate voltage (V_{GS}) for 58µm long polysilicon nanowire.

Promising sensitivity of the fabricated polysilicon nanowire sensors for ambient like vacuum, H_2, volatile IPA gases are demonstrated in figure 3 to 6. It should be noted that the change of ambient from air to vacuum results in a significant increase in the response current of nanowires. Adsorption of gas molecules from atmosphere has been reported to influence the electrical properties of semiconductor nanostructures [19] which was found to be process dependent [19-21]. Reduction of nanowire conductance in p-type silicon nanowires from air to vacuum was observed by ref. [19]. Accumulation of holes in nanowire by adsorbed OH- ions was attributed to the enhancement of the nanowire conductance in air [19]. However, ref. [20] reported that the base conductivity of silicon wafer is not inherited in nanowires prepared by etch and p-type silicon nanowires showed significantly reduced conductivity in air due to surface charge associated depletion. Conductance increase of silicon nanowires in vacuum was also observed by ref. [21]. The fabricated polysilicon nanowires of this work show similar result like ref. [20-21] where reduced conductivity is observed in air due to surface charge associated depletion in air. As it is well known

that donor type states deplete the p-type nanowires by locally accumulating positive charges at the surface it can be decided that these atmosphere dependent surface charges in our work are most likely donor type.

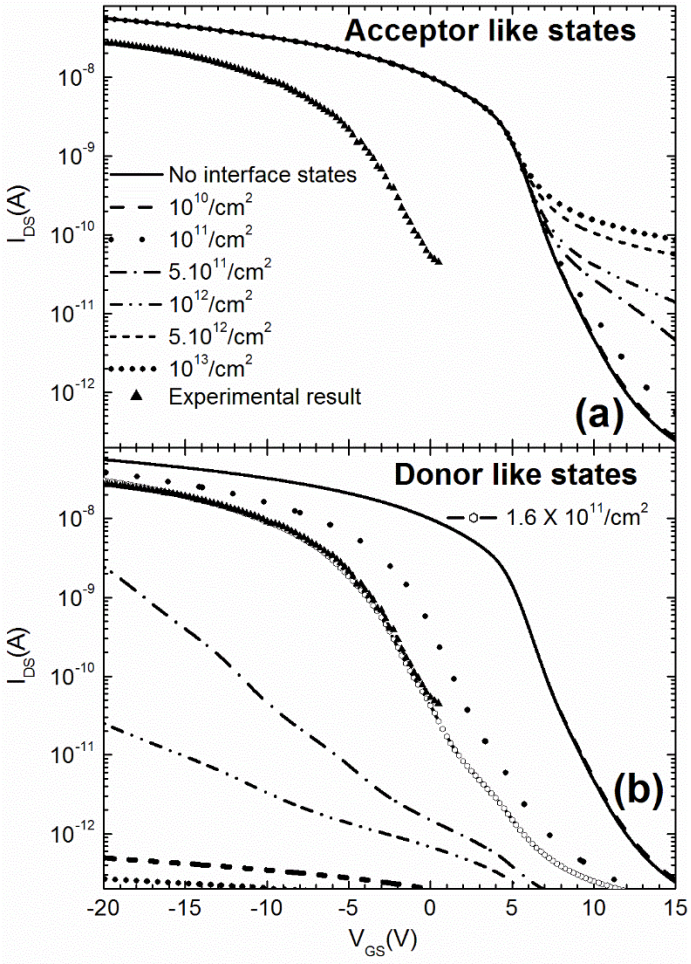


Fig. 7: The simulated and measured transfer characteristics for 14 µm long polysilicon nanowire in air for a nano-sensor block with different densities of a) acceptor like states and b) donor like states. V_{DS} value is 0.5V.

To get a deep insight on the surface/interface states of fabricated polysilicon nanowires in air, a 3D simulation of electrical characteristics of p-type polysilicon nanowires is done. For modelling polysilicon nanowires accurately, continuous trap states (both donor-like and acceptor-like) distribution model is used with parameters previously used in the literature for successful modelling of polyslicon thin film transistors [18]. For interface states, acceptor or donor type states with varying densities are invoked both in the top oxide/polysilicon and bottom nitride/polysilicon interface and tested. The simulated and measured transfer characteristics (I_{DS} as a function of V_{GS} for V_{DS} value of 0.5V) of the 14 µm long nanowire with different densities of acceptor like states is shown in figure 7a. It can be observed in figure 7a, acceptor like surface states and the reported continuous distribution of polysilicon defect states [18] are not enough to reproduce electrical characteristics of our polysilicon nanowires in air. The leakage current of p-type polysilicon nanowire is only affected by acceptor like surface states. The simulated and measured transfer characteristics of the 14 µm long nanowire with different densities of donor like states is shown in figure 7b. Obviously, polysilicon nanowire's sub-threshold and on current characteristics both are affected by donor type states. Acceptor like states are located at the upper half of the band gap whereas donor like states are located at the lower half of the band gap thereby showing different behaviour in different operation regime depending on the position of the fermi level.

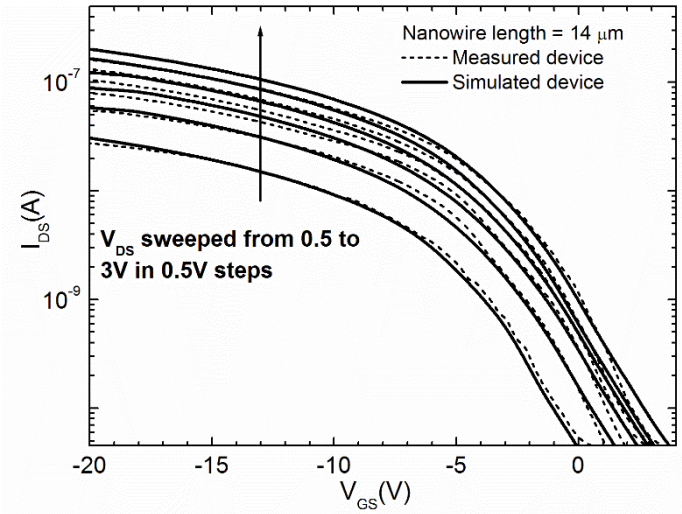


Fig. 8: Simulated and measured transfer characteristics for 14 µm long polysilicon nanowire in air. V_{DS} value is 0.5V.

The measured transfer characteristics (I_{DS} as a function of V_{GS}) for different V_{DS} values for the 14 µm long nano-sensor block is compared with simulation in figure 8. We are able to reproduce the measured sub-threshold characteristics of 14 µm long nanowires in air for the surface states density of $1.6 \times 10^{11}/cm^2$. This implies that the surface states of the fabricated polysilicon nanowires excluding grain boundary defects is around ~ $10^{11}/cm^2$. Similar results are reported for single crystal silicon nanowires reported by Jie et al ($4.3\times10^{11}/cm^2$) [19].

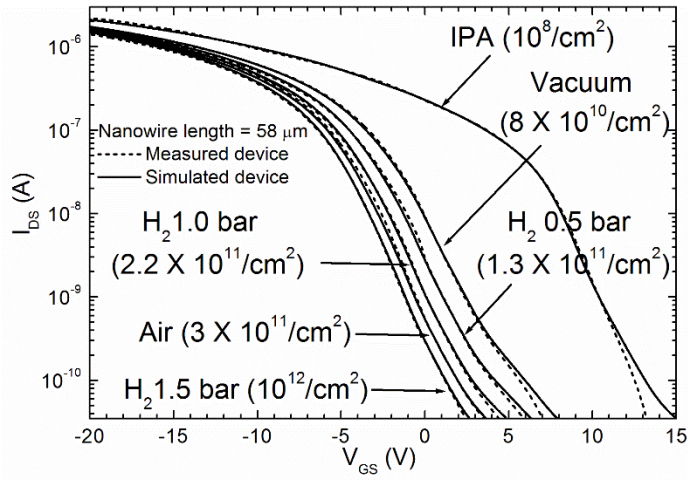


Fig. 9: Comparison of the measured transfer characteristics of the nano-sensor in different gas ambient with the simulated transfer characteristics for 58 µm long nanowires with different densities of surface states. V_{DS} value is 0.5V.

Figure 9 compares the measured sub-threshold characteristics of the nano-sensor in different gas ambient with the simulated transfer characteristics (I_{DS} as a function of V_{GS}) with different densities of surface states for 58 μm long nanowires. The donor type states in the top oxide/silicon interface for a density of $3 \times 10^{11}/cm^2$ is found to reproduce the measured sub-threshold characteristics in air which is similar to the density of $1.6 \times 10^{11}/cm^2$ that was used to reproduce the measured sub-threshold characteristics of 14 micron long nanowire in air (figure 8). This result implies that excluding the defect states contribution from grain boundaries, the surface states density of the fabricated polysilicon nanowires is around ~ $10^{11}/cm^2$. Donor type states with densities of $8 \times 10^{10}/cm^2$, $1.3 \times 10^{11}/cm^2$, $2.2 \times 10^{11}/cm^2$, $10^{12}/cm^2$ are able to reproduce the sub-threshold characteristics in vacuum and H_2 gases with pressures of 0.5, 1.0 and 1.5 bar respectively. Based on this result it can be concluded that $5 \times 10^{10}/cm^2$, $1.4 \times 10^{11}/cm^2$ and $9.2 \times 10^{11}/cm^2$ number of states are introduced when the ambient is switched from vacuum into H_2 gases with pressures of 0.5, 1.0 and 1.5 bar respectively and $2.2 \times 10^{11}/cm^2$ number of states is removed when the ambient is changed from air into vacuum. Donor type density of $1 \times 10^8/cm^2$ is able to reproduce the transfer characteristics in IPA, indicating that almost all of the surface states in air is effectively removed when the NW is exposed into IPA vapour. This result indicates the complex mechanism of IPA interaction with NWs that is reducing surface states and effectively working as passivation agent for NW surface states [22].

Table I

Comparison of the gas sensing figure of merit (i.e. conductance change, response/recovery times) of the fabricated polysilicon nanowire sensors with the literature.

Si NW size	Type/approach	Target	Conductance change	Response time	Recovery time	Ref.
Inorganic gas						
NA	Porous single crystal Si	H_2	18% at 5% H_2	-	-	23
			50% at 10% H_2	-	-	
d = 300 nm L = 5.25 μm	Single crystal Si NW synthesized by metal assisted chemical etch	H_2	60% at pure H_2	180 seconds	No recovery	24
d = 95 nm L = 58 μm	**Poly silicon nanowire by spacer etch (This work)**	**H_2**	**5% H_2 in air**	**100-120 seconds**	**101 seconds**	**This work**
			16% at 0.5 bar 32% at 1.0 bar 65% at 1.5 bar			
			5% H_2 in vacuum			
			70% at 0.5 bar 90% at 1.0 bar 97% at 1.5 bar			
Polar volatile organic compounds						
d = 30 to 70 nm L = 400 nm to 1 μm	Top down/bottom up grown single crystal Si NWs	Polar VOC (Methanol, Ethanol, Isopropanol)	50%	40 seconds	50 seconds	25
d = 95 nm L = 58 μm	**Poly silicon nanowire by spacer etch (This work)**	**Polar VOC (IPA)**	**3.6×10^5%**	**15 seconds**	**32 seconds**	**This work**

Table I compares the gas sensing figure of merit (i.e., response/recovery times and conductance change) of our polysilicon nanowire sensors with the only few reported silicon-based gas sensors for H_2 and polar volatile organic compound (i.e. IPA). To judge the capability of our nano-sensor to detect H_2 and IPA molecules fairly, we only consider pure silicon nanowires without any surface modifications. As can be seen from the table our polysilicon nanowire sensors exhibit much better response/recovery times in comparison to reported silicon based H_2/polar volatile organic compound sensors and demonstrate appreciable conductance change for H_2/IPA.

III. CONCLUSIONS

We have presented a thin film polycrystalline silicon nanowire sensor technology and have demonstrated room temperature gas detection using polysilicon nanowires. The nanosensors have been defined by spacer etch using low density plasma of SF_6/ O_2 gases from thin α-Si layer in a low cost reactive ion etch (RIE) equipment and well established 3 micron lithography. Comparative measurements in air, vacuum, hydrogen and isopropyl alcohol have demonstrated the potential of polysilicon nanowires as gas sensors in room temperature with remarkable gas sensitivity and impressive response/recovery times that are noticeably better than reported pristine silicon based nano-sensors. 3D numerical simulations of polysilicon nanowires have indicated the evolution of surface states in different gas ambient which is the testament of gas sensing ability of the polysilicon nanowires. The developed thin film polysilicon nanowire sensors show route towards simple, low cost, mass manufacturable nano-sensor fabrication which is compatible with glass or plastic substrates could open up opportunities in low cost disposable diagnostics and in portable, wearable or even implantable sensors.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support of the Center for Research & Training (CRT), East West University, Dhaka-1212, Bangladesh for funding this research Special thanks to clean room engineers of Southampton Nano-fabrication Center for assisting the fabrication/characterisation of nano-sensor chips and East West University Simulation Lab for providing TCAD simulation facility of nano-sensors.

REFERENCES

[1] M. E. Farnke, T. J. Koplin and U. Simon, "Metal and metal oxide nanoparticles in chemiresistors: does the nanoscale matters?" Small, vol. 2, pp. 36-50, 2006.

[2] S.A. Kharitonov and P.J. Barnes, "Exhaled markers of pulmonary disease," Am J Respir Crit Care Med., vol. 163, pp. 1693-1722, 2001.

[3] S.M. Studer et al., "Patterns and significance of exhaled-breath biomarkers in lung transplant recipients with acute allograft rejection," J. Heart and Lung Transplant, vol. 20, pp. 1158- 1166, 2001.

[4] T. H. Risby and S. S. Sehnert, "Clinical application of breath biomarkers of oxidative stress status," Free Radical Biol. Med., vol. 27, pp. 1182-1192, 1999.

[5] T. Karl et al., "Human breath isoprene and its relation to blood cholesterol levels: new measurements and modeling," J. Appl. Physiol., vol. 91, pp. 762-770, 2001.

[6] M. Phillips, J. Herrera, S. Krishnan, M. Zain, J. Greenberg, and R. N. Cataneo, "Variation in volatile organic compounds in the breath of normal humans," Journal of Chromatography B: Biomedical Sciences and Applications, vol. 729, no. 1-2, pp. 75–88, 1999.

[7] N. Bârsan et al., "Fundamental and practical aspects in the design of nanoscaled SnO_2 gas sensors: a status report," Fresenius' Journal of Analytical Chemistry, vol. 365, pp. 287–304, 1999.

[8] J. Huang and Q. Wan, "Gas sensors based on semiconducting metal oxide one-dimensional structure," Sensors, vol. 9, pp. 9903-9924, 2009.

[9] Rakesh K. Joshi and Ashok Kumar, "Room temperature gas detection using silicon nanowires," Materials Today, vol. 14, pp. 522011, 2011.

[10] D. P. Tran, T. T. T. Pham, B. Wolfrum, A. Offenhäusser and B. Thierry, "CMOS-Compatible Silicon Nanowire Field-Effect Transistor Biosensor: Technology Development toward Commercialization," Materials, ,vol. 11, pp. 785: 1-26, 2018.
[11] M. C. McAlpine, R. S. Friedman, S. Jin, K. H. Lin, W. U. Wang and C. M. Lieber, "High-Performance Nanowire Electronics and Photonics on Glass and Plastic Substrates," Nano Lett, vol. 3, pp. 1531-1535, 2003.
[12] Z. Q. Gao, A. Agarwal, A. D. Trigg, N. Singh, C. Fang, C. H. Tung, Y. Fan, K. D. Buddharaju, J. M. Kong, "Silicon Nanowire Arrays for Ultrasensitive Label-Free Detection of DNA," Anal. Chem., vol. 79, pp. 3291-3297, 2007.
[13] I. Park, Z. Li, A. P Pisano and R. S. Williams, "Top-down fabricated silicon nanowire sensors for real-time chemical detection," Nanotechnology, vol. 21, pp. 015501, 2010.
[14] C. Y. Hsiao, C. H. Lin, C. H. Hung, C. J. Su, Y. R. Lo, C. C. Lee, H. C. Lin, F. H. Ko, T. Y. Huang and Y. S. Yang, "Novel poly-silicon nanowire field effect transistor for biosensing application," Biosensors and Bioelectronic, vol. 24, pp. 1223-1229, 2009.
[15] C.-H. Lin, C.-Y. Hsiao, C.-H. Hung, Y.-R. Lo, C.-C. Lee, C.-J. Su, H.-C. Lin, F.-H. Ko, T.-Y. Huang and Y.-S. Yang, "Ultrasensitive detection of dopamine using a polysilicon nanowire field-effect transistor," Chem. Commun., vol. 44, pp. 5749-5751, 2008.
[16] You-Lin Wu et al., "Polysilicon Wire for the Detection of Label-Free DNA," Journal of the Electrochemical Society, vol. 157, pp. J191-J195, 2010.
[17] Silvaco international, *Atlas User's Manual Device Simulation Software*, Silvao International Ltd., Santa Clara, Dec. 2002.
[18] P. Walker et al., "Improved off-current and subthreshold slope in aggressively scaled poly-si TFTs with single grain boundary in the channel," IEEE Trans. On Electron Devices, vol. 51, pp. 212-219, 2004.
[19] Jiansheng Jie, Wenjun Zhang, Kuiqing Peng, Guodong Yuan, Chun Sing Lee, and Shuit-Tong Lee, "Surface-Dominated Transport Properties of Silicon Nanowires,"Advanced Functional Materials, vol. 18, pp. 3251-3257, 2008.
[20] S. Chen, J. G. Bomer, W. G. Van der Wiel, E. T. Carlen and A. Van den Berg, "Top-down fabrication of sub-30 nm monocrystalline silicon nanowires using conventional microfabrication,"ACS Nano, vol. 3, pp. 3485-3492, 2009.
[21] Hideo Fujii, Seigo Kanemaru, Takashi Matsukawa, and Junji Itoh, "Air-bridge-structured silicon nanowire and anomalous conductivity," Applied Physics Letters, vol. 75, pp. 3986-3988, 1999.
[22] J. K. El-Demellawi et al., "Room-Temperature Reactivity Of Silicon Nanocrystals With Solvents: The Case Of Ketone And Hydrogen Production From Secondary Alcohols: Catalysis?,"Applied Materials & Interface, vol. 7, pp. 13794-13800, 2015.
[23] N.K. Ali, M. R. Hashim, A. Abdul Aziz, "Effects of surface passivation in porous silicon as H_2 gas sensor," Soild-State Electronics, vol. 52, pp. 1071-1074, 2008.
[24] Hangbo Zhao, "Vertical silicon nanowire arrays for gas sensing," MSc Thesis, Massachusetts Institute of Technology, USA, February, 2014.
[25] A. O. Niskanen, A. Colli, R. White, H. W. Li, E. Spigone and J. M. Kivioja, "Silicon nanowire arrays as learning chemical vapour classifiers," Nanotechnology, vol. 22, pp. 295502, 2011.

Design and Analysis of high Efficient Latched Comparator

Dona Basak, Satyendra Nath Biswas*
Department of Electrical & Electronic Engineering
Ahsanullah University of Science & Technology (AUST)
141 & 142, Love Road, Tejgaon Industrial Area, Dhaka-1208, Bangladesh
(*Corresponding Author: sbiswas.eee@aust.edu)

***Abstract*—Keeping pace with the new era, the necessity for comparator circuit with low power and high speed profile is increasing. In this research, a new latched comparator is proposed and compared with the conventional model. The proposed circuit is designed using PTM model. This circuit is proposed to overcome the difficulties mentioned in the conventional circuits. Different feature sizes (180nm and 32nm) transistor are considered. We have calculated power and delay and they are compared with previously proposed circuits for demonstrating the efficiency of the proposed model. Its power consumption is low compared with others and delay is decreasing with feature size. Also it has higher resolution and load drivability than those conventional approaches.**

Index Terms*— *Dynamic latched comparator, Feature size, Common mode voltage, Positive feedback mechanism.

I. INTRODUCTION

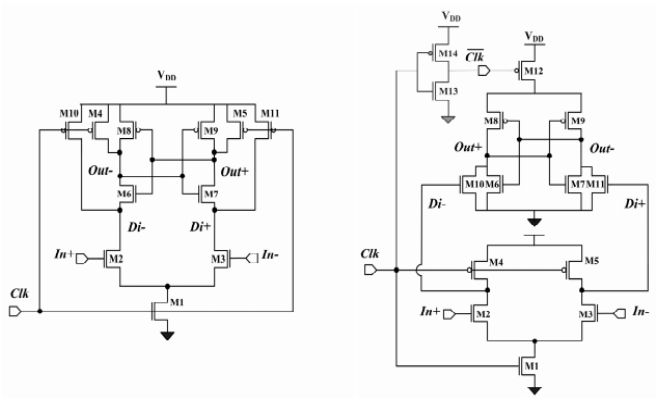
Comparator is one of the most widely used electronic devices after operational amplifiers in this modern era. Comparators are known as 1-bit analog-to-digital converter and for that reason they are mostly used in large abundance in analog to digital converter. Due to fast-speed, high input impedance and full-swing output, CMOS dynamic latched comparators are very attractive for many applications such as zero-crossing detectors, peak detectors, switching power regulators, analog to digital converters, memory sense amplifier, data transmission, and others. Nowadays, where demand for portable battery operated devices is increasing, a major emphasis is given towards low power methodologies for high speed applications. Also we have to minimize the power consumption by using smaller feature size processes. However, when we move towards power minimization, the process variations and other parameters will greatly affect the overall performance of the device. One such application where low power dissipation, low noise, high speed, less offset voltage are required is Analog to Digital converters for mobile and portable devices[3].

Dynamic comparators use positive feedback mechanism with one pair of back-to-back cross coupled latch in order to convert a small input-voltage difference to a full-scale digital level within a short time [1]. However, it has a larger input referred offset compared to linear amplifiers. The reason is that in addition to static V_{th} and β mismatch, it also suffers from dynamic offset due to imbalance of parasitic capacitors at internal nodes during evaluation [2]. A CMOS latch-type voltage sense amplifier was designed with a separated input and cross-coupled stage [7]. This topology has less stacking and can therefore operate at lower supply voltages. The double tail enables both a large current in the latching stage for fast latching independent of the V_{cm} and a small current in in the input stage for low offset .

Practically, there exist many papers, where some parameters of comparator such as speed, power consumption, offset voltage minimization etc. are analysed. Dynamic latched comparator and double tail comparators are modified by adding few transistors so that the positive feedback during the regeneration is strengthened, which results in remarkably reduced delay time. The simulated result is considered on 180nm CMOS technology with supply voltage of 1.2V and 0.6V [4]. However, input-referred offset voltage of 7.8mv and average power dissipation of 329µw at 500MHz frequency is found that is comparatively higher. Therefore, offset voltage compensation is very important to improve the performance of dynamic comparator. If the voltage gain of regenerative latch can be augmented to 22 V/V, it can demonstrate up to 24.6%less offset voltage and 30% less sensitivity of delay to decrease input voltage difference (17 ps/decade) than the conventional double tail latched comparator for same area and power consumption [6].In order to minimize input referred offset voltage a number of transistors are used that made the circuit more complex. Again, a comparator with a modified latch is proposed [7], which is different from the conventional circuit. This circuit shows an effective result at low power supply voltage (0.65V) as well as the conventional power supply (1V or 1.2V). Though the paper has shown that it gives rail to rail output, we have found a different result by simulation in LTspice. It can even detect a change of input signal within clock pulse but can not give rail to rail output as the capacitors take time to charge or discharge. However, we need an optimum design where power, speed, offset and size altogether gives an efficient result [5].

In this paper, a new dynamic latched comparator is proposed and simulated with various feature size. The circuit is simulated using LTspice at 180nm and 32nm PTM technology[9] for high performance application and 0.8V power supply. Following parameters are used for simulation using LTspice. PMOS parameters are: v_{thp}=-0.42V and v_{thn}=0.3999V for 180nm, v_{thn}=-0.4939V and v_{thp}=-0.4915V for 32nm. Clock frequency is considered as 0.33GHz for 180nm and 2.4GHz for 32nm because 180nm technology does not show desired result at high frequency such as 2.4 GHz .

II. CONVENTIONAL MODELS



(a) **(b)**

Fig 1(a) Conventional dynamic latched comparator [10]; (b) Double-tail dynamic comparator [8]

Though there are several advantages such as fast speed, ideally zero static power consumption, high input impedance and full-swing output, the dynamic latched comparator shown in Figure 1(a) has been most widely used [10]. However, in order to increase the drive current of the latch, it is inevitable to size up the tail transistor M1 since this comparator has only one tail transistor. If the size of tail transistor is increased, the drain currents of both input transistors M2 and M3 will increase during the evaluation phase (*clk*=V_{DD}). This means the reduction of the time duration for transistor M2 and M3 being operated in saturation region. *Di* nodes discharge from V_{DD} to ground in a very short period. Consequently, lower amplification of the input voltage difference will be made. And such a small *Vth* variation from mismatch between transistor M6 and M7 can yield high input-referred offset. In addition, since it shows the strong dependency on speed and offset with different common-mode input voltage V_{com} [10], it is less attractive in applications with wide common-mode ranges such as ADCs [8].

In order to circumvent these drawbacks, the comparator with separated input-gain stage and output-latch stage shown in Figure 1 (b) was introduced in [8]. This type of design made it possible for this comparator to operate at a lower supply voltage (V_{DD}) and have more stable offset voltage over wide common-mode voltage (V_{cm}) ranges. However, since it requires both *clk* and *clkb (inverted clk)* for its operation, high accuracy of *clkb* is required. Because the output-latch has to detect the voltage difference at *Di* nodes at very limited time. If a simple inverter replaces the *clkb*, *clk* has to be able to drive an additional large inverter (heavier clock load) in order to drive the largest transistor M12 in a small delay. If *clkb* is lagging the *clk*, it results in increased delay and if *clkb* is leading the *clk*, it results in increased power dissipation due to existing the short circuit current path from M12 to M10/M11 though M8/M9. And it can even increase the latch offset voltage if the device mismatch between M8 and M9 is significant. The comparator from [11] without offset calibration technique, shown in Figure 2, resolved the problem by replacing *clkb* with *Di* nodes. As a result, *clk* load was lessened and the input referred offset was reduced.

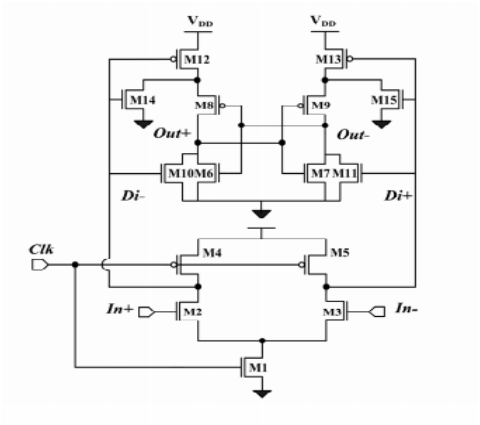


Fig 2: Dynamic comparator[11]

However, the improved offset has to trade off with the increased delay since the current drivability of the output load was weakened due to the fact that transistor M12 and M13 use *Di* node voltages as their *clkb* signal. That show slow exponential decay shape, and that the maximum drive current of each *Out* node was reduced to half of the single output tail current of M12 comparing with Comparator 1, since it was separated into two transistors M12 and M13

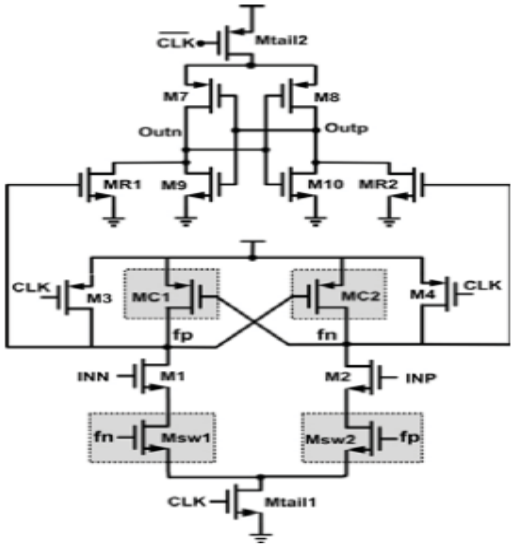


Fig 3: Dynamic comparator with reduced delay[4]

The above mentioned circuit of figure 3, improves the speed and reduces the delay and static power consumption of conventional dynamic and double tail comparator by adding two control transistor MC1 and MC2 as cross coupled inverter in input section. Two NMOS switches below two input transistors (MC1 and MC2), as a latch structure increases the voltage difference between fp and fn nodes within a very short time and passes that difference to intermediate stage transistors MR1 and MR2. This phenomenon furthermore improves the regeneration time of output latch stage, thus speed improves. It also enhances the effective transconductance (g_{meff}) of the latch, and delay is reduced. On the other hand, two NMOS switches M_{sw1} and M_{sw2} are included in order to avoid static power consumption due to current flow from V_{DD} to ground via two input transistors. However, it has shortcomings in case of offset mismatch and

kickback noise. When the input differential voltage is very small, threshold voltage and current factor mismatch of controlled transistors augment input referred offset voltage. Kickback noise is also increased in comparison with conventional double tail structure.

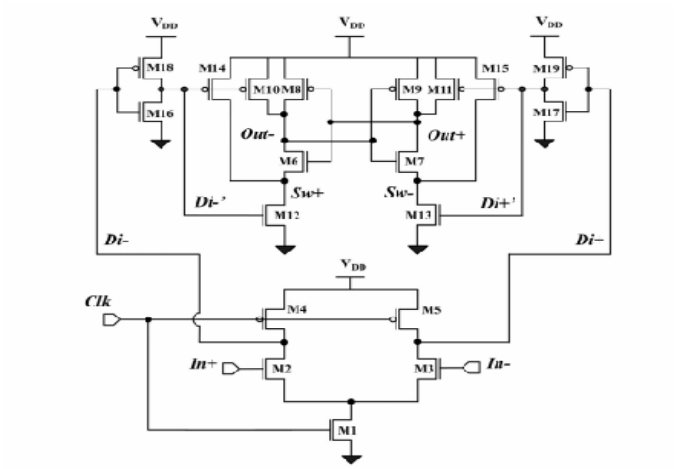


Fig 4: Dynamic comparator with reduced offset[6]

To resolve the problem of offset voltage another comparator circuit is proposed [6]. The given circuit offers lower input referred offset voltage and faster operation at the same area and power consumption. Two inverters are included between input and output stage for amplification of weak difference signal so that input referred offset voltage is attenuated by the voltage gain of that stage. On the other hand, pre-charge state voltage mismatch is overlooked by using M14 and M15 transistors which reset sw nodes. Therefore, V_{th} mismatch between M6 and M7 is removed. It has many stacked transistors and higher load drivability. The circuit was simulated on HSPICE on 90nm PTM technology with V_{DD}=1V.

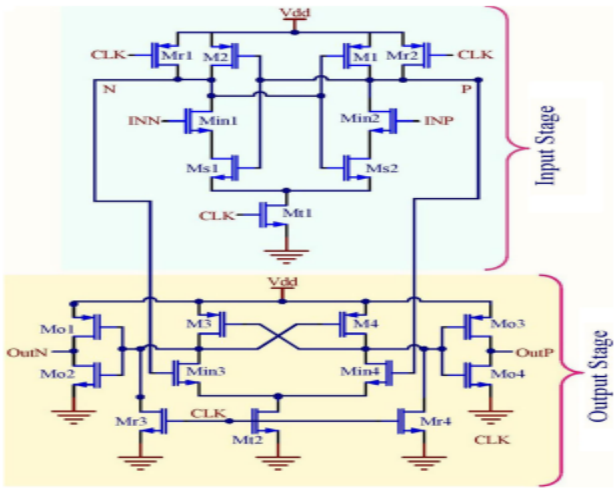


Fig 5: Dynamic comparator with optimized design[5]

To achieve a best possible design for double tail latch comparator power, speed, offset and size are optimized. The circuit is simulated on 180nm and power, delay is indicated. It depends on discharging of output capacitance to indicate rail to rail output swing.

III. PROPOSED DYNAMIC COMPARATOR

Fig .6 demonstrates the schematic diagram of the proposed dynamic comparator. During pre-charge (reset) phase when Clk = 0V, tail transistor M_{t1} and M_{t2} remains off while two PMOS transistors P3,P6 becomes on .It makes N4 and N3 switch on to pull both of the output to zero. During evaluation (decision making) phase when ClK=1, both M_{t1} and M_{t2} turns on and P3,P6 turns off. As a result, the capacitor at the gate of N3 and N4 starts to discharge according to the difference in input. Consider $V_{INP}>V_{INN}$, thus the drain of N1 discharges faster than that of N4.When drain voltage of N1 drops below V_{thn} the corresponding NMOS transistors N3 turns off gradually. At that time, voltage of OutN is still zero, so it turns on P1,makes Outp towards V_{DD}. On the other hand, N4 remains on before the gate volage falls off its threshold level. When the gate voltage of N4 drops below v_{thn}, N4 turns off. P2 was off before, therefore, the capacitor at OutN holds it to ground while OutP remains at V_{DD}. The case is vice versa for $V_{INP}<V_{INN}$.

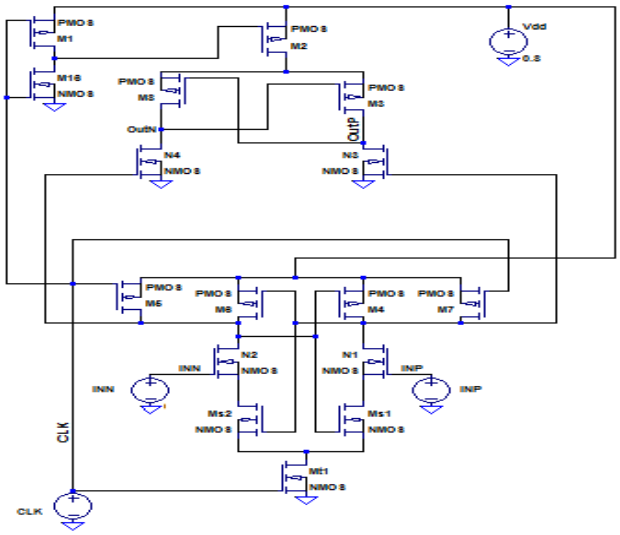


Fig 6: Schematic diagram of proposed comparator at 180nm CMOS Technology

IV. SIMULATION RESULTS

The circuit is simulated on LTspice using 180nm CMOS technology modeled with parameter (v_{thp}=-0.42V and v_{thn}=0.3999V. Fig. 7 is the simulated waveform for proposed circuit at 180nm CMOS technology. The same circuit is simulated with 32nm and the wave characteristics are shown in figure 8. Also, Fig. 3,Fig. 4 and Fig. 5 are simulated at 32nm and compared with proposed one.

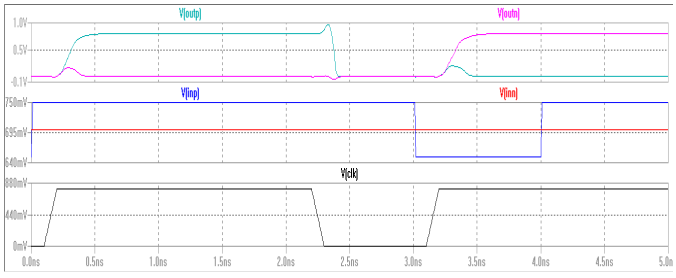


Fig 7:Simulation result of Fig .6 (ΔV_{in}=50mv{from INN to highest peak of INP},f_{clk}=0.33GHz at 180nm CMOS technology)

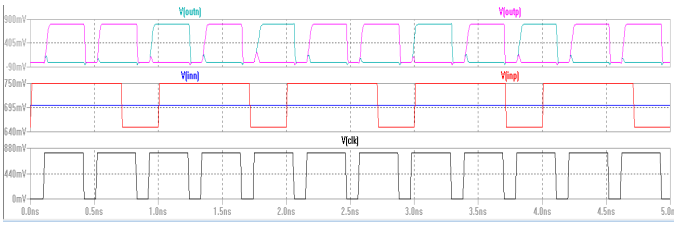


Fig 8:Simulation result of Fig .3 for 32nm CMOS technology at 2.4GHz

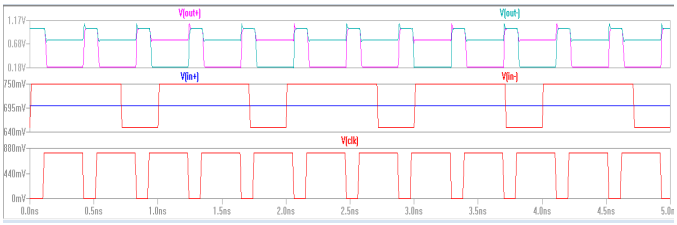


Fig 9:Simulation result of Fig .4 for 32nm CMOS technology at 2.4GHz

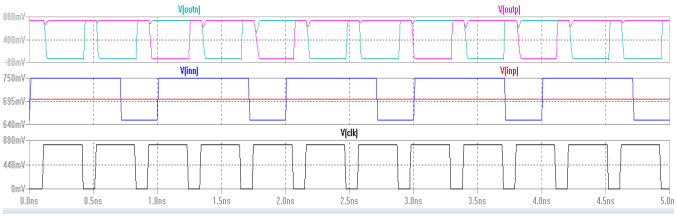


Fig 10:Simulation result of Fig .5 for 32nm CMOS technology at 2.4GHz

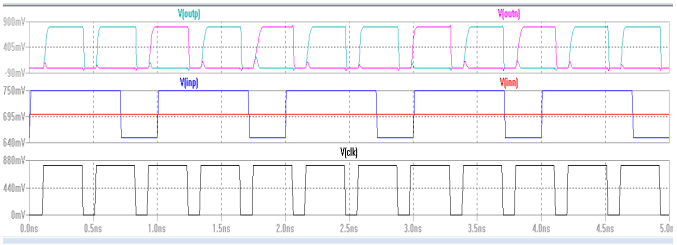


Fig 11:Simulation result of Fig .6 for 32nm CMOS technology at 2.4GHz

V. PERFORMANCE COMPARISON

In order to compare the proposed comparator with the referenced comparator given [4, 5, 6], the other circuits have been simulated in 180nm CMOS technology with V_{DD}=0.8V except Fig. 4. where supply is 1V. This is because, for Fig. 4 if the supply voltage becomes low, the top level of output alienate more from the V_{DD} level, which is not desired. Now, consider the circuits for 180nm.Maximum clock frequency is proportional to supply voltage. If the supply voltage is high ,we can use high clock frequency or vice versa. i.e. for Fig. 3 if V_{DD} is 1.2 V ,clock frequency can be elevated upto 2.5GHz. For V_{DD}=0.8V, here in Fig. 3 the maximum clock frequency is 1.69GHz.The same result is for proposed circuit where for Fig. 5 it is 1.66GHz. For Fig. 4 clock frequency is 3GHz with V_{DD}=1V. But, output can not reach to V_{DD} as high level. We need to enhance that frequency in order to get a good digital output. On the other hand, the ac performance tends to degrade if the number of transistor becomes larger due to parasitic capacitances. In our proposed circuit the number of transistor also has lessen compared to others. Time delay and power dissipation is also shown in Table I and it is clearly visible that for 180nm technology, our proposed design shows better performance.

Table II shows the same features for 32nm technology as we are moving towards smaller feature size. Clock frequency can be increased and we have used 2.4GHz. Power consumption and delay are analyzed with two threshold voltages to observe the variation. And again our proposed one gives less delay and power performance.

TABLE I. PERFORMANCE COMPARISON WITH REFERENCED COMPARATORS AT 180NM

Comparator structure	Ref. [4]	Ref. [5]	Ref. [6]	Proposed Comparator
Technology	180nm	180nm	90nm	180nm
No. of transistor	16	20	19	14
Supply voltage(V)	0.8V	0.8V	1V	0.8V
Clock frequency	Max 1.69GHz at 0.8V	1.66 GHz	3GHz	Max 1.69GHz At 0.8V
Delay at ΔV_{in}=50mv	Rise time:172.3ps	Fall time:49ps	Delay[ps]/log(ΔV_{in})=17	Rise time:97.41ps
Power dissipation at 0.33GHz and V_{cm}=0.7V	2.5908µW for simulation time 5ns	11.862µW for simulation time 5ns	162 µW at 3GHz and V_{cm}=0.6V	2.2929µW for simulation time 5ns

TABLE II. PERFORMANCE COMPARISON WITH REFERENCED COMPARATORS AT 32NM FOR 5NS SIMULATION TIME

Comparator structure	Dynamic Comparator[4]		Dynamic Comparator[5]		Dynamic Comparator[6]		Proposed Comparator	
	V_{thn}= 0.49V V_{thp}= -0.49V	V_{thn} = 0.35 V V_{thp} =-0.24 V	V_{thn}= 0.49V V_{thp}=-0.49V	V_{thn}= 0.35V V_{thp}= -0.24V	V_{thn}= 0.49V V_{thp}= -0.49V	V_{thn}= 0.35V V_{thp}= -0.24V	V_{thn}= 0.49V V_{thp}= -0.49V	V_{thn}= 0.35V V_{thp}= -0.24V
Technology	32nm		32nm		32nm		32nm	
Supply voltage (V)	0.8V		0.8V		1V		0.8V	
Clock frequency	2.4GHZ		2.4GHZ		2.4GHZ		2.4GHZ	
Delay at ΔV_{in}=50 mv	Rise time: 29.7 ps	Rise time 11.25 ps	Fall time: 12.34ps	Fall time 4.5 ps	Fall time: 11.37 ps	Rise time: 6.73 ps	Rise time: 28.10 ps	Rise time 7.86 ps
Power dissipation at 2.4GHz, V_{cm}= 0.7V	3.2542µW	2.6798µW	22.023 µW	26.11µW	54.27 µW	40.35 µW	3.0883 µW	2.3688µW

VI. CONCLUSION AND FUTURE WORK

In this paper, comprehensive analysis for clocked dynamic comparator is presented and analysed for various feature size. The present research is mainly concentrated on power consumption and delay of the proposed dynamic comparator. The comparator works perfectly according to the requirement and demonstrates better performance as compared to other referenced models. Moreover, the size of CMOS transistors as well as threshold value are varied and analysed their effect on the performance. The analysis shows clearly that the proposed

circuit reduces both the time delay and power dissipation in comparison with referenced others.

ACKNOWLEDGMENT

The first author would like to thank her parents. They also would like to thanks to Department of EEE of AUST to give the chance of doing analysis on it and encourage for doing further research work.

REFERENCES

[1] Shashank Shekhar, Dr. S.R.P Sinha. "Design and Analysis of Dynamic Comparator with Reduced Power and Delay" *International Journal of Science and Research(IJSR)*, vol. 4, Issue 11, November 2015.

[2] Jun He, Sanyi Zhan, Degang Chen, and R.L. Geiger, "Analyses of Static and Dynamic Random Offset Voltages in Dynamic Comparators," IEEE Trans. Circuits Syst. I: Reg. Papers, vol. 56, pp. 911-919, May 2009.

[3] Smriti Shubhanand, Dr.H.P. Shukla ,"Design and Simulation of a High Speed CMOS Comparator", International Journalof Electronics and Communication Engineering Vol. 6, Number 1 (2013), pp. 75-80..

[4] Samaneh Babayan-Mashhadi,Reza Lotfi,"Analysis and Design of a Low-Voltage Low-Power Double-Tail Comparator, " IEEE Trans. On very large scale integration(VLSI) systems, January 15,2013.

[5] Ehsan Yaquibi, Seyed Hamid Zahiri, "Optimum design of a double-tail latch comparator on power,speed,offset and size " in Analog Integr Circ Sig Process, DOI 10.1007/s10470-016-0903-1, December 21,2016

[6] HeungJun Jeon, Yong -Bin Kim,"A novel low- power,low-offset. And high-speed CMOS dynamic Latched comparator,Analog Integr. Circ. Sig. Process, DOI 10.1007/s10470-011-9687-5,July 17,2011.

[7] B. Goll and H. Zimmermann, "A comparator with reduced delay time in 65-nm CMOS for supply voltages down to 0.65," *IEEE Trans. Circuits Syst. II, Exp. Briefs,* vol. 56, no. 11, pp. 810–814, Nov. 2009.

[8] D. Shinkel, E. Mensink, E. Klumperink, E. van Tuijl, and B. Nauta,"A double-tail latch-type voltage sense amplifier with 18ps Setup+Hold time," in *Proc. IEEE Int. Solid-State Circuits Conf., Dig. Tech.Papers,* Feb. 2007, pp. 314–315 .

[9] http://www.eas.asu.edu/ptm/latest.htm

[10] B. Wicht, T. Nirschl, and D. Schmitt-Landsiedel, "Yield and speed optimization of a latch-type voltage sense amplifier," *IEEE J. Solid-State Circuits*, vol. 39, pp. 1148-1158, July 2004

[11] M. Miyahara, Y. Asada, P. Daehwa and A. Matsuzawa, "A Low-Noise Self-Calibrating Dynamic Comparator for HighSpeed ADCs," *in Proc. A-SSCC*, pp. 269-272, Nov. 2008.

Performance comparison between Graphene Nano-Ribbon FET & conventional CMOS based on Arithmetic Logic Unit (ALU)

Md. Omar Faruque,Department of Electrical and Electronic Engineering , Ahsanullah University of Science and Technology , mohammadomarfaruque584@gmail.com

Tanvir Hasan, Department of Electrical and Electronic Engineering , Ahsanullah University of Science and Technology , tanvirhasan334@gmail.com

Safayat Al Imam, Department of Electrical and Electronic Engineering , Ahsanullah University of Science and Technology ,imam.safayat@gmail.com

***Abstract*—. This paper presents the implementation and analysis of GNRFET and Conventional CMOS based 1-bit and multi-bit Arithmetic Logic Unit (ALU). The simulation results calculated using HSPICE have reported to show 2.79% reduction in Propagation Delay (PD), 57.59% reduction in Average Power Consumption, 61.98% reduction in Power Delay Product (PDP) and 66.44% reduction in Energy Delay Product (EDP) for GNRFET base 1-bit ALU as compared to Conventional CMOS based 1-bit ALU at 32nm technology**

Keywords—GNRFET, Power delay product, Energy delay product

I. INTRODUCTION

The advancement in IC technology rendering area optimized and fast ICs is primarily attributed to the MOSFET scaling theory. As the process technology reached the nanometer regime, silicon CMOS started developing short-channel effects which led to increased power dissipation. Much research has been conducted to find a potential replacement for Conventional CMOS. Graphene Nano-Ribbon Field-Effect Transistor (GNRFET) proved to be a potential replacement to traditional CMOS. Thus, analysis of the circuits built using GNRFET is necessary to demonstrate its merits. The driving engine for the exponential growth of digital information processing systems is scaling down the transistor dimensions. Although there are some improvements still at the end of the International Technology Roadmap for Semiconductors (ITRS) [1] roadmap industry needs more improved and fast switching devices compare to the current trend. With the introduction to new techniques in fabrication nano materials with carbon group are emerging in Field-effect transistors [2],[3]. They have very good electrical properties along with capabilities to be integrated. The most studied are carbon nanotube FETs (CNFETs) and graphene nanoribbon FETs (GNRFETs). Compared to cylindrical CNTs, GNRs can be grown through a silicon-compatible, transfer-free, and in situ process [4],[5]. Graphene [6] and carbon nanotube (CNT) [7] are the two carbon allotropes, which have become prominent contenders to substitute silicon in post-CMOS technology. With fast switching capability as 2D electron gas in graphene gives high velocity and concentration in carrier which makes large mobility and hence fast switching compared to traditional Si-based CMOS technology [8]. It makes the transistor current quite high as electrons are injected from the source and transit to the drain terminal [7],[8]. In this paper, an ALU has been designed with the emerging Graphene Nano Ribbon-based FET (GNRFET) at 32nm technology node. The results have been optimized by varying the number of graphene nanoribbons (GNRs). Lastly, the results are compared with a CMOS based ALU to get a realistic idea of its performance, as compared to the state of art technology. Delay, Power & Power-Delay Product (PDP) and energy-delay product (EDP) has been chosen as parameters of comparison between the aforementioned devices. The MOSFET type GNRFET model has been used for simulation purposes on the HSPICE platform.

II. GNRFET MODELING

This paper explores circuit-level simulations by using the SPICE models of MOS-GNRFETs from [10],[11]. We simulated the ideal cases of MOS-GNRFETs and Si-based CMOS technology. In MOS-GNRFET circuits, a contact resistance of 20 kΩ is added to all graphene-metal vias. The 32nm Si-CMOS libraries from Predictive Technology Models (PTM) are adopted as comparisons, and the GNRFETs are set to have matching dimensions. The impacts of supply voltage and process variation are evaluated on 1-bit arithmetic and logical unit and the performance of the design is evaluated. Table I defines the default values and their definitions.

TABLE I. DEVICE PARAMETER DEFINITIONS AND USED VALUES

Device Parameter	Description	Used values
L	Physical channel length	32nm
T_{OX}	The thickness of top gate dielectric material	0.95nm
nRib	The number if GNRs in the device.	6
N	The number of dimmer lines in the GNR lattice	12
dop	Source and Drain reservoirs doping fractions	0.001
T_{OX2}	Oxide thickness between channel and substrate/bottom gate	20nm
2*sp	The spacing between the edges of two adjacent GNRs within the same device.	2.0nm

III. ARITHMETIC LOGIC UNIT DESIGN

ALU is the heart of any processor. It performs basic arithmetic and logical operations. The ALU that have been implemented in this paper performs arithmetic operations such as addition & subtraction and logical operations such as AND & OR. In the 1-bit ALU shown in Fig. 1, the logical operations are performed by the 2 input AND & OR gate while the arithmetic operations are performed by the full adder-subtractor unit. The output 'S/D' of the full adder-subtractor unit indicates

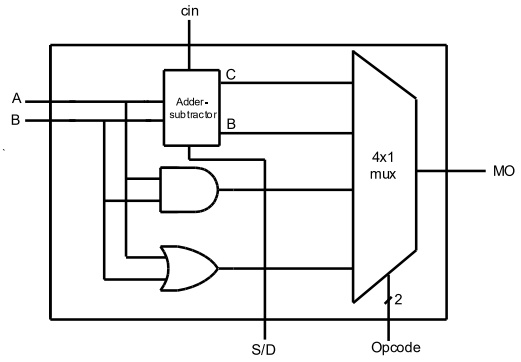


Fig. 1. *Block Diagram of 1-bit ALU.*

'Sum/Difference' bit as they are both the same in the case of addition and subtraction operations. The output 'C' of the full adder-subtractor unit indicates 'Carry' bit in case of addition operation and output 'B' indicates 'Borrow' bit in case of subtraction operation. The 2-bit opcode of 4x1 mux selects the result of either logical or arithmetic block which will be the output 'MO' of the ALU. The following Table II summarizes the select lines and operations performed by the ALU.

TABLE II. ALU OPERATION

Opcode		Operations
S_0	S_1	
0	0	ADD
0	1	SUB
1	0	AND
1	1	OR

IV. RESULT AND DISCUSSION

Timing diagrams of Fig. 2 and Fig. 3 represents the arithmetic operations of GNRFET and MOSFET based ALU. In arithmetic unit operation A, B & Cin are the three input pins and S_0 and S_1 are the opcodes. The unit operates addition when the opcode is in 00 combinations and works subtraction when it is in 01 combinations.

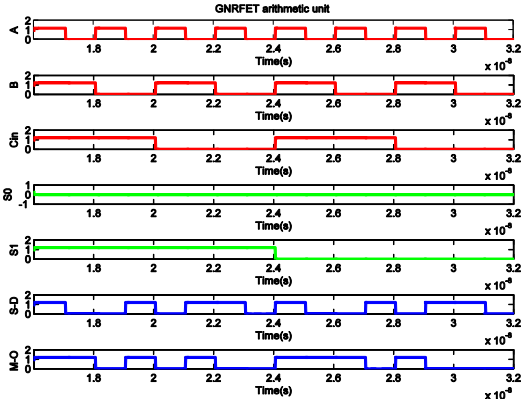


Fig. 2. *Arithmetic unit timing diagram for MOSFET*

The 1-bit ALU based on the University of Illinois (UIUC) MOS-GNRFET model is implemented and analyzed at 32nm technology. Another 1-bit ALU based on 32nm Berkeley Short-channel IGFET Model (BSIM4) for bulk CMOS is also implemented and analyzed for comparison, both having the same design. The ALU circuits are designed using static logic and simulated in HSPICE with a supply voltage of 1.2V. In logical unit operation, both opcodes are varied to get different logical operations. ALU performs AND operation where the output is high only when both inputs are high. In addition to different combinations, the unit performs OR operation with high output with one or more of its input is high. The resultant output performs different performance analysis with two sets of unlike structure. The results are compared at the nominal voltage of 1.2V and the number of nanoribbons is kept equal to 6 for making dimensions of GNRFET comparable to that of Si-CMOS. In this case, the delay of GNRFET based ALU is found to be 1.003ns whereas the delay of the Si-CMOS based is 1.031ns. Propagation delay has been taken between the 50% transition points of the input and output and hence calculated. Power consumption is another figure of merit for this ALU.

TABLE III. COMPARISON BETWEEN GNRFET AND MOSFET

Figure of Merit	GNRFET (32nm)	MOSFET (32nm)
Propagation Delay (PD)(s)	1.003×10^{-09}	1.031×10^{-09}
Average Power consumption (W)	6.498×10^{-06}	1.024×10^{-05}
Power delay Product (PDP)(J)	6.517×10^{-15}	1.056×10^{-14}
Energy delay Product (EDP)(Js)	6.537×10^{-24}	1.088×10^{-23}

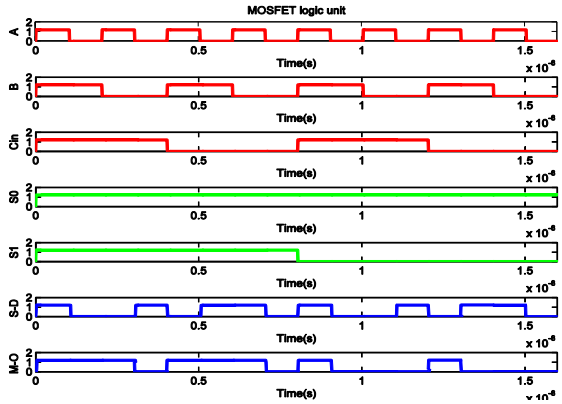


Fig. 3. *Logical unit timing diagram for GNRFET.*

The results show that GNRFET based 1-bit ALU has 2.79% reduction in Propagation Delay (PD), 57.59% reduction in Average Power Consumption, 61.98% reduction in Power Delay Product (PDP), 66.44% reduction in Energy Delay Product (EDP) as compared to Conventional CMOS based 1-bit ALU. As the simulation results have illustrated a substantial improvement in the performance of GNRFET based design over CMOS based design.

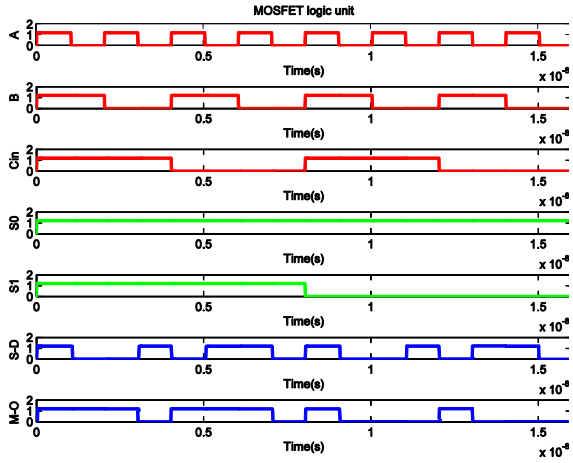


Fig. 4. *Logical unit timing diagram for MOSFET.*

so GNRFETs can be considered as potential candidates to replace conventional CMOS in the future.

TABLE IV. COMPARISON BETWEEN GNRFET(32NM) AND GNRFET(10NM) 1 BIT ALU.

Figure of Merit	**GNRFET (32nm) 1 bit**	**GNRFET (10nm) 1 bit**
Propagation Delay (PD)(s)	1.003×10^{-09}	1.002×10^{-09}
Average Power consumption (W)	6.498×10^{-06}	2.118×10^{-06}
Power delay Product (PDP)(J)	6.517×10^{-15}	2.122×10^{-15}
Energy delay Product (EDP)(Js)	6.537×10^{-24}	2.126×10^{-24}

Simulation results tabulated in Table IV shows that GNRFET (10nm) based 1-bit ALU has 0.09% reduction in Propagation Delay (PD), 67.41% reduction in Average Power Consumption, 67.44% reduction in Power Delay Product (PDP), 67.48% reduction in Energy Delay Product (EDP) as compared to GNRFET (32nm) based 1-bit ALU. From Table V, in case of GNRFET (32nm) based 2-bit ALU has 1.22% reduction in Propagation Delay (PD), 28.14% increase in Average Power Consumption, 27.28% increase in Power Delay Product (PDP), 26.38% increase in Energy Delay Product (EDP) as compared to Conventional CMOS based 2-bit ALU.

TABLE V. COMPARISON BETWEEN GNRFET(32NM) 2 BIT AND MOSFET(32NM) 2 BIT ALU.

Figure of Merit	**GNRFET (32nm) 2 bit**	**MOSFET (32nm) 2 bit**
Propagation Delay (PD)(s)	5.507×10^{-09}	5.574×10^{-09}
Average Power consumption (W)	2.874×10^{-05}	2.065×10^{-05}
Power delay Product (PDP)(J)	1.583×10^{-13}	1.151×10^{-13}
Energy delay Product (EDP)(Js)	8.716×10^{-22}	6.416×10^{-22}

V. CONCLUSIONS

The performance analysis of a single-bit and multi-bit ALU is done with two different transistor structures and technology node in this paper. In case of 1-bit ALU based on 32nm technology, the paper reports a small percentage of reduction in PD along with more than 60% of average power consumption with a significant amount of PDP reduction on GNR based device over the conventional one. Reduction of feature size (to 10nm for GNRFET) leads to small reduction in PD and significant amount of reduction in EDP, PDP and average power as compared to 32nm based GNRFET. But in case of 2-bit 32nm ALU GNRFET based design shows a small reduction only in PD as compared to conventional 32nm based CMOS design. So, in the case of the 1-bit combinational unit, the performance parameter of graphene-based architecture has the upper hand over the Conventional CMOS based unit.

VI. REFERENCES

[1] L. Wilson, International technology roadmap for semiconductors (ITRS),Semiconductor Industry Association, 2013.

[2] J. Zhang et al., "Robust Digital VLSI using Carbon Nanotubes," Keynote paper, IEEE Trans. CAD, 2012.

[3] S. Chilstedt, C. Dong, and D. Chen, "Carbon nanomaterials transistors and circuits, transistors: Types, materials and applications," Nova Science Publishers, 2010.

[4] W. de Heer et al., "Pionics: the emerging science and technology of graphene-based nanoelectronics," IEDM, 2007.

[5] M. P. Levendorf et al., "Transfer-Free Batch Fabrication of Single Layer Graphene Transistors," Nano Letters, 2009.

[6] K. S. Novoselov, V. Fal, L. Colombo, P. Gellert, M. Schwab, and K. Kim, A roadmap for graphene,*Nature*, 490(7419) , 192-200, 2012.

[7] S. Iijima, Helical microtubules of graphitic carbon,*Nature*, 354(6348), 56-58, 1991.

[8] F. Schwierz, Graphene transistors: status, prospects, and problems,*Proc. of the IEEE*, vol-101(7), pg-1567-1584, 2013,

[9] K.C. Yung, W.M. Wu, M.P. Pierpoint & F.V. Kusmartsev." Introduction to graphene electronics – a new era of digital transistors and devices", *Contemporary Physics*, vol-54:5, pg-233-251,

[10] M. Gholipour, Y.Y. Chen, A. Sangai, and D. Chen, *" Highly accurate SPICE-compatible modeling for single-and double-gate GNRFETs with studies on technology scaling",Proc. of IEEE on Design, Automation & Test in Europe. European Design and Automation Association*, vol-120, 2014,

[11] Ying-Yu Chen, Morteza Gholipour, Deming Chen, "Flexible transition metal dichalcogenide field-effect transistors: A circuit-level simulation study of delay and power under bending process variation and scaling", *Design Automation Conference (ASP-DAC) 2016 21st Asia and South Pacific,* pp. 761-768, 2016.

Numerical Assessment of the Backward Facing Step for NACA 0015 Airfoil using Computational Fluid Dynamics

Asif Kabir
Dept. of Mechanical Engineering
Bangladesh University of Engineering and Technology (BUET)
Dhaka, Bangladesh
asifkabir121@gmail.com

Md. Saadbin Chowdhury
Dept. of Mechanical Engineering
Bangladesh University of Engineering and Technology (BUET)
Dhaka, Bangladesh
saadshovon@gmail.com

Md. Jahirul Islam
Dept. of Mechanical Engineering
Bangladesh University of Engineering and Technology (BUET)
Dhaka, Bangladesh
j.i.munna319@gmail.com

Mehran Islam
Dept. of Offshore Engineering
Bangabandhu Sheikh Mujibur Rahman Maritime University (BSMRMU)
Dhaka, Bangladesh
mehran.islam.92@gmail.com

***Abstract*—The objective of this work was to investigate the lift and drag characteristics of a NACA 0015 airfoil with backward facing step. For a free stream velocity of 30m/s, the point of separation was determined via CFD analysis for increasing angle of attack. A backward facing step was then introduced just prior to the point of separation which extended till the trailing edge. Computational studies were conducted with the goal of enhancing the aerodynamic performance of the airfoil by inserting several backward facing steps of gradually increasing depth. The CFD analysis was conducted for 30 m/s wind speed, with steps placed just prior to the point of separation for 18° angle of attack. The results produced show that as much as 26% reduction in drag is possible by introducing the step. But at the same time, about 12% reduction in lift was also observed. A wind tunnel test on a NACA 0015 was also conducted to verify the results from CFD analysis. Predominantly the goal of this thesis was to numerically determine whether or not the aerodynamic performance of an airfoil can be enhanced by introducing a backward facing step on the upper surface of the airfoil.**

Keywords—backward facing step, computational fluid dynamics, life force, drag force, separation delay

I. INTRODUCTION

An Airfoil is a structure, piece, or body designed to obtain a useful reaction upon itself in its motion through the air. An airfoil may be no more than a flat plate, but usually it has a cross section carefully contoured in accordance with its intended application or function. Airfoils are employed on aircraft as wings to produce lift or as propeller blades to produce thrust. Both these forces are produced perpendicular to the air flow. Drag is a consequence of the production of lift/thrust and acts parallel to the airflow. Airfoil shapes are designed to provide high lift values at low drag for given flight conditions. The airfoil creates a situation where the velocity above the airfoil is higher than the velocity below the airfoil. According to Bernoulli's principle, a liquid or gas, areas with high relative velocity create lower pressure systems, and areas with low relative velocity create high pressure systems. So, the upper surface pressure is lower than the lower surface pressure. This pressure difference creates a resultant force. This resultant force has two components. The vertical component is known as lift and the horizontal component is known as drag. Fig 1.1 shows the forces acting on an airfoil.

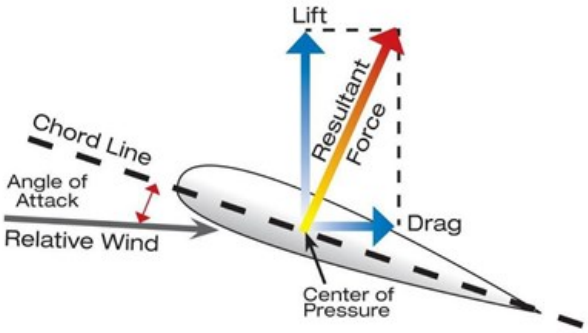


Fig. 1.1: Force vectors on an airfoil

The drag FD and lift FL are represented in terms of two dimensionless coefficients: the drag coefficient and Lift coefficient, defined as

$$C_D = \frac{F_D}{\frac{1}{2}\rho_\infty {U_\infty}^2 A} \quad (1)$$

$$C_L = \frac{F_L}{\frac{1}{2}\rho_\infty {U_\infty}^2 A} \quad (2)$$

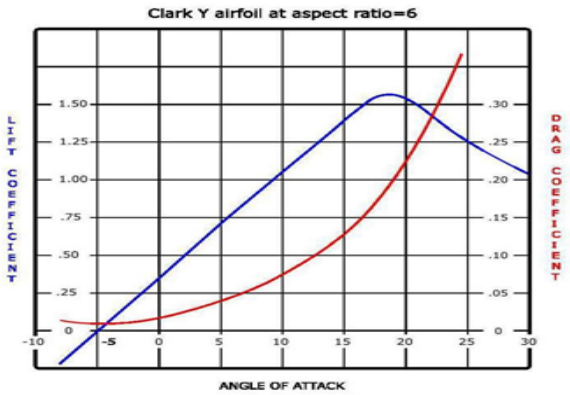


Fig. 1.2: Lift and drag curves for typical airfoil.

The lift generated by an airfoil depends on such factors as: Speed of the airflow, Density of the air, Total area of the

airfoil, Angle of attack. Fig 1.2 shows a lift and drag curve obtained in wind tunnel testing for different attack of angle.

The curve represents an airfoil with a positive camber so some lift is produced at zero angle of attack. With increased angle of attack, lift increases in a roughly linear relation. At about 18 degrees this airfoil stalls, and lift falls off quickly beyond that. The drop in lift can be explained by the action of the upper-surface boundary layer, which separates and greatly thickens over the upper surface at and past the stall angle. The thickened boundary layer's displacement thickness changes the airfoil's effective shape, in particular it reduces its effective camber, which modifies the overall flow field so as to reduce the circulation and the lift. The thicker boundary layer also causes a large increase in pressure drag, so that the overall drag increases sharply near and past the stall point. So, flow separation occurs over the airfoil top surface when the angle of attack (AOA) exceeds a certain value- the "stalling" angle of attack.

In other words, flow separation occurs when the boundary layer travels far enough against an adverse pressure gradient that the speed of the boundary layer relative to the object falls almost to zero. The fluid flow becomes detached from the surface of the object, and instead takes the forms of eddies and vortices. In aerodynamics, flow separation can often result in increased drag, particularly pressure drag which is caused by the pressure differential between the front and rear surfaces of the object as it travels through the fluid.

II. Literature Review

Boundary layer separation is a phenomenon that designers normally attempt to avoid, as it leads to a decrease in lift and an increase in drag. This separation may occur anywhere an adverse pressure gradient exists, such as on wings, airfoils, turbo machinery, and engine inlets. This problem is especially prevalent at low Reynolds numbers, when laminar flow is prevalent.

One of the earliest investigations of the flow features near surfaces structured with spherical dimples was made by Bearman and Harvey [1]. This study demonstrated the effectiveness of dimples as a technique of causing early boundary layer transition without the drag penalties associated with sand roughness. Unlike a smooth or roughed round ball the dimpled ball drag curve remains almost constant at the post critical Reynolds number. This confirms the dimples have a more beneficial effect on the drag reduction and laminar-turbulent transition than sand roughness.

Experiments carried out with the dimpled circular cylinder [2] have the shallow dimples cause the critical drag reduction event to occur at a lower Reynolds number than that for a smooth cylinder. It was concluded that the dimples affect the flow around a circular cylinder in a similar manner as the flow around a golf ball. Experimental studies of a cross-flow dimpled cylinder [3], [4] have revealed that the separation point is moved in the downstream direction. This results in a smaller wake region and a reduction in the pressure drag.

A two-dimensional groove is the limiting case of a closely spaced span wise array of dimples. The background relevant to the rectangular-shaped groove can be found in many existing publications. The data obtained by Zhak [5] have shown the laminar-turbulent flow transition inside rectangular cavity can occur via the specific mechanism of the Taylor-Görtler instability The available publications demonstrate that steps engraved into a solid surface generate 3D or 2D vortex flow structures that energize the boundary layer and reduce pressure losses over a wide range of Reynolds numbers. Besides Mahdi [6] and Asif [7] have all also used a wide range of Reynolds numbers for NACA 0012. As air flows over an indented surface treatment, such as a dimple or groove, a complex flow pattern develops. The flow inside the step is highly time dependent. This unsteadiness causes instability with a laminar boundary layer, causing the boundary layer to begin transition to turbulent flow. The turbulent flow promotes the mixing of high momentum flow in the boundary layer, which helps to combat the adverse pressure gradient.

Lake [8], Rouser [9], and Khalatov [10] have all previously used these techniques in the design of improved passive flow separation control techniques for cases of laminar separation. The indentations do not affect separation of a turbulent boundary layer, as the effects of mixing have already been realized through the turbulent nature of the flow. The effectiveness of the recessed technique depends on several factors, including the location of the indentation relative to the separation bubble, the geometry of the step, and the pre-recess boundary layer thickness. Although several studies have assessed the effectiveness of indented surface treatments for delaying boundary layer separation, definitive design guidelines have never been published. This study examined the previous literature with an eye towards determining the optimum placement, depth, and length of the steps. The best location for the grooves or dimples is just prior to the region of separation. So, our objective was to introduce a backward facing step just prior to the point of separation on the upper surface of a NACA 0015 airfoil and investigate its aerodynamic performance using Computational Fluid Dynamics (CFD) and verifying the CFD results by experimental approach.

III. Methodology

A. Formulation of the Flow Problem

The flow problem was set up as an airfoil fixed at a certain angle of attack with air flowing over it at a given free stream velocity. The objective was to study the aerodynamic performance of the airfoil using backward facing step that extends till the trailing edge of the airfoil. The configuration of a step is defined by its location on the airfoil chord and its depth expressed in the terms of airfoil maximum thickness. The employment of step to control the flow field around the airfoil comprises the passive flow control used in this study.

B. Parameter Range

There are several parameters which can affect the flow field around an airfoil. Both the airfoil and step profiles constitute the geometric parameters. The most important ones are the shape of airfoil (symmetric or cambered); chord length, which was taken as 19 cm; angle of attack; chord-wise location of the step; its length and depth on the upper surface. Fig. 3.1 shows a sketch of a sample airfoil configuration with backward facing step defining the main geometric parameters.

The flow parameters are the free stream velocity, density and viscosity grouped together as the Reynolds number (Re). For this flow problem the free stream velocity was taken as 30 m/s, density as 1.166 kg/m^3 and the viscosity as 1.8599×10^{-5} kg/s-m, which gave a corresponding Reynolds number of 357,341 while α was 18 degrees.

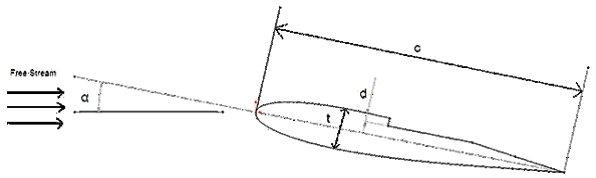


Fig. 3.1: Geometric parameters associated with airfoil with backward facing step

C. Computational Details

All of the computational work was done using commercial software packages. The airfoil profile was generated using SOLIDWORKS 2016. The profile was then imported into ANSYS 17.0 FLUENT. All the Pre-processing comprising generating meshes for all the airfoil configurations, quality control, setting up of the boundary conditions was done using FLUENT. FLUENT is a finite volume method based CFD tool which solves the governing equations of conservation of mass and momentum. Post processing including generation of plots and figures was done using FLUENT.

The problem was set up as a two dimensional flow problem in single phase. The Cartesian co-ordinate axes were fixed at the leading edge of the airfoil in each case. The computational domain comprises of four blocks for all base airfoil configurations, while the backward facing step configurations had six mesh blocks. Fig. 3.2 and 3.3 shows the grid distribution around a NACA 0015 airfoil. The flow was modeled as steady, incompressible in the flow domain comprising several blocks. The fluid chosen was air at 30° C and 1 atm pressure which were set as the reference values along with the free-stream velocity set corresponding to the flow Re for each simulation run for the preliminary cases of study. The inlet was set as "Velocity Inlet" which includes all the front, top and bottom edges of the flow domain. The x and y-velocities are "u", which is defined by the Reynolds number used and "v". "Spalart-Allmaras" one equation turbulence model was used in FLUENT for all computations. This computational model is well suited for modeling wall driven flows which transpire most commonly over aerodynamic bodies. All the exit boundaries combined were set as "Pressure Outlet". All the airfoil edges together were set as "Wall" with no-slip condition. The flow solver used the SIMPLE algorithm for pressure-velocity coupling. Least Squares Cell Based method was used for gradient computations. Pressure was set as second order, Momentum, and Modified Turbulent Viscosity were set as Second Order Upwind for Spatial Discretization computations.

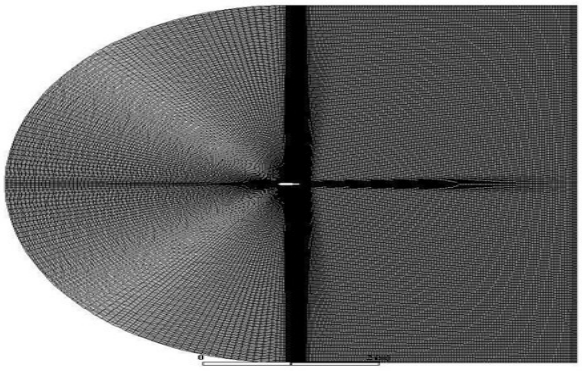


Fig. 3.2: Grid distribution around a NACA 0015 airfoil

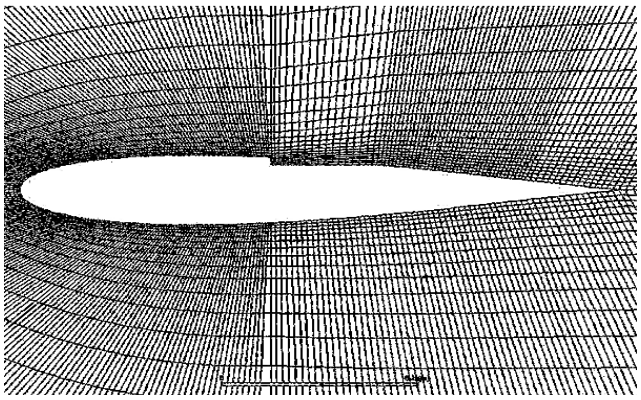


Fig. 3.2: Grid distribution around a NACA 0015 airfoil with backward facing step (close).

IV. Wind Tunnel Testing

A subsonic wind tunnel test on a NACA 0015 airfoil was conducted to validate some of the results obtained with numerical studies conducted as a part of this research effort. The experimental studies were conducted in the Fluid Mechanics and Machineries Laboratory located in the Department of Mechanical Engineering at Bangladesh University of Engineering and Technology. This section presents the approach followed for the wind tunnel testing of the NACA 0015 airfoil.

The airfoil was contained in a clear sided duct in 'closed ends' arrangement; thus airflow was only across the curved surfaces of the airfoil. Pressure taps were placed along the top and bottom surfaces of the airfoil. Each of pressure taps were numbered according to their placement. Pressure taps on the upper surface of the airfoil were numbered 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 and 16 while taps on the lower surface of the airfoil were numbered 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32.

First the initial readings were taken from each tube of the manometer. Second readings were taken after the flow over the airfoil reached steady condition which was the final value in centimeter. The differences were converted into meter. Comparing the experimental data for a NACA 0015 airfoil at 18° AOA with the CFD result, we can see that our simulation result matches very closely to the experimental results. The slight deviation close to the leading edge may result from friction during the experimental procedure.

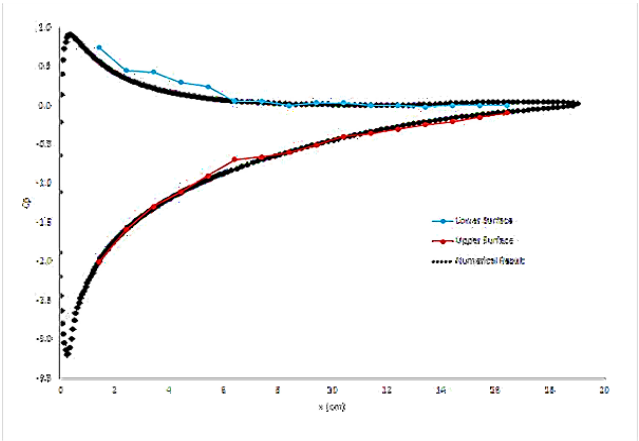


Fig. 4.1: Experimental results for 18° AOA, 30 m/s free stream velocity.

V. Results and Discussion

CFD analysis was conducted on a NACA 0015 airfoil with a chord length of 19 cm and a free stream velocity of 30 m/s. The temperature taken was 30° C which gave us air density, ρ = 1.166 kg/m3 and kinematic viscosity, μ = 1.8599 × 10-5 kg/s-m. This gave us a Reynold's number of 357,341 based on the chord length. Based on this Reynold's number, the turbulent Spalart-Allmaras model was chosen for the CFD simulation.

At first, for a base NACA 0015 airfoil, the point of separation was determined for increasing angle of attack from 8° up to 20° and found that for the above conditions, the stall angle was at 18°.The determination of the separation point was necessary because from the literature review, it was found that the best location for the abrupt change in geometry is just prior to the region of separation. So, a backward facing step was introduced just prior to the point of separation on the upper surface of the NACA 0015 airfoil and the lift and drag characteristics were investigated by CFD analysis.

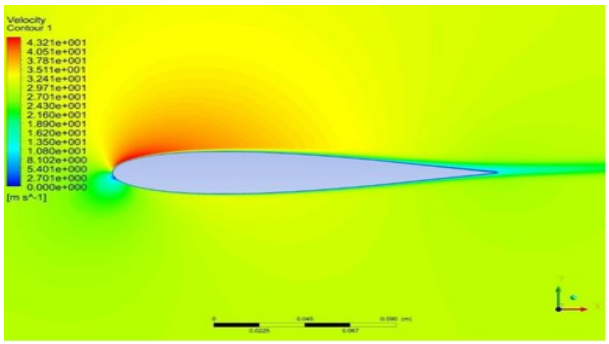


Fig. 5.1: Velocity contour of NACA 0015 at 8° AOA, 30 m/s free stream velocity

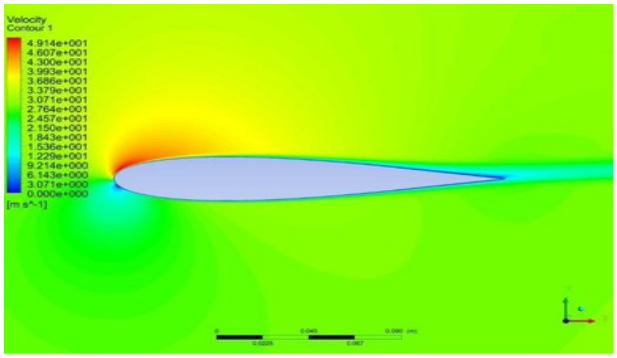


Fig. 5.2: Velocity contour of NACA 0015 at 12° AOA, 30 m/s free stream velocity

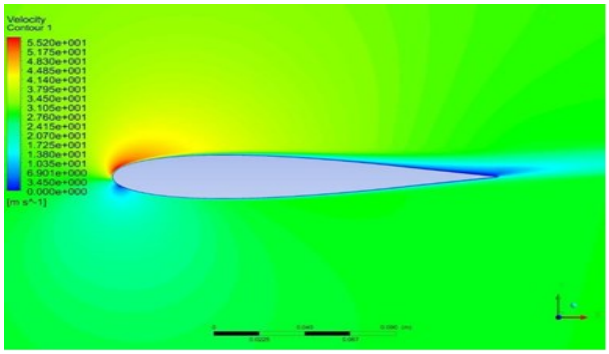


Fig. 5.3: Velocity contour of NACA 0015 at 16° AOA, 30 m/s free stream velocity

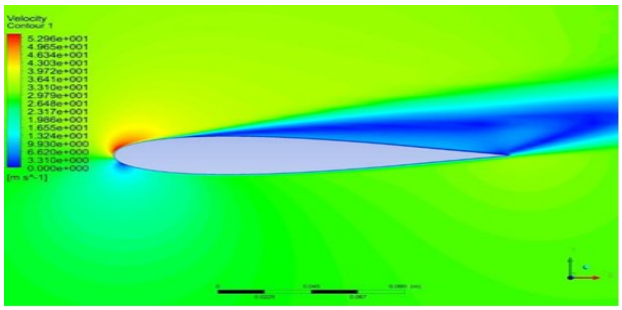


Fig. 5.4: Velocity contour of NACA 0015 at 20° AOA, 30 m/s free stream velocity

From 8° to 10° AOA, no separation can be observed. However, a separation can be observed near the trailing edge at 12° AOA. From here on, the separated region grows and slowly moves towards the leading edge as the angle of attack increases.

Near the trailing edge, a small separated region can be noticed at 14° AOA. The separation point can be located where the x-wall shear becomes negative. In this case, from the CFD analysis it was found to be at x/c = 0.83 from the leading edge.

Stall conditions were reached at 18° AOA. After which the flow over the upper surface of the airfoil becomes more separated. A very large increase in the separated region can be noticed when the angle of attack was increased from 18° to 20° due to exceeding the stall angle.

In order to place the backward step, the angle of attack was chosen to be 18° and the free stream velocity to be 30 m/s. For 18° AOA, the separation point was found at x/c = 0.47 from the leading edge. The step was placed at point prior to the separation point; at x/c = 0.42 from the leading edge.

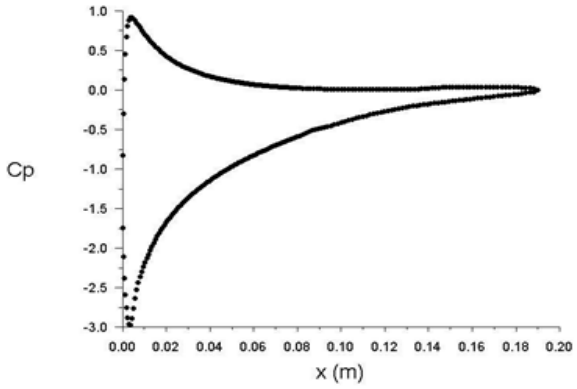


Fig. 5.5: Pressure distribution over NACA 0015 with step depth of 1 mm.

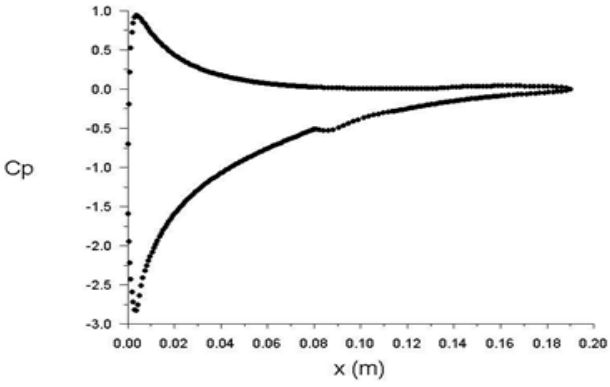


Fig. 5.6: Pressure distribution over NACA 0015 with step depth of 3 mm.

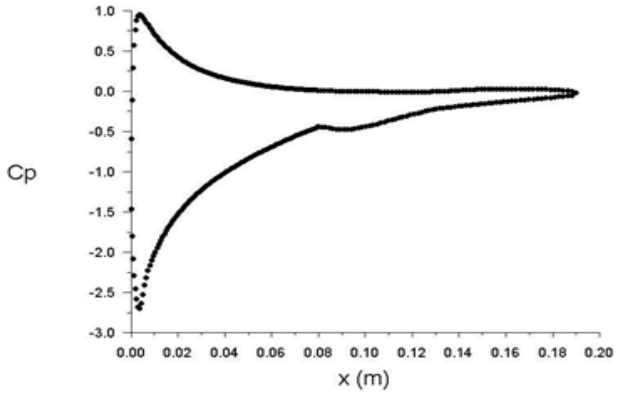


Fig. 5.7: Pressure distribution over NACA 0015 with step depth of 5 mm.

The pressure distribution curve follows similar trend to the base airfoil. But a small pressure drop can be noticed when the flow immediate to upper surface of the airfoil passes the step. This disturbance in the pressure distribution becomes larger with increasing step depth. A 4% reduction in lift and a 9% reduction in drag can also be observed due to the addition of the step on the upper surface.

For a step depth of 3 mm, a sharper disturbance can be observed in the pressure distribution curve. The pressure starts to drop sharply from location of the step at x/c = 0.42 until x/c = 0.44, and then the pressure sharply increases until x/c = 0.52; after which the curve follows similar trend to the base airfoil.

For a backward facing step of 5 mm depth, a higher pressure drop can be observed that starts from x/c = 0.42 and stops at x/c = 0.48, followed by a noticeable pressure increase until x/c = 0.68. From there the curve starts to follow similar trend to the base airfoil. The highest lift reduction at 12% and the highest drag reduction at 26% can be observed. With increasing step depth, the step length that extends till the trailing edge also increases.

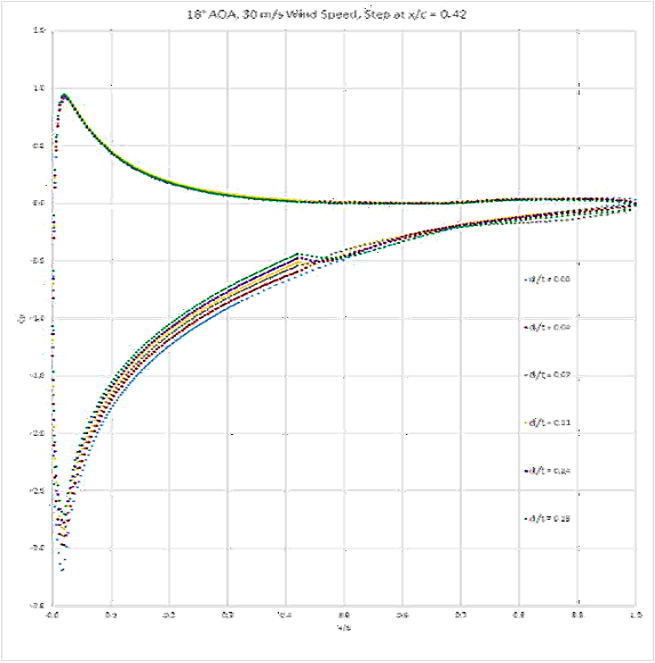


Fig. 5.8: Comparison of Cp for increasing step depths.

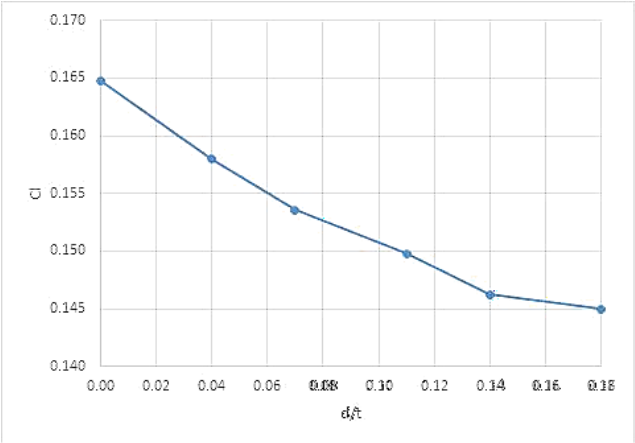


Fig. 5.9: Reduction in lift coefficient with increasing step depth.

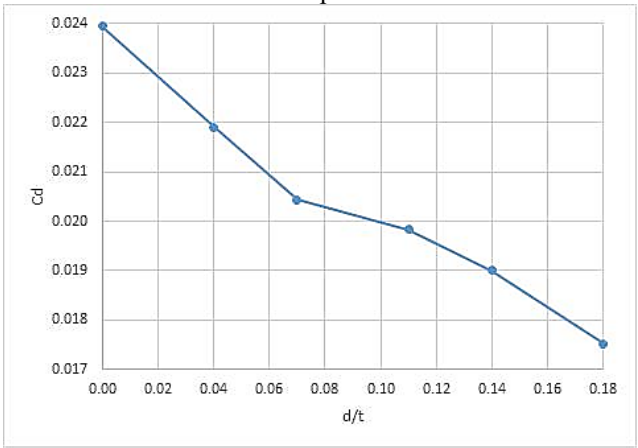


Fig. 5.10: Reduction in lift coefficient with increasing step depth.

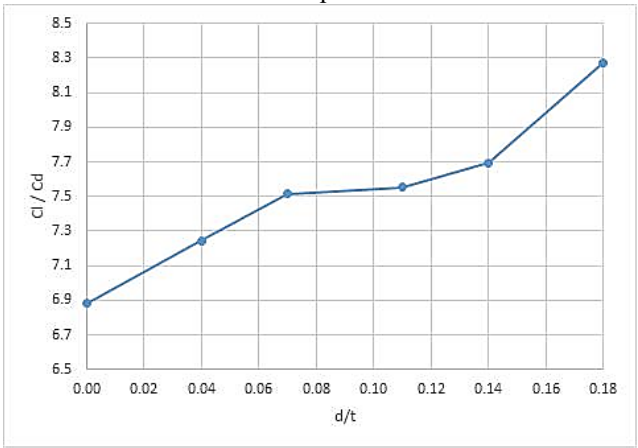


Fig. 5.11: Increase in lift to drag ratio with increasing step depth.

With the increase of the step depth, both the lift and the drag decreases. From fig. 5.9 and 5.10, it is evident that both lift and drag reduction is linear until d/t = 0.07 and becomes irregular as d/t increases. However, it is clear that the reduction in lift is much smaller than the reduction in drag. As a result, even though the lift is decreasing, for a larger decrease in drag, we get higher lift to drag ratio with increasing step depth. Similar to the trend that can be observed in fig. 5.9 and 5.10, it can be seen that the increase in lift to drag ratio is also linear up until d/t = 0.07 and becomes irregular as the d/t increases.

VI. Conclusion

The objective of this study was to investigate the aerodynamic performance of a NACA 0015 airfoil with backward facing step for a chosen velocity and angle of attack. The airfoil with a backward facing step on the upper surface was shown to have the potential to enhance the aerodynamic characteristics by reducing the drag coefficients and increasing the lift-to-drag ratios considerably. The findings of this study may be summarized as following:

- It is possible to enhance the aerodynamic performance of an airfoil by introducing a backward facing step extending till the trailing edge placed on the upper surface prior to the separation point.
- Drag coefficient can be reduced by the introduction of the step.
- Drag coefficient decreases with the increase of step depth.
- A drag reduction of up to 26% was observed in this study.
- However, lift coefficient also decreases with the increase of step depth. A lift reduction of up to 12% was observed.

Although the reduction in drag is much higher than the reduction in lift. As a result, lift to drag ratio actually increases with the increase in step depth.

VII. Recommendation

In this work, we have studied the aerodynamic performance of a NACA 0015 airfoil with a backward facing step extending till the trailing edge. The chord length was taken to be 19 cm, free stream velocity was chosen to be 30 m/s. The step was placed at x/c = 0.42 at 18° AOA. The depth of the step was varied from d/t = 0 to d/t = 0.18.

- The aerodynamic performance of an airfoil with a backward facing step can be investigated for different AOA.
- The backward facing step studied in this work was perpendicular to the chord. Steps of different angle can be studied to enhance the aerodynamic performance of the airfoil.
- The effect of different geometry for surface modification, for example, a circular step extending till the trailing edge, can be investigated.
- The aerodynamic characteristics of an asymmetrical airfoil with backward facing step should be investigated.

Further increase in step depth should be investigated. A backward facing step can be introduced at the lower surface of the airfoil.

References

[1] P. W. Bearman and J. K. Harvey, "The Golf Ball Aerodynamics," *Aeronautical Quarterly,* vol. 27, pp. 112-121, 1976.

[2] P. W. Bearman and J. K. Harvey, "Control of Circular Cylinder Flow by the use of Dimples," *AIAA Journal,* vol. 31, pp. 1753-1756, 1993.

[3] A. P. Kozlov, A. V. Shchukin and R. S. Agachev, "The Hydrodynamics Effects from Spehrical Depressions on the Cross Flow Cylinder," *Izevtia Vuzov, ser. Aviatsionnaya Technika,* pp. 27-34, 1992.

[4] A. S. Kovalenko and A. A. Khaloatov, "Fluid Flow and Heat Transfer Features at a Cross-Flow of a Dimpled Tube in Confined Space," in *ASME*, 2003.

[5] V. D. Zhak, "The Taylor- Görtler Vortices and Three-Dimensional Flow Evolution in Cavity," *Russian Journal of Engineering Thermophysics,* vol. 5, pp. 165-176, 1995.

[8] J. P. Lake, "Flow Separation Prevention on a Turbine Blade in Cascade at Low Reynolds Number," Wright-Patterson AFB, June 1999.

[9] K. Rouser, "Use of Dimples to Suppress Boundary Layer Separation on a Low Pressure Turbine Blade," Wright-Patterson AFB, December 2002.

[10] A. A. Khalotov, I. S. Varganov, G. V. Kovalenko and I. I. Borisov, "Blade of Turbine, Compressor, or Fan with a Flow Separation Control". Ukraine Patent 44532A, 15 February 2002.

A Spatiotemporal Analysis and Forecasting of Electricity Generation-Mix in Bangladesh

Sujoy Barua
Dept. of Electrical and Electronic Engineering
Chittagong Institute of Engineering and Technology
Chittagong, Bangladesh
sujoy.baruaa@gmail.com

Anik Nath
Dept. of Electrical and Electronic Engineering
Chittagong University of Engineering and Technology
Raozan, Chittagong, Bangladesh
anik.nathnu@gmail.com

Fahim Shahriyar
Dept. of Electrical and Electronic Engineering
University of Science and Technology Chittagong
Foy's Lake, Chittagong, Bangladesh
fahimshahriyar.fsr@gmail.com

Nur Mohammad
Dept. of Electrical and Electronic Engineering
Chittagong University of Engineering and Technology
Raozan, Chittagong, Bangladesh
nur.mohammad@cuet.ac.bd

***Abstract*—An analysis is presented to show time series data of electricity generation mix and forecasting by 2030 in Bangladesh. The comparative studies have been analyzed using spatiotemporal data of Germany, Australia and Bangladesh. The spatiotemporal data has been taken out from World Bank data bank for analysis. A Linear regression technique is applied for forecasting electricity generation mix from 2015 to 2030. The result shows the rise of renewable energy sources, coal and oil, and the diminution of natural gas gradually.**

Keywords— spatiotemporal, forecasting, electricity, energy mix, solar energy

I. INTRODUCTION

Power Generation mix refers to the combination of the various primary energy sources (fossil fuels, nuclear energy, renewable energy sources) used to meet electricity according to geographical region. For each region or country, the composition of the energy mix depends on the availability of usable resources domestically or the possibility of importing it. The Policy defines by historical, economic, social, demographic, environmental and geopolitical factors [1]. These differences can be appreciated by taking a look at the electricity generation and consumption information for a country.

World Energy Statistics presents comprehensive world electricity generation mix statistics on all energy sources: coal, gas, oil, electricity, renewable sources and waste. Coal is the highest source (39.3%) of electricity generation mix in the world.

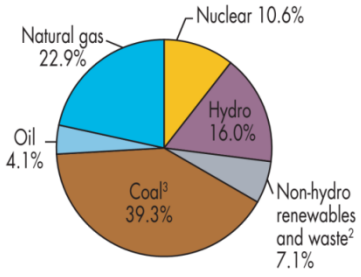


Fig.1. Global Electricity Generation Mix

TABLE 1. ELECTRICITY GENERATION MIX IN BANGLADESH

Gas	8810 MW (64.99%)
Furnace Oil	2785 MW (20.55%)
Diesel	880 MW (6.49%)
Power Import	600 MW (4.43%)
Hydro	230 MW (1.7%)
Coal	250 MW (1.84%)
Total	**13, 555 (100%)**

It is followed by natural gas which is contributed 22.9% in World electricity generation. Hydro (16%), nuclear (10.6%), Renewable energy excluded hydro (7.1%) and oil (4.1%) are respectively in the list of electricity generation mix in the world. Fig.1. shows the percentage of different sources of electricity generation [2].

In Bangladesh, the dominant electricity generation-mix or sources are largely reliant on gas and imported oil. Table 1 describes percentage share of electricity generation mix by sources [3]. The Contribution of gas and oil are significant which accounted for 60.99% and 20.55% respectively of the country's total power generation in 2017, while gas occupying 89.22%, and oil 3% [4]. Electricity generation mix changes over time in Bangladesh. By 2030, coal and nuclear power plant will be operation in which have the ability to generate a huge amount of electricity at a time [5]. So, power generation mix would be changed to renewable energy due to global warming. Even solar energy is potential in Bangladesh as a sub-tropical region [6] and govt. has already planned 2000 MW by 2021 [7].

An agriculture earnings based Bangladesh is one of world's most populated countries. Demand of energy creates with rapid urbanization and industrialization in Bangladesh. So, Bangladesh government should ensure electrical energy for the people that will be affordable and environmentally friendly. Energy sector in Bangladesh is reviewed, and it is hard to accomplish Millennium development goals (MDG) and Sustainable Development Goal (SDG) without improving energy sector [8]. Bangladesh government has already taken step to break the tax for investors in renewable energy [9]. Though Renewable currently play a small role in Bangladesh's fuel mix, despite domestic roof-top solar power units have brought electricity to over a million homes which are not on the grid. There are lots of opportunities in urban building to establish on grid solar power. Ensuring

electricity for all, Bangladesh government is establishing coal and uranium based power plants, as these are able to produce huge electricity energy at a time. Electricity market and cost of energy are hard to determine for power plant project. So, the competitive electricity market is optimized by game theory for revealing existence Nash equilibrium for network incommodity [10]. Last ten years, there is no any wind energy project in Bangladesh. This uncertainty of wind generation and market is managed proposing wind induced demand response (WIDR) using bi-level optimization frameworks [11]. The levelized cost of electricity (LCOE) including import and export is discussed in [12] for the each country. Next ten years, energy sector is expecting new generation-mix in electricity market carrying out the MDG and SDG towards development of Bangladesh.

In this paper, a spatiotemporal data of electricity generation mix of Bangladesh is analyzed and forecasted electricity generation sources for future indication using MS Excel and MatLab. And a comparative study has been done for showing how generation mix is varied with geography location and changed with time being in three different geography locations (Germany, Australia and Bangladesh). Linear regression is used for forecasting to indicate future electricity generation mix in Bangladesh. Linear regression can find linear relationship between two continuous variables. One is predictor and other is response or dependent variable. It looks for statistical relationship between two variables is said to be deterministic if one variable can be accurately expressed by the other [13]. In research, electric power generation mix is analyzed for long-term indication deploying model in Malaysia [14]. World Bank organization maintain data bank for such kind research and analysis [15]. The objective of this paper is to explore electricity generation sources which are an effective generation-mix towards the long-term energy security of Bangladesh considering geographic, demographic, environment, socio-economical factors for the national sustainable development.

TABLE 2: COUNTRY PROFILE OF AUSTRALIA, GERMANY AND BANGLADESH [14]

Country Profile (2016)			
	Australia (AUS)	**Germany (DEU)**	**Bangladesh (BGD)**
Population, total (millions)	24.21	82.35	162.95
Population growth (annual %)	1.5	0.8	1.1
Surface area (sq. km) (thousands)	7,741.2	357.4	147.6
Population density (per sq. km of land area)	3.2	357.4	1,251.8
Urban population growth (annual %)	1.6	1.1	3.3
CO_2 emissions (metric tons per capita)	15.37	8.89	0.46
Electric power consumption (kWh per capita)	10,059	7,035	310
Access to electricity (% of population)	100	100	75.92
Access to electricity, rural (% of rural population)	100	100	68.85
Access to electricity, urban (% of urban population)	100	100	94.01

II. METHODOLOGY

A. Country profile and Study

In this study, two countries are selected according to geographical location and development. As Developed countries their profile and status of electricity generation mix help to find result of developing country Bangladesh. Table 2 highlights information about the three countries which are indicated the present status of Australia, Germany and Bangladesh. Most of Points of table carry the information of electricity status phenomena over the countries that ease to do comparative study among them.

Bangladesh has topped the list of total population having 162.95m (m=Millions), 2nd to Germany (82.35m) and Australia (24.21m) is the last. As for population growth, it is the highest in Australia having 1.5%, Bangladesh is the second having 1.1% and Germany is the 3rd having 0.8%. Regarding population growth, land area determines the density of population of the country. Bangladesh is 1251.8 persons per sq. km of land area whereas Germany (357.4 person) and Australia (3.2 person) is the least number. Population of the country describes access to electricity and electricity 100% in both rural and urban but Bangladesh could not reach CO_2 emission of country. Australia and Germany access that point because of density of population in the country, rural and urban consumers consume 68.85% and 94.01% respectively.

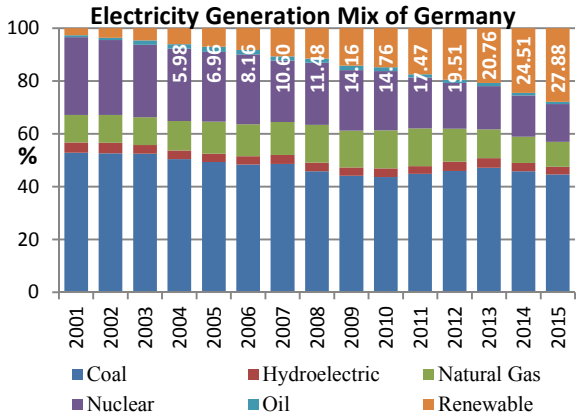


Fig.2. Percentage of Electricity Generation Mix in Germany

Electricity generation mixes of Germany in the period of 2001-2015 are shown in figure 2 where Coal is the largest percentage share of generation sources in the period of 15 years. There are two significant progresses reducing uranium sources and increasing renewable energy sources from electricity generation mix share over last 15 years (2001-2015).

The figure 3 shows percentage of the electricity generation of Australia about previous 15 years from 2011 to 2015. Australia is also highest percent source (80%-60%) using coal to produce electricity over the period. There is trend including natural gas in mix and renewable sources are

placing gradually over coal in last 10 years (2005-2015). There is no nuclear plant over this period significant of electricity generation sources in Australia.

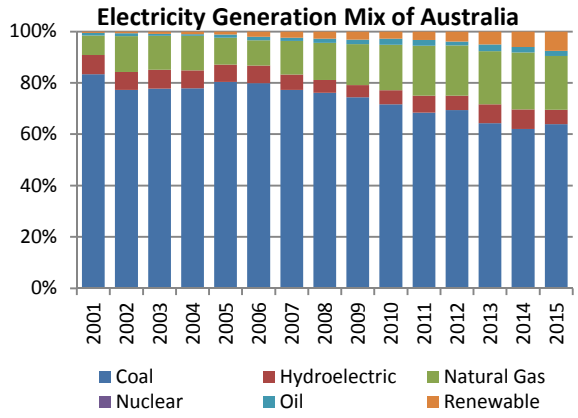


Fig.3. Percentage of Electricity Generation sources in Australia

The figure 3 shows percentage of the electricity generation of Australia about previous 15 years from 2011 to 2015. Australia is also highest percent source (80%-60%) using coal to produce electricity over the period. There is trend including natural gas in mix and renewable sources are placing gradually over coal in last 10 years (2005-2015). There is no nuclear plant over this period significant of electricity generation sources in Australia.

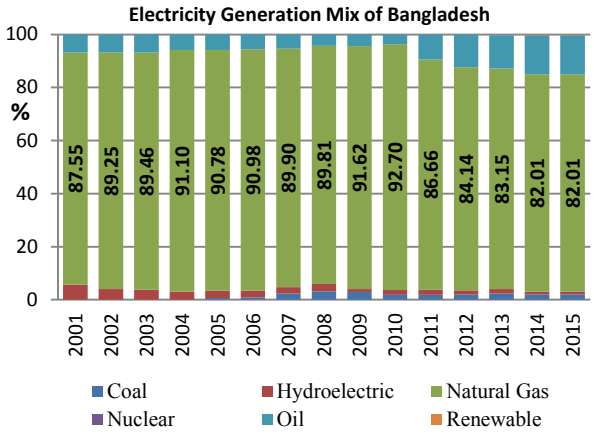


Fig.4. Percentage of Electricity Generation Mix in Bangladesh

Bangladesh fully depends on natural gas which is above 80% of electricity generation mix over the period of 2001 to 2015. Figure 4 shows the oil contribution in generation mix is gradually increasing 2010 to 2015. Coal started to share as electricity generation mix from 2005 and it would be increased more by 2030 [16]. Because of it has characteristic to generate huge amount electricity at a time. Nowadays, nuclear energy source is the most reliable to produce electricity in developed countries for producing burst amount of electricity at a time. But it would be in operation at Rooppur by 2023[17]. Renewable energy sources are the most preferable sources to produce energy. They are abundant and safe. But the figure 4 shows that from the year 2001 to 2015, Bangladesh get this very tiny amount share in electricity generation as sources. So there is lots of opportunity to build renewable energy plant, especially solar energy and bio energy are feasible for Bangladesh.

B. Theory of linear Regression

In statistical, linear regression is a linear approach to modeling the relationship with future value by using existing values and a set of statistical processes for estimating the relationships among variables or data. The known values are existing x-values (independent array or range of data.) and y-values (dependent array or range of data) and the new value is predicted by using linear regression. The equation (1) is for forecasting [18] [19]: $\bar{x}$ *and* $\bar{y}$ are mean value of *x* and *y*.

$$a = \bar{y} - b\bar{x} \quad \text{..... (1)}$$

$$\text{Here } b = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$$

C. Data Analysis and Simulation

Data is collected for World Bank data indicator for analysis and Microsoft excel and MATLab are used for data analysis and evaluate the result. Forecast function is used to get future predictable electricity generation mix share in Bangladesh. There are remarks to simulate the data, these are: if input data is nonnumeric data, output generate error value. If they are empty or contain a different number of data points, function returns the error value.

III. RESULT AND DISCUSSION

A. Comparative Result

In the figures, Comparative studies have been observed between developed (Germany and Australia) and developing country Bangladesh to get forecasting electricity generation mix in Bangladesh. Forecasting is observed considering some factors like environment, global warming, and stability, socio-economic and geographical location.

Electricity generation mix from gas sources in Germany & Australia is pretty low which is between 10-20% for both countries. But Bangladesh share almost 90% of their electricity from gas, which did not come below 80 in past 15 years in figure 5. In figure 6, Electricity production source from coal, Australia is at the top, even though the production source is decreased 80% to 60% and Germany is keeping a constant as a mix around 50% in 2001-2015 and both countries are downwards trends. Whereas Bangladesh produces around 2% electricity from coal and it is upward trend as coal power plant has ability to produce big amount of power. Bangladesh has planned to generate large amount of electricity from coal by 2030. Next Figure 7 shows electricity production from nuclear source where only country exists, Germany which is decreasing their production from 30% to 12% in 2001-2015 time period. Bangladesh has started commission Nuclear power plant by 2024. Following figure 8 is displaying the electricity generation mix from renewable sources, where Germany made great advance from the initial 5% to 30% by 2001-2015. Even though Australia made slight improvement from 8% to 15% in 15 years, Bangladesh is yet to make visible improvement in this sector in spite of there are lots of opportunities on renewable energy in Bangladesh. Eventually, over the time period electricity generation mix changes and near future renewable energy, coal, uranium sources would be added accountable amount in electricity generation mix in Bangladesh.

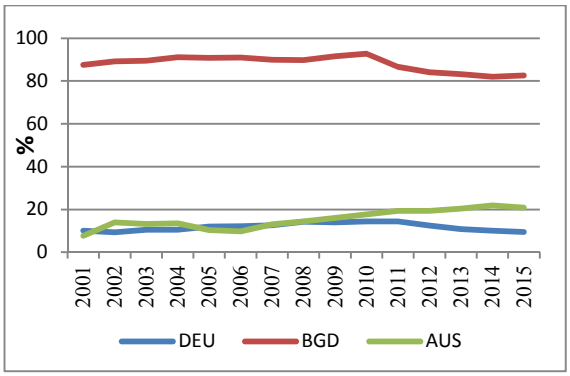


Fig 5. Electricity production from natural gas sources (% of total)

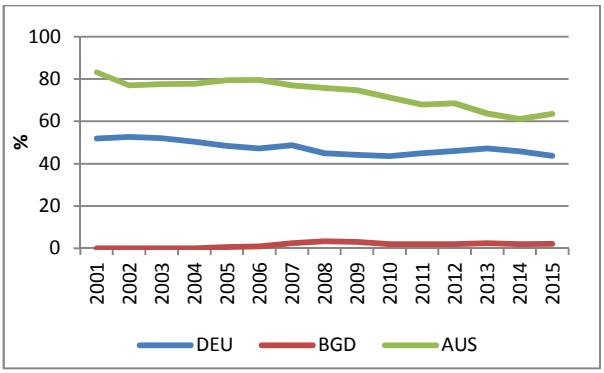


Fig.6. Electricity production from coal sources (% of total)

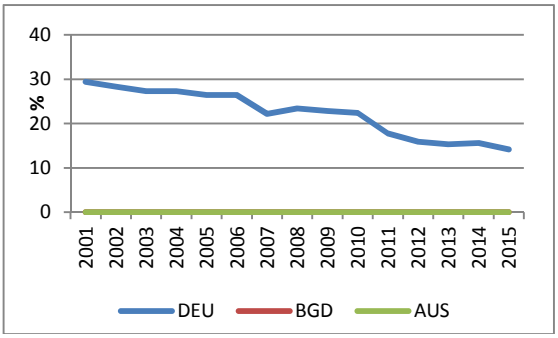


Fig.7. Electricity production from nuclear sources (% of total)

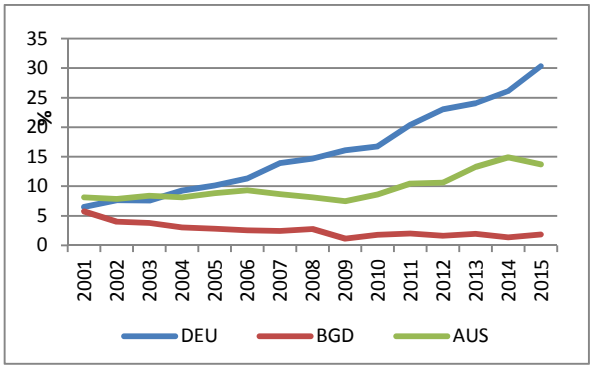


Fig.8. Electricity production from renewable sources (% of total)

B. Simulation Result

The simulation result shows forecasting of electricity generation mix 20016 to 2030 in figure 9. There is significant reduction of natural gas which reaches at 70% from 80% share in electricity generation over the period 2015-2030. Renewable energy share improve to 1.5% in 2030 and same as improvement of coal. So it is the position sign for country to reduce natural gas and improve renewable energy sources, and oil share is also gradually increasing but it should not be. As nuclear source was not in share, simulation function could not predict the future value. But as plan of Bangladesh, nuclear share would be added.

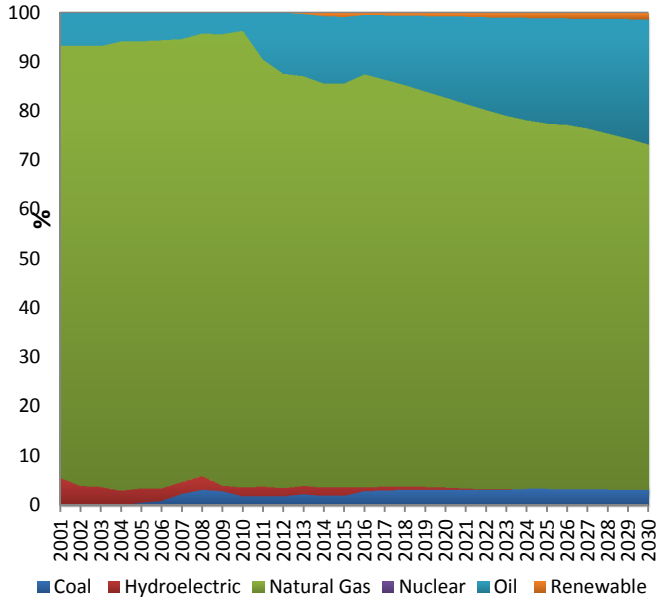


Fig.9. Electricity Generation Mix Forecasting in Bangladesh (2016-2030)

IV. CONCLUSION

In Bangladesh, the population is rapidly increasing and fossil fuel is depleting day by day. The consumption of electricity 310KWh per capita and 76% population get electricity in Bangladesh. Electricity demand is increasing demand day by day. Coal and Uranium will be added as a high percentage in electricity generation mix in Bangladesh. Both sources have the ability to produce a huge amount of electricity to fulfill the demand. Forecasting and comparative result have shown same output and prediction. Renewable sources will be mentionable share in electricity generation sources in Bangladesh according to forecasting result. End of the paper, it is clear that Electricity generation mix build prosper nation considering education environment, global warming, stability, socio-economic and geographical location.

REFERENCES

[1] Mazbahul Ahamad, Fahian Tanin, "Next power generation-mix for Bangladesh: Outlook and policy priorities", Energy Policy 60 (2013) 272–283

[2] Key World Energy Statistics 2017 by IEA Available at: www.iea.org

[3] Annual Report 2016-2017 Available at www.bpdp.govv.bd

[4] Annual Report 2010 Available at www.bpdp.govv.bd

[5] A. Nath, S. Barua, and N. Mohammad, "Electric Power Generation-Mix for Bangladesh and Its Future," in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1–4. DOI: 10.1109/ECACE.2019.8679245

[6] S. Barua, R. Arun Prasath and Dwipen Boruah, "Rooftop Solar Photovoltaic System Design and Assessment for the Academic Campus Using PVsyst Software", International Journal of Electronics and Electrical Engineering Vol. 5, No. 1, February 2017, doi: 10.18178/ijeee.5.1.76-83

[7] Shariful Islam Sharif, Md. Anisur Rahman Anik, Md. Al-Amin, Md. Abu Bakr Siddique, "The Prospect of Renewable Energy Resources in Bangladesh: A Study to Achieve the National Power Demand" Energy and Power 2018, 8(1): 1-6 DOI: 10.5923/j.ep.20180801.01

[8] Saiful Islam, Md. Ziaur Rahman Khan, "A review of energy sector of Bangladesh", 1st International Conference on Energy and Power, ICEP2016, 14-16 December 2016, RMIT University, Melbourne, Australia

[9] David Milligan, "The Future of the power market in Bangladesh", Asian Power Online, Available at: asian-power.com/power-utility/commentary/future-power-market-in-bangladesh

[10] N. Mohammad and Y. Mishra, "Competition driven bi-level supply offer strategies in day ahead electricity market," *2016 Australasian Universities Power Engineering Conference (AUPEC),* Brisbane, QLD, 2016, pp. 1-6. doi: 10.1109/AUPEC.2016.7749349.

[11] Nur Mohammad, Yateendra Mishra, "Coordination of wind generation and demand response to minimise operation cost in day ahead electricity markets using bi-level optimisation framework", IET Gener. Transm. Distrib., 2018, Vol. 12 Iss. 16, pp. 3793-3802, doi: 10.1049/iet-gtd.2018.0110

[12] M. Schäfer, L. Schwenk-Nebbe, J. Hčorsch, B. Tranberg and M. Greiner, "Allocation of nodal costs in heterogeneous highly renewable European electricity networks," *2017 14th International Conference on the European Energy Market (EEM),* Dresden, 2017, pp.1-6. doi: 10.1109/EEM.2017.7981964

[13] Linear Regression—Detailed View, Available at: towardsdatascience.com Access: 10 jan 2018

[14] Rina Haiges, YD Wang,A. Ghoshray, AP Roskhilly, "Optimization of Malaysia's power generation mix to meet the electricity demand by 2050"9th International Conference on Applied Energy, ICAE2017, 21-24 August 2017, Cardiff, UK, Energy Procedia 142 (2017) 2844-2851

[15] World Bank Indication, Available at: databank.worldbank.org.

[16] M. K. Hasan and N. Mohammad, "An Outlook over Electrical Energy Generation and Mixing Policies of Bangladesh to Achieve Sustainable Energy Targets -Vision 2041," *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Cox'sBazar, Bangladesh, 2019, pp. 1-5. doi: 10.1109/ECACE.2019.8679446

[17] S. Hossain, M. Rahaman, I. Tasnim and N. Mohammad, "Optimal Energy Mix and Operation Cost in the Presence of Nuclear and Solar PV Generation," *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Cox'sBazar, Bangladesh, 2019, pp. 1-6. doi: 10.1109/ECACE.2019.8679472

[18] Microsoft Office, Available at http://support.office.com

[19] Anthony Hayter, "Probability and Statistics for Engineers and Scientists", fourth edition, ISBN- 13:978-1-111-82704

Compensation of Frequency Fluctuation in a Single Area Power System Using LQR-VC

Fariya Tabassum
Dept. of Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
f.tabassum@ieee.org

M. S. Rana
Dept. of Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
sohel@eee.ruet.ac.bd

***Abstract*—To maintain the power system stability frequency fluctuation should be eliminated within the shortest possible time. A new control scheme combining linear quadratic regulator (LQR) along with a vibration compensator (VC) is proposed in this paper for the robust control of frequency fluctuation due to load variation in a single area generating system. The proposed LQR controller is used for compensation of error deviation and the augmented VC to damp the frequency fluctuation. The performance of the proposed controller is compared with the automatic generation control (AGC) scheme, conventional proportional-integral-derivative (PID) and LQR controller to demonstrate its effectiveness.**

Keywords—Load frequency control (LFC), linear quadratic regulator (LQR), vibration compensator (VC), cost function.

I. INTRODUCTION

A reliable electrical power system is required for fulfilling the increasing demand of electricity. Stable power system is the main key to maintain the continuity and quality of the supply. With the varying demand of the consumers, the power is changing. Any change in real power affects mainly the system frequency. Maintaining constant frequency during supply is desired in order to maintain the stability and efficiency of the system. The conventional way of controlling the frequency of a generating unit is by using the load frequency control (LFC) loop [1]-[2]. It controls the system frequency by controlling the turbine input [3]. LFC plays an important role in maintaining constant frequency to ensure the reliability of electric power.

A complete mathematical model of LFC along with the necessary diagrams has been discussed in [4]. The step response of frequency deviation due to sudden load variation has been analyzed there and it was shown that using only the basic LFC loop to mitigate the frequency fluctuation was not wise. There was steady-state frequency deviation and the response of the basic LFC loop was oscillatory. The settling time was approximately 8 s and the overshoot was 54.8 % which is pretty large [4]. Settling time is a critical factor for maintaining power system stability and the power system should regain its steady-state condition as soon as possible after any parametric variation.

In literature, various techniques have been adopted to improve the performance of LFC to maintain nominal value of system frequency. One way to restore the frequency to its rated value quickly is to add an integrator with the basic LFC loop. The integral unit monitors the average error over a period of time and will overcome the offset. Thus as the system load changes continuously the generation is adjusted automatically to restore the frequency. This system is called automatic generation control (AGC). It is a feedback control system that regulates the output power of generators to maintain a specified system frequency. Mathematical modeling of AGC system and its step response of frequency deviation has been briefly discussed in [4]. It was shown that the response of AGC was certainly an improvement from the LFC but the AGC system also took a long time to be settled down. The settling time of the AGC was approximately 10 s [4].

Frequency deviation even for a small amount of time is very harmful for the stability of the power system, hence the settling time of the system is needed to be improved. In order to improve the system dynamics even more, a PID controller is proposed in [5]. Proposed controlled system had a negligible overshoot but settled down after approximately 5 s. The controller gains, K_P, K_I and K_D were tuned there by trial and error technique. By using more suitable values of K_P, K_I and K_D, with the help of advanced optimization techniques a better response can be obtained from the controller which was showed in [6]. It was shown that a PID controlled system with advanced optimization technique would settle down the frequency response much quicker and there were no steady-state error.

Moreover, optimal control techniques can be used for better performance. It deals with designing controls for dynamic systems by minimizing a performance index that depends on the system variables. Linear quadratic regulator (LQR), which is an optimal control method, was suggested in [7]. The response of the LQR controller was very satisfactory. The response settled down at approximately 0.7 s. Even though the settling time was reduced but there was a remarkable steady state error after the implementation of the LQR controller.

LQR is an optimal control regulator that tracks a reference trajectory more effectively as compared to traditional controllers like PID. It is a well-known method that provides optimally controlled feedback gains to enable the closed-loop stable and high performance design of system. For the effective control of system frequency an optimal control technique based on LQR is proposed here. Moreover, a vibration compensator (VC) incorporated with the LQR is used to improve the damping by reducing overshoot. The optimal LQR-VC is capable enough to

control the LFC loop effectively in order to minimize the settling time, tackle the overshoot issue and reduce the steady-state frequency deviation.

The remainder of the paper is organized as follows: modeling of the LFC system is discussed in section II. Control methodologies are discussed in section III. Implementation of controllers on LFC system is provided in section IV. Whereas, comparison among the proposed controller and the existing control techniques are discussed in section V. Finally the paper is concluded in section VI.

II. MODELING OF LFC

In electrical power system, nominal frequency depends significantly on the balance of produced and consumed active power. When active power imbalance occurs in any part of the system, it results in changes in the frequency of the entire system. Loads connected to electrical power system are very sensitive to the quality of electrical energy, mainly the frequency component. The relationship between system frequency and load is inversely proportional [8]. The alternation in frequency is sensed by LFC control loop.

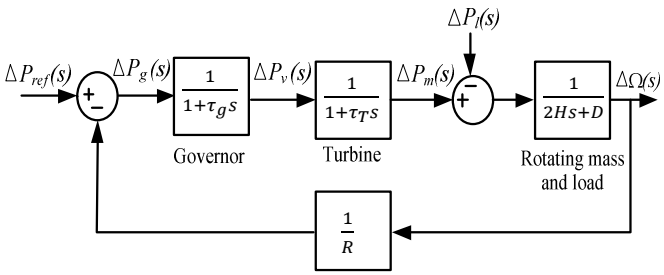


Fig. 1: LFC block diagram of an isolated power system.

The change in frequency and real power are sensed as a measure of the change in rotor angle δ. So the error Δδ needs to be corrected to eliminate the frequency fluctuation.

TABLE I. PARAMETERS OF AN ISOLATED SYSTEM [7]

Parameter	Value
Turbine time constant, τ_T	0.5 sec
Governor time constant, τ_g	0.2 sec
Governor inertia constant, H	5 sec
Frequency sensitive load coefficient, D	0.8
Change in load, ΔP_L	0.2 per unit
Governor speed regulation, R	0.05

This work is done by the primary LFC loop which consists of generator model, load model, prime mover model and governor model. Block diagram of the basic LFC control loop is shown in Figure 1 [7]. The LFC model parameters used for an isolated system are presented in (Table I) [7]. With the primary LFC loop, a change in the system load will result in a steady-state frequency deviation, depending on the

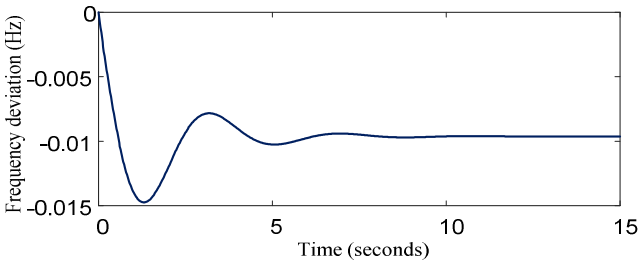


Fig. 2: Frequency deviation step response of LFC.

in a steady-state frequency deviation, depending on the governor speed regulation. The response is highly oscillatory and contains a steady-state error. So using alone the LFC loop to control the frequency is not wise and the response of the system is not satisfactory. Figure 2 shows the response of the LFC loop due to sudden load change of 0.2 p.u.

III. CONTROLLER DESIGN

A. Design of Integral controller

The integral controller is used along with the basic LFC loop to restore the frequency to its nominal value. This scheme is known as AGC. The block diagram of AGC for an isolated system is shown in Figure 3 [7]. The integral controller gain K_i must be adjusted for satisfactory response.

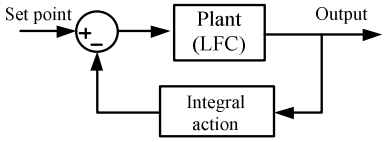


Fig. 3: AGC block diagram of an isolated power system.

B. Design of PID controller

The PID controller is used to improve the dynamic response as well as to reduce or eliminate the steady-state error. PID controllers being popular for its performance over wide range and simplicity, is largely used for controlling the LFC loop [5].

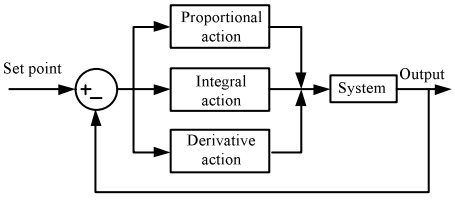


Fig. 4: Block diagram of PID controller.

A simple block diagram of a PID controlled system is showed in Figure 4 [5]. The transfer function of the PID controller is

$$G_C(s) = K_P + \frac{K_I}{S} + K_D s \quad (1)$$

C. Design of Vibration Compensator (VC)

Vibration compensator is used to improve system parameters like reducing steady-state error, reducing resonant peak, improving system response by reducing rise

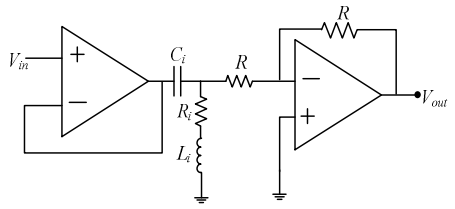


Fig. 5: Structure of the vibration compensator [9].

time. The VC as shown in Figure 5 is introduced in this paper to achieve better damping performance [9]. It helps to enhance the responsiveness and stability of the system. It is used to improve damping and thus reduce overshoot and improve settling time. The feedback loop for the LFC along with the VC is shown in Figure 6. By selecting the proper values of *Li*, *Ri*, and *Ci*, the damping of the LFC loop is improved. The general form of the damping compensator is [9].

$$M = \sum_{i=1}^{N} - K_i \frac{C_i s(R_i + L_i s)}{L_i C_i s^2 + R_i C_i s + 1} \quad (2)$$

where K_i, i = 1, 2,..., N, is the compensator gain of the corresponding mode.

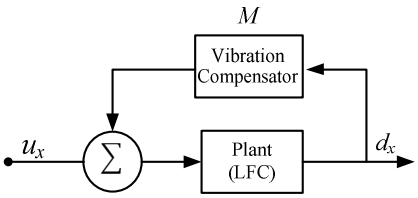


Fig. 6: Feedback loop of the vibration compensator for the LFC.

D. Design of (LQR)

LQR controller is a form of optimal control that provides optimally controlled feedback gains to enable the closed-loop stable and high performance design of systems. Its objective is to determine the optimal control law which minimizes a performance index [6]. In LQR, the system dynamics are described by a set of linear differential equations and the cost is described by a quadratic function and hence the name is linear quadratic regulator. The plant is described by

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t); \end{cases} \quad (3)$$

where, x is called the state vector, $x(t) \in \mathbb{R}^n$, y is called the output vector, $y(t) \in \mathbb{R}^q$, u is called the input (or control) vector, $u(t) \in \mathbb{R}^p$, A is the state (or system) matrix of $n \times n$, B is the input matrix of $n \times p$. C is the output matrix of $q \times n$, D is the feedthrough (or feedforward) matrix (in case where the system model does not have a direct feedthrough, D is the zero matrix), of $q \times p$. $\dot{x}(t)$ is the first derivative of x(t). The problem is to find the vector $K(t)$ of the control law

$$u(t) = -K(t)x(t) \quad (4)$$

that minimizes the value of a quadratic performance index

$$J = \int_0^\infty (x^T Q x + u^T R u) dt. \quad (5)$$

Here Q is an $n \times n$ positive semidefinite matrix, also called the state weighting matrix, R is an $p \times p$ positive definite matrix known as control weighting matrix [10].

$$K = R^{-1} B^T P \quad (6)$$

where P is the unique positive definite solution of the algebraic Riccati equation (ARE) [9].

$$A^T P + PA + PBR^{-1} B^T P + Q = 0 \quad (7)$$

From the state-space representation of the plant values of A, B, C and D matrices can be found. By proper selection of the Q and R matrix the value of P and therefore K is determined which will place the close-loop poles of the system in the most perfect place using the LQR method with respect to the cost function. A simple block diagram of an LQR controlled system is shown in Figure 7.

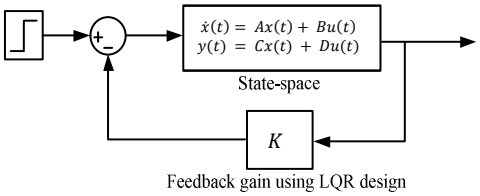


Fig. 7: Block diagram of an LQR controlled system.

IV. Implementation of Controllers on LFC

An isolated system with parameters given in TABLE I is considered and from Figure 2 it is seen that the frequency deviation step response of basic LFC loop is oscillatory and contains a steady state error.

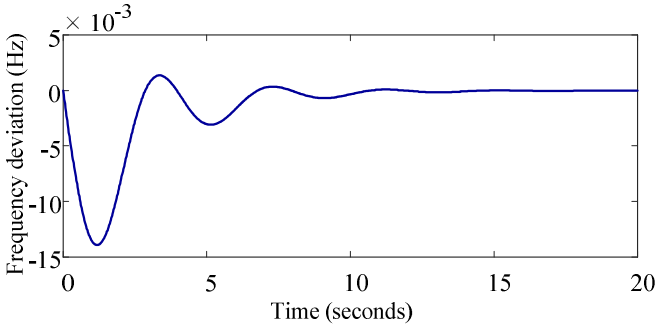


Fig. 8: Frequency deviation step response of AGC.

By using an integral controller and tuning the value of gain K_i=7 a response is obtained which has less steady-state error but both the overshoot and the settling time are quite

large. From Figure 8 it is seen that settling time is approximately 10 s. By using PID controller and tuning the values of K_P, K_I and K_D, a better response is obtained with negligible steady-state error and a smaller settling time.

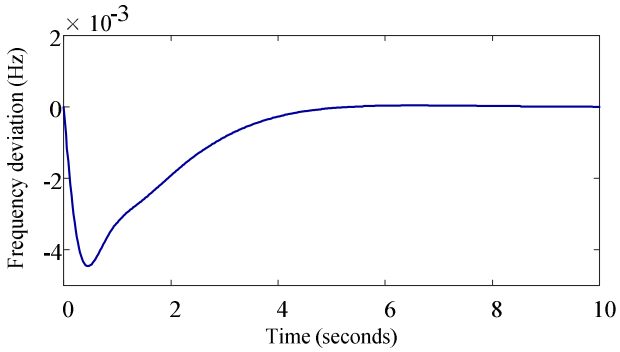


Fig. 9: Frequency deviation step response for PID controller.

When the values of controller gains are considered as $K_P = 30$, $K_I = 25$ and $K_D = 30$, then the response of the system is as shown in Figure 9 which has a settling time of approximately 5 s. The frequency deviation step response after using LQR controller is shown in Figure 10 which indicates a settling time of approximately 0.63 s but suffers from large steady-state error.

Figure 11 shows that when the proposed control scheme is applied to the same isolated system then settling time is reduced to a value of approximately 0.01 s with a negligible steady-state error and overshoot.

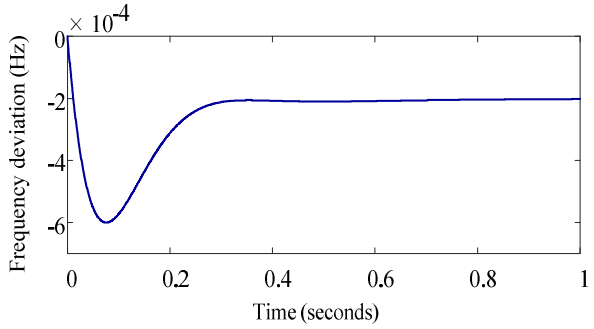


Fig. 10: Frequency deviation step response for LQR controller.

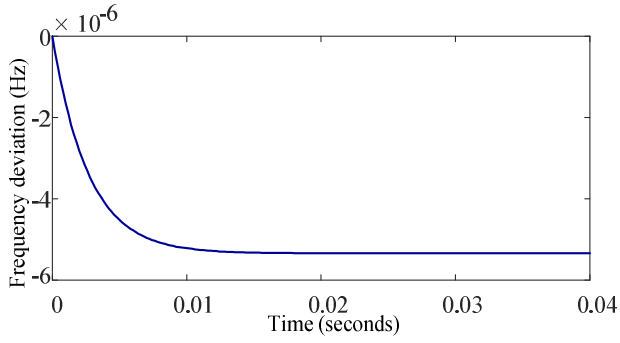


Fig. 11: Frequency deviation step response for proposed controller.

V. COMPARISON AND RESULT

A relative comparison among different control techniques for controlling the frequency of an isolated system is shown in Figure 12.

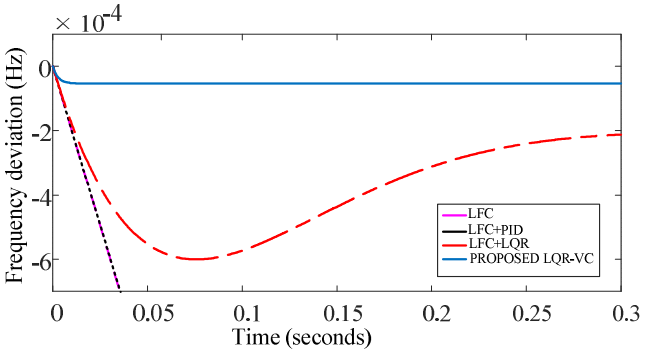


Fig. 12: Comparison among different control techniques.

From the comparison it is clear that by using proposed LQR controller along with a VC the best result is obtained compared to the existing control techniques. The response settles down after approximately 0.01 s which is approximately 60 times faster than the existing LQR control technique. So it can be concluded that proposed controller is certainly an improvement from the existing controllers. TABLE II summaries the results for different control techniques implemented on LFC and hence justifies that the proposed controller gives better result.

TABLE II. Performance parameters of different control techniques

Controller	Settling Time (seconds)	Overshoot (%)	Steady-State Error
PID	4.67	0.0001	1.0558e-07
LQR	0.63	0.02	2.0228e-04
Proposed LQR-VC	0.0105	0	5.3422e-05

VI. CONCLUSION

In this paper, design and implementation of an optimal LQR controller augmented with a VC is proposed to mitigate the frequency disruption of an electric power generating unit. The performance of the proposed controller is compared with the existing control techniques. The proposed controller eliminates the frequency fluctuation approximately 60 times faster than the existing LQR control method. The protective scheme is designed and applied in an isolated system but with some modifications it can be applied to the interconnected power grid.

REFERENCES

[1] Thanh Ngoc Pham, Hieu Trinh and Le Van Hien, "Load Frequency Control of Power System With Electrical Vehicles and Divers Transmissions Links Using Distributed Functional Observers," *IEEE Transactions on Smart Grid,* vol.7, no. 1, pp. 238-252, Jan. 2016.

[2] M. A. Rahman, M.S. Rana and M.S. Anower, "Indemnity for Frequency Disruption in a smart Grid during Cyber-Attack," in 2nd International Conference on Electrical & Electronic Engineering (ICEEE), Rajshahi, 2017, pp. 1-4.

[3] Miaomiao Ma, Chunyu Zhang, Xiangjie Liu and Hong Chen, "Distributed Model Predictive Load Frequency Control of the Multi-Area Power System After Deregulation," *IEEE Transactions on Industrial Electronics,* vol. 64, no. 6, pp. 5129-5139, June 2017.

[4] Orosun. M. M, Orosun R. O. and Adamu S. S., "Modeling and Simulation of Automatic Generation Control System for Synchronous Generator with Model Predictive Controller," *Zimbabwe Journal of Science & Technology*, vol. 11, pp. 142-157, Nov. 2016.

[5] Pedro Mercader, Karl Johan Åström, Alfonso Baños and Tore Hägglund, "Robust PID Design Based on QFT and Convex–Concave Optimization," *IEEE Transactions on* Control Systems Technology, vol. 25, no. 2, pp. 441-452, Mar. 2017.

[6] Wen Tan "Unified Tuning of PID Load Frequency Controller for Power Systems via IMC," *IEEE Transactions on Power Systems,* vol. 25, no. 1, pp. 341-350, Feb. 2010.

[7] Hatem Elaydi and Mohammed Wadi, "Optimal Controller for Single Area Load Frequency via LQR and Legendre Wavelet Function," *Journal of Automation and Control,* vol.3, no. 2, pp. 43-47, Apr. 2015.

[8] H. Gozde, M. C. Taplamioglu, MC, I Kocaarslan and M. A. Ve Senol, "Particle swarm optimization based pi-controller design to loadfrequency control of a two area reheat thermal power system," *Journal of Thermal Sciences & Technologies*, vol.30, no. 1, pp. 13-21, Jan. 2010.

[9] M. S. Rana, H. R. Pota and I. R. Petersen, "Nonlinearity effects reduction of an AFM piezoelectric tube scanner using MIMO MPC," *IEEE/ ASME Transactions on Mechatronics,* vol. 20, no. 3, pp. 1458-1469, June 2015.

[10] Le Wei and Fang Fang, "H_∞ -LQR Based Coordinated Control for Large Coal-Fired Boiler–Turbine Generation Units," *IEEE Transactions on Industrial Electronics,* vol. 64, no. 6, pp. 5212-5221, June 2017.

Low Power Consuming Solar Assisted Vapor Compression Refrigerator To Preserve Emergency Medicines And Vaccines In Rural Areas

Akbar Hossain
Institute of Photonics
University of Eastern Finland
Joensuu Campus
Joensuu,Finland
akbarhossainakbar8888@gmail.com

Shovan Saha
Department of Electrical & Electronic Engineering
American International University-Bangladesh
Dhaka, Bangladesh
shovansaha20162016@gmail.com

Md Arifur Rahman
Department of Electrical & Electronic Engineering
American International University-Bangladesh
Dhaka, Bangladesh
arif1622ccr@gmail.com

Mahbub Alam Himel
Department of Electrical & Electronic Engineering
American International University-Bangladesh
Dhaka, Bangladesh
mahbubalamhimel@yahoo.com

***Abstract*—Solar Assisted Refrigerator (SAR) can be manufactured using elements like a solar panel, evaporator, expansion valve, compressor, and the inverter, etc. This paper gives an evidence to the concept of using solar energy to preserve emergency lifesaving medicines and vaccines at low temperature. The target storage temperature range was between -15°C and 30°C. The model refrigerator was designed and constructed totally by using local technology and placed in outdoor conditions to observe the effect of different weather conditions. By monitoring the temperature variation and compressor current after every one-hour, power consumption and the coefficient of performance (COP) were determined. The overall coefficient of performance (COP) was calculated 4.89 for the refrigerant "Iso-Butane(R600a)". The temperature and pressure profile have been observed using a highly accurate digital thermocouple thermometer and accurate pressure gauge to find out the heat transfer rate, suction and discharge pressure. Total power consumption per day was also calculated by observing the temperature and pressure profile. It proved that the SAR can be implemented in rural areas and the areas where grid-connected electricity is not available to preserve emergency lifesaving medicines and vaccines at low temperature. This could be a means to save millions of people's lives in the disastrous regions and at the refugee camps.**

Keywords—Solar Energy, Refrigerator, Refrigerant, Temperature and Pressure profile, Power Consumption, Heat loss, C.O.P, Storage, Rural Area, and Refugee Camps.

I. INTRODUCTION

Solar Assisted Refrigerator is a refrigeration system which works with the help of solar energy which is harvested from the sunlight. In hot climates and in those regions where grid-connected electricity is not available at all, people can use the SAR to preserve food so it does not rot. Alternatively, SAR can also be used to store vaccines at the proper temperature. When sunlight is not available, at that time the SAR can work using the power from batteries to run the refrigerator.

Solar energy has become very attractive for cooling purposes nowadays. It can provide a cheap and clean source of electricity all over the world for cooling and refrigeration purpose. Solar refrigeration technology collects the energy from the radiation of the sun and uses that energy for cooling purpose. This system is a lucrative choice for medicine and vaccine storage facility where an adequate amount of sunlight is available but has instability in electricity supply. It is also suitable for the areas which are located far from the grid. The solar system can be divided into two main categories, such as thermal driven and electrically driven systems. Photovoltaic panels are used as an electrically-driven system to produce electricity which is then used conventionally to run the refrigeration machine. The system consists of four basic components: PV modules, a battery storage system, charge regulator, and vapor- compression unit. In practical applications, solar PV refrigeration units are used in luxurious boats and/or in remote areas for food and vaccine preservation.[1]

Some research works are already done regarding Solar Assisted Refrigerator, but at that time the works had some limitations. In paper [2], researchers used solar panel, charge controller, battery for power storage and a DC powered cooling unit but it did not have sufficient battery back-up. On the other hand, in research work [3], the refrigeration system had enough amount of battery back-up for cloudy day protection. The problem of this both [2], [3] design was that there was no scope of running the system using AC power. If there was an option of running the system using AC power and had more battery storage then it would have been a better system. In paper [4], they used the MPPT unit but they did not keep any battery storage option. In paper [5], they worked based on Peltier device and that refrigerator has the ability to warm up the food too. Peltier device was used for cooling as a replacement of compressor which is commonly used in a traditional refrigeration system. The problem of using a Peltier device is that although it works well for cooling a small amount of food or medicine, it is not good for cooling a large load. Peltier device was also used by the researchers of paper [6] and their proposed refrigerator can only perform cooling for 1 hour 35 minutes from full charge condition which is not suitable for practical use.
The proposed Solar Assisted Refrigerator (SAR) is capable of solving all limitations of previous works and it is a very good solution for storing the medicine and vaccines at low temperature.

II. WORKING PRINCIPLES AND EXPERIMENTAL SETUP

The refrigeration cycle is very simple to understand. In figure 1, the refrigeration cycle is shown. Firstly, a cold liquid is passed around the object which needs to be cooled. This process will remove heat from the object. There is a throttling device that helps to reduce the temperature of the liquid. A throttling device is actually nothing but a capillary tube. For efficient throttling, an inlet liquid must be high pressured liquid. Throttling device essentially reduces the pressure of the

liquid. As the pressure decreases, the boiling point of the refrigerant comes down and it starts to evaporate. For the continuation of evaporation, the temperature drops down. If the temperature is measured at the inlet and the outlet of the capillary tube, a temperature drop will become noticeable. Then the evaporating refrigerant is passed through the evaporator tube of the refrigerator. After that, the refrigerant starts to absorb heat from the objects stored in the refrigerator. Thus, the refrigerant further evaporates and turns into pure vapor. This creates the required refrigeration effect.

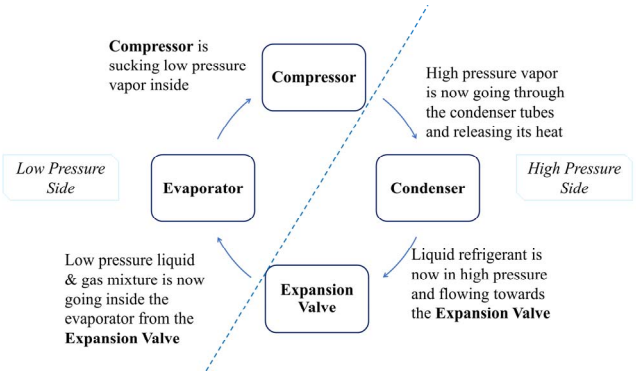


Figure 1: Refrigeration Cycle [8]

After absorbing heat from the refrigerator chamber, the refrigerant is in high temperature and low-pressure state. In order to repeat the refrigeration process, the pressure of the refrigerant has to increase. For this purpose, a compressor is used. The compressor increases the pressure of the refrigerant to its high-pressure state. Since it is compressing gas, the temperature of the refrigerator will also increase. After the compression, the refrigerant turns into a high pressure and high-temperature vapor. To transform the refrigerant from the vapor state to the liquid state, a heat exchanger is used. The heat exchanger is placed outside of the refrigerator and known as a condenser. It releases heat on the environment and as a result, the refrigerant will get condensed and turn into the liquid from its vapor state. Finally, the refrigerant will return to its beginning state and that is high-pressure liquid. This cycle is repeated again and again for running the refrigerator. This cycle is recognized as a vapor compression cycle.

In figure 2, a block diagram of the Solar Assisted Refrigerator is shown. The model is created using PV cell, charge controller, refrigerator, inverter and battery.

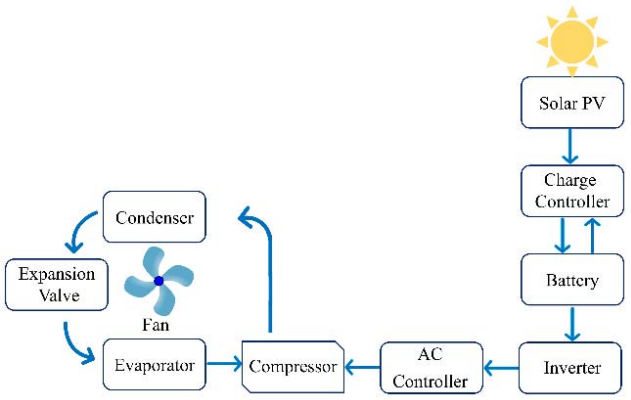


Figure 2: Block Diagram of the SAR [9]

At first, the PV cells will collect the irradiation from the sunlight and then produce electrical energy which is direct current (DC). This energy is stored in the battery and a charge controller is used to control the charging of the battery. An inverter is used to convert the direct current (DC) to alternating current (AC) for running the compressor. Since, the compressor is an AC compressor, an inverter is required for the SAR. The SAR can run by using the power from the battery storage or directly from the national grid as the compressor uses alternating current.

III. MATHEMATICAL MODELING

The Solar Assisted Refrigerator (SAR) was designed and constructed by using local technology and locally available materials the project was successfully implemented.

TABLE I

Description	Specifications
Outside Dimension	2.5 feet × 2 feet × 2 feet (H × W × L)
Refrigerant Type	Iso-Butane (R600a)
Charged Refrigerant Mass	120 grams (Approx.)
Compressor Type	Hermetically sealed
Compressor Oil Charge	200 ml (Approx.)
Power Source	220 Volt AC, 50 Hz
Total Weight	50 kg. (Approx.)
Compartment Type	1 Unit
Air Circulation	Forced Convective Fan
Calibration	Temperature Monitoring
Compressor Control	On/ Off Single Speed
Temp., Pressure & Humidity During Test	20°,102 kPa,60% RA [Based On 2017 January]
Antifreeze	Defrost*
Refrigerant Circuit (RC)	Total 4 'T' Connector Used for Compressor & Expansion Device

In table I, all specifications are given which are required for fabricating the SAR.

TABLE II

Four states				
	State 1	State 2	State 3	State 4
T	2.5°C	50.9°C	46.3°C	2.5°C
P	320 kPa Suction Pressure	1200 kPa Delivery Pressure [Provided by Manufacturer]	1200 kPa [No Loss Assumed]	320 kPa [Ideal Value is Assumed]
h	252 kJ/kg	279.3 kJ/kg	117.8 kJ/kg	117.8 kJ/kg
S	0.930 kJ/kgK	0.930 kJ/kgK	0.4245 kJ/kgK	S4 = 0.4436
X	1	1	0	x4 = 0.3

Here,

m' means mass in $\frac{kg}{h}$

h means heat in $\frac{kJ}{kg}$

1. Work Done by the Compressor = $m'(h_2 - h_1) = 3\ (274.3 - 252) = 82.3\ kW$
2. Cooling Load In the Evaporator = $m'(h_1 - h_4) = 3\ (252.0 - 117.8) = 402.30\ kW$

3. Heat Rejected by the Condenser = $m'(h_1 - h_3) = 3\ (279.3 - 117.8) = 484.6\ kW$
4. Heat Rejected = (Work Done + Cooling Amount) = $82.3 + 402.30 = 484.6\ kW$

$X_4 = \frac{117.8-55.2}{251.9-55.2} = 0.318 \cong 0.32$

$S_4 = S_f + X_4(S_g - S_f) = 0.2165 + 0.32(0.93 - 0.2165) = 0.4436$

Coefficient of Performance,

C.O. P = $\frac{Cooling\ Load}{Work\ Done} = \frac{402.3}{82.3} = 4.89$

A. Equipment Rating Selection

Before starting to build the Solar Assisted Refrigerator (SAR), the necessary calculation was done to determine the compressor power, pipe length for the compact condenser, evaporator and for the capillary tube. The rating of solar panel, battery and inverter size also calculated carefully and precisely.

B. Compressor Power Rating

The total volume of the chamber inside the SAR= 86.5 liter. Since the medicine weighs lighter than the same volume of water, the weight of 35-liter water is considered.

Q = Heat Energy
m = Mass
ΔT = Change in Temperature
W = Work
P = Power
H = Efficiency
ΔH_f = latent heat of fusion (kJ/kgK)
C_{p1} = specific heat of ice (kJ/kgK)
C_{p3} = specific heat of water(kJ/kgK)

Calculation for determining the heat required to transform -10°C ice to 0°C ice:

ΔT = 273 K – 263 K= 10 K

Q_1 = m×C_{P_1}×ΔT
=35 × 2.108 ×10
=737.8 kJ/s

Calculation for determining the heat required to transform 0°C ice to 0°C water:

ΔT = 273 K – 273 K= 0 K

Q_2 = m×C_{P_2}
= 35×335
= 11725 kJ/s

Calculation for determining the heat required to transform 0°C water to 10°C water:

ΔT = 283 K-273 K= 10 K

Q_3 = m×C_{P_3}×ΔT
= 35×4.2×10
=1470 kJ/s

Total Heat Energy,

Q=Q_1+Q_2+Q_3
= 737.8+11725+1470 = 13932.8 kJ/S
= $\frac{13932.8}{24\times60}$ kJ/m [For 1 day]
= 9.68 kJ/minute

C.O.P = 4.89

Efficiency, η = 30%

It is known that, C.O.P = $\frac{Q}{W}$

W= $\frac{Q}{C.O.P}$
= $\frac{9.68}{4.89}$= 1.98 kJ/minute

P = $\frac{W}{\eta} = \frac{1.98}{60\times0.3}$
= 0.10992 kw
= 109.92 watts

So, compressor power rating should be minimum 109.92 watts. Here, the power rating was calculated by considering the water, but for the medicine the C_P value is less than water. So, for the medicine, a less powerful compressor can be used. If less than 35 kg medicine is stored, then the compressor power rating will also decrease accordingly.

Considering the above-mentioned facts, a 110 watts compressor was chosen which can easily run the SAR without any trouble.

C. Pipe Length for the Compact Condenser, Evaporator and for the Capillary Tube

After calculation, it was determined that exactly fifty (50) feet evaporator pipe is sufficient for cooling the SAR. Ten (10) feet capillary tube and thirty (30) feet copper tube also used for condenser to get the optimum output from the SAR. The diameter of the copper tube was ¼ inch for both the compact condenser and the capillary tube.

D. Rating of Solar Panel and Battery

Solar Panel Selection:

The total energy required for the SAR = 133 watts

Total rating of the SAR= Total Connected Load × Operating Hours = 133×6 = 798 watt-hours [Compressor remains ON 15 minutes per hour] .[7]

So, 200-watt peak (Wp) PV panel was selected.

Battery Selection:

Power output of the PV panel = Peak Power Rating × Operating Factor = 200×0.75 = 150 watts

Actual power obtained (It is less than the Power Output due to low combined efficiency of the system) = Power output of the panel × combined efficiency = 150 × 0.85 = 127.5 watts

Energy produced by one 200 W_p panel in a day = Actual power output of the panel × 7 hours/ day = 127.5×7 = 892.5 watts-hour .[7]

Rating of the Battery = $\frac{\text{Total watt–hours rating of the SAR}}{\text{12 V battery}} = \frac{798}{12}$
= 66.5 ampere hours of reserve battery power

To preserve medicine at least for two days = 66.5 × 2 = 133 ampere hours of reserve battery power is required.

150 ampere hours of reserve battery was selected considering various losses.

E. Rating of the Inverter & Solar charge controller

Inverter Selection:

Total load = 133 Watts

Let's assume power factor is 0.8, so 1.18 VA = 1 Watt

Required inverter size = 1.18 × 133 = 156.9 VA

200 VA inverters was selected because 156.9 VA inverter is not available in the market.

Solar Charge Controller Selection:

Solar charge controller rating = $\frac{\text{Panel's wattage rating}}{\text{Voltage rating}} = \frac{133}{12}$
= 11.08 Ampere

So, 15 Amps solar charge controller was used.
To implement this SAR, walls with the thickness of four (4) inch was used along with several layers of polyurethane foam and 97% reflective aluminum foil.
A digital N-type thermostat was used for maintaining the precise temperature of the SAR. Furthermore, four pressure gauges were installed for analyzing the pressure profile.



Figure 3: Attachment of Evaporator Tube with the SAR chamber



Figure 4: Refrigerator Chamber with Wheels



Figure 5: Pressure Gauges

In figure 3 and figure 4, the evaporator and the refrigerator chamber are visible. Wheels were added with the SAR for easy portability of the system. Four (4) pressure gauges were used to find out suction pressure and discharge pressure. These gauges can measure up to 300 psi. Their picture can be seen of figure 5.
High accuracy digital thermocouple thermometer was used at the time of observing the temperature profile. On figure 6, the thermometer, along with the refrigerator can be seen. The thermometer probes can be seen on figure 7.

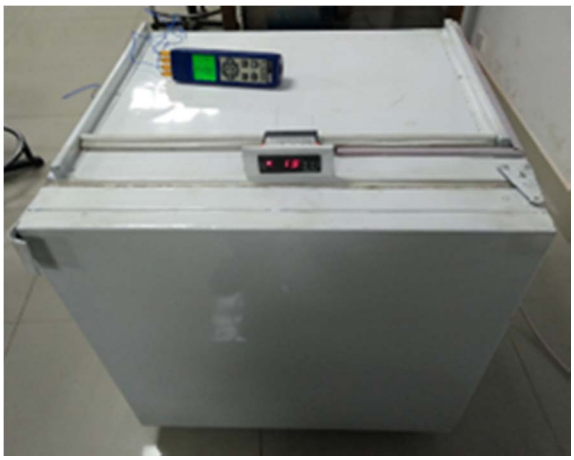


Figure 6: Top Side View of the Refrigerator

IV. Results Analysis & Critical Design Review

From figure 7 it can be seen that, T_1 was placed at the top of

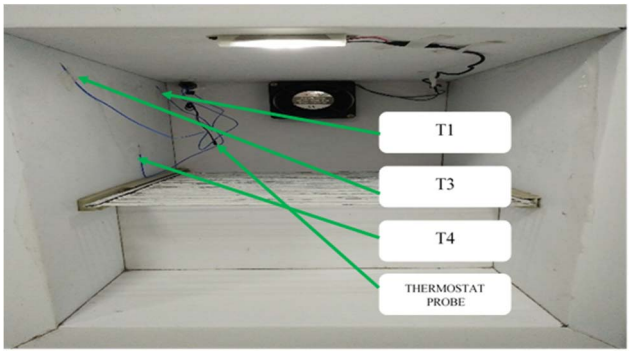


Figure 7: Themometer Probes Placed at Various Points Inside the Refrigerator Chamber

the left side wall near the back wall. T_3 was placed on the left side wall, close to the door. T_4 was placed at the bottom part of the same left wall. T_2 was placed at outside of the refrigerator.

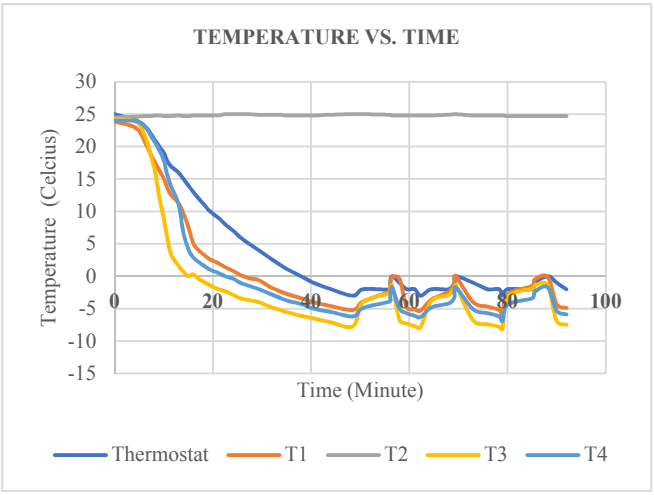


Figure 8: Temperature Versus Time

The figure 8 indicates that compressor remains off from 48.56 minutes to 56 minutes, 62.3 minutes to 69.16 minutes and 79 minutes to 85.55 minutes respectively. This graph indicates that outside temperature was almost 24.8°C and the desired set temperature was -2°C. At that condition, approximately 7 minutes to were needed to achieve the desired temperature -2°C to -1°C. It is the evidence that the heat loss is very low for this Solar Assisted Refrigerator.

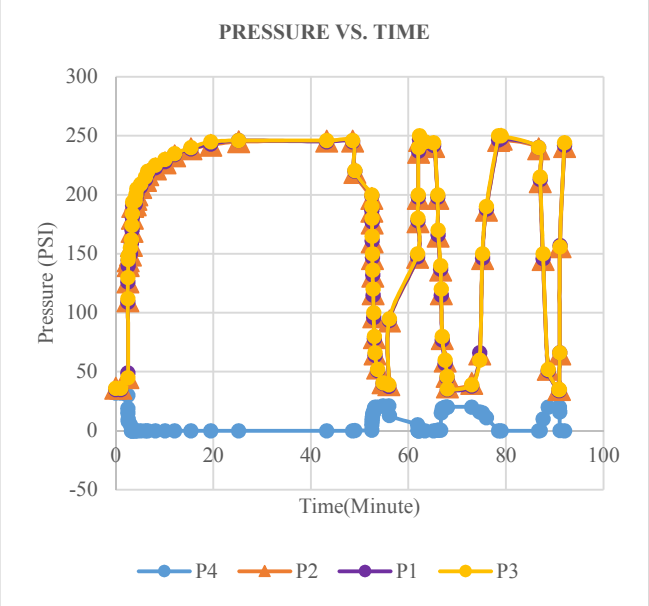


Figure 9: Pressure Versus Time

The figure 9 represents the *Pressure Vs Time Curve* (P_1, P_2, P_3, P_4,) all together. This figure represents the pressure P_1, P_2 , P_3 as the discharge pressure and P_4 as the suction pressure. All discharge pressure values are almost same and show the approximately same characteristics. Discharge pressure always increases when compressor turn ON. Suction pressure value always increases when the compressor turns OFF. Only the suction pressure will be zero when the compressor turns ON. Figure 9 also indicates that when discharge pressure value is at maximum position, at that time suction pressure value is at minimum position. On the other hand, when discharge pressure value is at minimum position at that time suction pressure value is at maximum position. This figure clearly indicates that discharge pressure and suction pressure is vice versa.

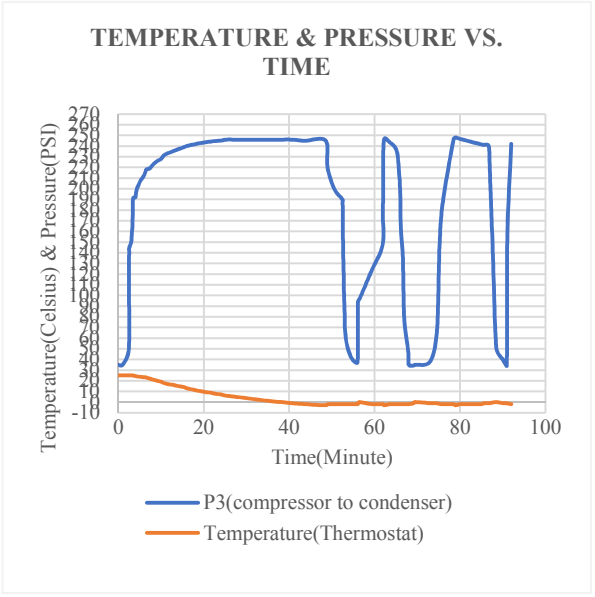


Figure 10: Temperature and Pressure (Compressor To Condenser) Versus Time

Thermostat temperature and the pressure for compressor to condenser is shown in figure 10. The figure 10 indicates that when the temperature was decreasing to the desired set temperature -2°C, at that time, discharge (compressor to condenser) pressure was increasing up to 248 psi. So, it is clear that compressor was also at working condition at that time. Additionally, it proves that when compressor is at running condition, at that time temperature is decreasing but oppositely discharge pressure is increasing.

A. Heat Loss

Outside temperature, $\boldsymbol{T_{out}}$ = 24.5 °C
Inside temperature of the SAR, $\boldsymbol{T_{in}}$ = -2°C
$\boldsymbol{T_{out}} - \boldsymbol{T_{in}}$ = 24.5°C – (-2°C) =26.5°C = 299.5 K
Area, A= 30 inch × 24 inch = 0.464 m^2
Wall thickness, x= 4 inch =0.1016 m
Thermal conductivity of Aluminum sheet, $\boldsymbol{K_1}$= 205 W / m K
Thermal conductivity of 97% Aluminum Foil, $\boldsymbol{K_2}$= 0.34 W / m K [8]
Thermal conductivity of Polyurethane Foam, $\boldsymbol{K_3}$=0.03 W / m K
Thermal conductivity of Air, h=0.025 W / m K

$$Q = \frac{A(T_{out} - T_{in})}{\frac{1}{h_{out}} + \frac{1}{h_{in}} + \frac{x}{K_1+K_2+K_3}}$$

$$Q = \frac{0.464(299.5)}{\frac{1}{0.025} + \frac{1}{0.025} + \frac{0.1016}{205+0.34+0.03}}$$

Q = 1.73 watt
So, heat loss is 1.73 watt.

B. Energy Consumption

Current Consumption, I = 0.7 Ampere [Measured with the help of clamp meter]
Voltage, V = 220 Volt
Power factor = 0.8
Compressor works total 6 hours per day
So, total Energy consumption per day = $\frac{220\times0.7\times0.8\times6}{1000}$ = 0.7392 kilowatt-hour
Total energy consumption per month = (0.7392 × 30) = 22.176 kilowatt-hour
It is expected that the SAR will consume 22.176 kilowatt-hour energy per month.

C. COP Calculation (Practical)

Efficiency, η = 30%
Total Q = 9.68 kJ/minute
P = 110 watt
W = P η = 110 Watt × 0.3 = 33 J/ Minute
C.O.P = Q/W
$= \frac{9.68\times1000}{33\times60}$
= 4.89
COP = 4.89

D. Verification of Using the Correct Rating of Solar Panel, Battery and Inverter Size

Measured energy consumption = 220×0.7×0.8 = 123.2 watts.

At the time of implementing the SAR, a 200-watt peak (Wp) PV panel was used that can produce 127.5 watts.

As the measured energy consumption is slightly less than the produced energy per second by the PV panel, the SAR can run properly. Since the measured energy consumption is perfectly matched with the produced energy by the PV panel, so the solar charge controller, battery and inverter rating selection was also correct for the SAR for running it properly.

V. Cost Analysis

In the following table, the cost of the project is described. Solar PV panel, battery and inverter are the most expensive items of the project. If these components could have procured at a lower price, then the cost of this project would have reduced significantly.

The total cost of the project is 22,410 Bangladeshi Taka (Approximately). The cost is very affordable. Locally available material and technology were used for keeping the construction cost low. If the SAR unit is mass produced, the cost of each unit will reduce drastically. Then the price of per unit of this SAR system will be in the grasp of the poor rural hard-working people of Bangladesh who are deprived of grid connectivity.

TABLE III

Serial	Name of The Equipment	Price (BDT)
1	Compressor	1,900/-
2	Digital Thermostat	1,100/-
3	Polyurethane Foam	900/-
4	Metal Sheet	1,200/-
5	Wheels	350/-
6	Compact Condenser	400/-
7	Insulation Material	350/-
8	Evaporator Tube	960/-
9	Refrigerant (R-600a)	500/-
10	Solar Panel (PV Array)	4,000/-
11	Charge Controller	750/-
12	Inverter	4,800/-
13	Battery	5,000/-
14	Others	200/-
	TOTAL:	22,410/-

VI. Conclusion

By using traditional technologies and locally available materials, the fabrication of the Solar Assisted Refrigerator was implemented successfully. The temperature profile for different points at inside and outside of the SAR was analyzed. In addition, the pressure profile for discharge pressure and suction pressure was also analyzed. The C.O.P of the system is 4.89 that is the highest C.O.P for the refrigerant "Iso-Butane (R600a)". The heat loss and the power consumption were also calculated for the SAR. Calculations show that the heat loss and the power consumption are very low as the four-inch thickness wall combined with several layers of polyurethane foam and 97% reflective aluminum foil were used. It can be said without a doubt that the proposed SAR model is one of the best choices for saving energy consumption as the heat loss is very low and the coefficient of performance is very high. The price of the SAR very affordable. If this SAR can be made available to the people of rural areas where there is no grid connectivity, then the SAR can contribute in saving a lot of lives by preserving live saving medicines at accurate temperature and thus contributing into the prosperity of the country.

References

[1] C. Del Per, F. Buter, M. Buffoli, L. Piegari, L. Capolongo and M. Fattore, "Feasibility study of a Solar Photovoltaic Adaptable Refrigeration Kit for remote areas in developing countries," IEEE, pp. 701-708,2015.

[2] Fatehmulla, A., Al-Shammari, A. S., Al-Dhafiri, A. M., & Al-Bassam, A. A. (2011, April). Design of energy efficient low power PV refrigeration system. In Electronics, Communications and Photonics Conference (SIECPC), 2011 Saudi International (pp. 1-5). IEEE.

[3] Henriques, J. J., Foster, B. T., Schnorr, W. G., & Barton, R. (2012, October). Implementation of a Mobile Vaccine Refrigerator with Parallel Photovoltaic Power Systems. In *Global Humanitarian Technology Conference (GHTC), 2012 IEEE* (pp. 128-131). IEEE.

[4] Du, J., Yang, F., & Wu, J. (2012, May). Design and analysis of semiconductor refrigeration system powered by PV Cells. In Industrial Electronics (ISIE), 2012 IEEE International Symposium on (pp. 286-291). IEEE.

[5] Attavane, P., Arjun, G. B., Radhakrishna, R., & Jadav, S. R. (2017, May). Solar powered portable food warmer and cooler based on peltier effect. In Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2017 2nd IEEE International Conference on (pp. 1975-1978). IEEE.

[6] Arjun, K. G., Pruthviraj, B. G., Chethan, K. Y., & Rashmi, P. (2017, May). Design and implementation of peltier based solar powered portable refrigeration unit. In Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2017 2nd IEEE International Conference on (pp. 1971-1974). IEEE.

[7] P. Kumar and D. Thanki, "Synchronized Operation of Grid Power, Solar Power, and Battery for Smart Energy Management", *Handbook of Research on Power and Energy System Optimization*, pp. 255-291. Available: 10.4018/978-1-5225-3935-3.ch008 [Accessed 14 April 2019].

[8] Yuyan Zhang and S. Tadigadapa, "A Novel Immunosensing Technique Based on the Thermal Properties of Biochemicals", *IEEE Sensors, 2005.*. Available: 10.1109/icsens.2005.1597632 [Accessed 14 April 2019].

[9] "Appliances", Real Simple, 2018. [Online]. Available: https://www.realsimple.com/foodrecipes/toolsproducts/appliances/how-does-refrigerator-work.2018. [Accessed: 04- Apr- 2018].

[10] "For Refrigeration Problems, a Magnetically Attractive Solution", NIST, 2018. [Online].Available: https://www.nist.gov/newsevents/news/2009/01/refrigeration-problemsmagnetically-attractive-solution. [Accessed: 04- Apr- 2018].

Robust Control of Terminal Voltage of an Isolated Electric Power Generating Unit

Fariya Tabassum
Dept. of Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
f.tabassum@ieee.org

M. S. Rana
Dept. of Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
sohel@eee.ruet.ac.bd

***Abstract*—The performance of an automatic voltage regulator (AVR) system controlled by an optimal linear-quadratic-Gaussian (LQG) controller augmented with an integral action is investigated in this article for maintaining constant output voltage. The aim of this controller is to keep stable terminal voltage of a single area power system during sudden load change. To verify its efficacy, a comparison is conducted with some existing controllers for instance the rate feedback stabilizer, proportional-integral-derivative (PID) controller, and linear quadratic regulator (LQR). The comparison is done on the basis of some important transient response characteristics and the proposed control scheme shows strong robustness against voltage variation due to load change.**

Keywords—Automatic voltage regulator (AVR), linear-quadratic-Gaussian (LQG) controller, rate feedback stabilizer, linear quadratic regulator (LQR), single-area power system model, settling time.

I. INTRODUCTION

With the increasing demand of electricity it is becoming more and more difficult to keep the power system stable [1]. The electricity supplier has to provide quality power to the consumer. Every piece of electrical equipment has its own operating voltage range which should be maintained for effective operation. When the voltage level falls outside the operational range, a device may be unable to start its operation or it may malfunction. Furthermore, the device may be damaged. So, it is very essential to minimize the variations in voltage in order to protect the equipment using electricity. Load of a power generating unit is continuously varying and a change in reactive power affects the system voltage. The basic way of controlling the voltage of a generating unit is by using the automatic voltage regulator (AVR) loop [2], [3]. It is designed to automatically control, adjust and maintain a constant voltage level to avoid the costs associated with equipment damage and downtime caused by poor voltage levels. AVR controls the system voltage by controlling the generator's excitation system through a buck or a boost action. Any change in reactive power influences terminal voltage to be changed. Hence, any mismatch between generation and demand deviate the system voltage from scheduled values. As, this deviation is not desired an effective control is necessary to restore the terminal voltage. Traditionally, voltage regulation in power system is achieved by balancing generation and demand.

In the past literature, various investigations have been carried out on AVR with their mathematical model [4], [5]. The transient response characteristics of the basic AVR loop was not satisfactory. It shows approximately 83 % overshoot along with steady state error of 9 %. Moreover, the settling time, a crucial parameter for power system stability, is approximately 19 s which is quite high. For better response during load variation, an improved dynamics of AVR is demanded. In order to improve the system dynamics a rate feedback stabilizer was implemented in [5] which reduces the overshoot to around 4 %, steady-state voltage variation was 0.09 and minimizes settling time to 8 s. but it is not wise to use as the settling time is again large.

Proportional-integral-derivative (PID) controller based AVR control technique was introduced in [6]. The step response of PID controlled AVR loop gives smaller overshoot as compared to open-loop AVR, steady-state error was reduced to 0.001 with a lower settling time of about 4.5 s. From all the above literature it is clear that, the control system for AVR have to be more efficient hence optimal control techniques can be a better solution of the problem which can track the reference trajectory more effectively.

Optimal control is a branch of modern control theory that deals with the problem of finding a control law for a given system. The control problem includes a cost function which is a function of state and control variables of the concerned system. An optimal control is actually a set of differential equations that describes the paths of the control variables to minimize the cost function. Linear quadratic regulator (LQR), an optimal control technique, has been introduced in [7] and [8] to mitigate the voltage deviation due to load variation in power generating system. LQR helps to enhance the transient stability under different operating conditions. When a step response is obtained for the LQR controlled AVR loop it gives around 0.6 % of overshoot with approximately 0.9 s settling time.

All the above mentioned literatures provide a large time to be settled down for a power generating unit which motivated for further research to improve the transient response characteristics. A robust controller is needed to effectually improve the system dynamics. The main goal of

the controller should be to give better transient response to a sudden change of load. If a significant mismatch between load and generation is not properly coordinated it is possible for the system to become unstable. Furthermore, high voltage deviation may lead to system collapse. Encouraged by the importance of controlling AVR loop a robust control scheme based on linear-quadratic-Gaussian (LQG) controller augmented with an integral control action is proposed here which will help to improve the performance of the control loop. It will be proved as an accurate and fast acting optimal controller to maintain constant nominal terminal voltage accompanied with a short settling time.

The remainder of the paper is organized as follows: In Section II, the model and step response of AVR is developed for a single-area power system. The controller design is presented in Section III and implementation of these controllers on AVR system is provided in Section IV. Section V includes a comparison of the proposed controller with the existing control techniques and finally, section VI draws conclusion.

II. Modeling of AVR

An increase in the reactive power load of the generator goes along with a drop in the terminal voltage. The function of AVR is to regulate the voltage. It takes the fluctuate voltage and changes them into a constant one. It works on the principle of detection of errors. The output voltage of an AC generator obtained through a potential transformer and then it is rectified, filtered and compared with DC reference. The difference between the actual voltage and the reference voltage is known as the error voltage. The error signal is then amplified which increases the terminal voltage of the exciter. So the generator field current is increased.

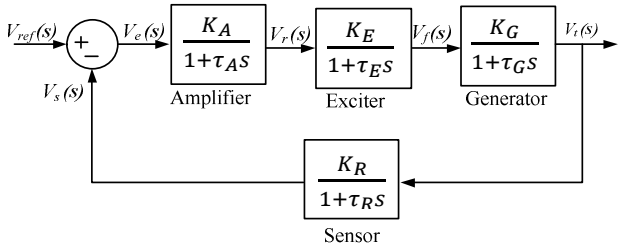


Fig. 1: AVR block diagram of an isolated power system.

TABLE I. PARAMETERS OF AN ISOLATED SYSTEM [5]

Elements	Gain	Time constant
Amplifier	K_A=10	τ_A = 0.1
Exciter	K_E = 1	τ_E = 0.4
Generator	K_G = 1	τ_G = 1.0
sensor	K_R = 1	τ_R = 0.05

As a result, the reactive power generation is increased accordingly and hence the terminal voltage is raised to the desired value. The block diagram of an AVR system is shown in Figure 1 [5]. The AVR model parameters used for an isolated system are presented in (Table I) [5].

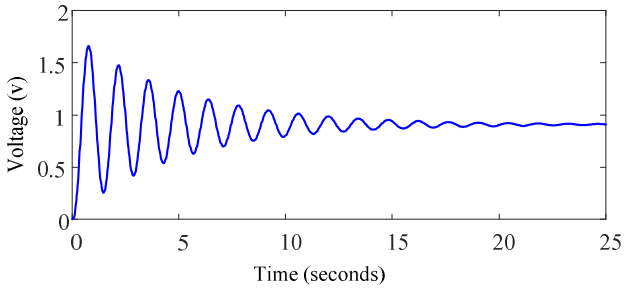


Fig. 2: Terminal voltage step response of AVR.

With the primary AVR loop, a change in the system load will result in a steady-state voltage deviation. Moreover, the response is highly oscillatory and it takes approximately 19 s to be settled down. Figure 2 shows the step response of AVR loop. So using alone the AVR loop to control the voltage is not wise as the response of the system is not satisfactory.

III. Controller Design

A. Design of Rate Feedback Stabilizer

In order to improve system dynamics a rate feedback stabilizer was introduced in [5]. It adds a zero to the AVR open loop transfer function which helps to remove the oscillation from the step response. Figure 3 shows the block diagram of the AVR system with rate feedback stabilizer.

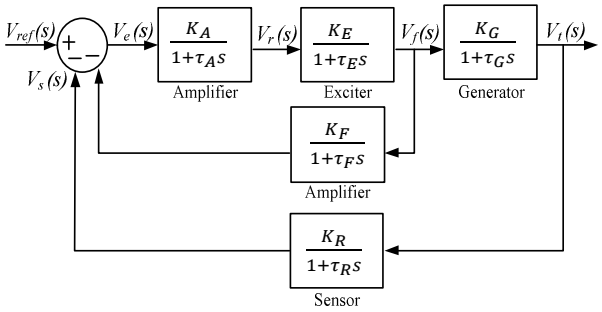


Fig. 3: Block diagram of an AVR system with rate feedback stabilizer.

B. Design of PID controller

PID controller is widely used in industrial control systems. It continuously calculates an error value as the difference between a desired set-point and a measured process variable then applies a correction to the plant based on proportional, integral and derivative action. The transfer function of the PID controller is

$$G_C(s) = K_P + \frac{K_I}{s} + K_D s \qquad (1)$$

A simple block diagram of a PID controlled system is showed in Figure 4.

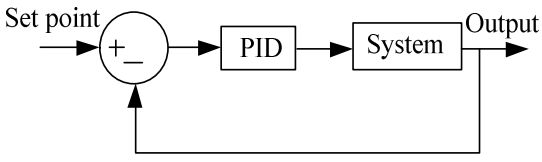


Fig. 4: Block diagram of PID controller.

C. Design of LQR

LQR controller is a form of optimal control that provides optimally controlled feedback gains to enable the closed-loop stable and high performance design of systems. It is a well-known design technique that provides practical feedback gains. In LQR, the plant dynamics are described by a set of linear differential equations and the cost is described by a quadratic function, so the name is linear quadratic regulator. If the plant described by

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t); \end{cases} \quad (2)$$

where, x is called the state vector, $x(t) \in \mathbb{R}^n$, y is called the output vector, $y(t) \in \mathbb{R}^q$, u is called the input (or control) vector, $u(t) \in \mathbb{R}^p$, A is the state (or system) matrix of $n \times n$, B is the input matrix of $n \times p$. C is the output matrix of $q \times n$, D is the feedthrough (or feedforward) matrix (in case where the system model does not have a direct feedthrough, D is the zero matrix), of $q \times p$. $\dot{x}(t)$ is the first derivative of x(t). The problem is to find the vector $K(t)$ of the control law

$$u(t) = -K(t)x(t) \quad (3)$$

which minimizes the value of a quadratic performance index

$$J = \int_0^\infty (x^T Qx + u^T Ru)dt. \quad (4)$$

Here Q is an $n \times n$ positive semidefinite matrix, also called the state weighting matrix. R is an $p \times p$ positive definite matrix known as control weighting matrix [7].

$$K = R^{-1}B^T P \quad (5)$$

here P is the unique positive definite solution of the algebraic Riccati equation (ARE) [7].

$$A^T P + PA + PBR^{-1}B^T P + Q = 0. \quad (6)$$

From the state space representation of the plant the value of A, B, C and D matrices can be found. The control scheme is

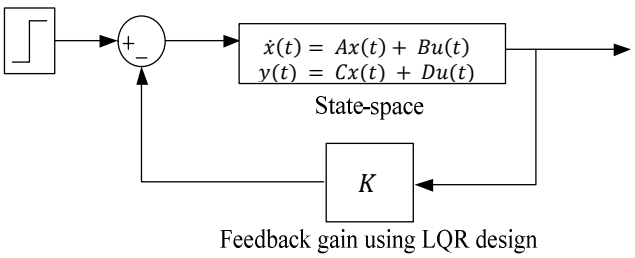


Fig. 5: Block diagram of an LQR controlled system.

based on placing the closed-loop poles in a suitable region in the complex plane to ensure desired transient response. By proper selection of the Q and R matrix the value of P and therefore the value of K can be obtained which will place the closed-loop poles of the system in the most perfect place using the LQR method with respect to the cost function. Figure 5 shows a simple block diagram of an LQR controlled system.

D. Design of LQG

LQG is one of the most fundamental optimal controllers that concerns linear systems driven by additive white noise. The name LQG arises from the use of a linear model, an integral quadratic cost function, and Gaussian white processes to model disturbance signals and noise. The problem is to determine an output feedback law that is optimal in the sense of minimizing the expected value of a quadratic cost criterion.

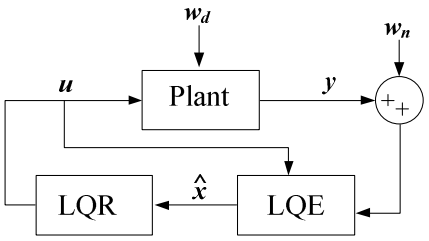


Fig. 6: Block diagram of an LQG controlled system.

Output measurements are assumed to be corrupted by Gaussian noise and the initial state is assumed to be Gaussian random vector. The LQG controller is simply a combination of a linear-quadratic estimator (LQE), together with a LQR. The separation principle states that the state estimator and the state feedback can be designed independently. Figure 6 shows a block diagram of LQG controlled system. For LQG the stste-space plant model is described as

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + w_d \\ y(t) = Cx(t) + Du(t) + w_n; \end{cases} \quad (7)$$

where the variables x, y, A, B, C and D represents the same as for LQR. w_d and w_n are additive white Gaussian system noise and Gaussian measurement noise respectively. Now the problem is to find the optimal control u(t) which minimizes the cost function as given in [9], [10].

$$J = \lim_{T \to \infty} E \int_0^T (x^T Qx + u^T Ru)dt. \quad (8)$$

Where E denotes the expected value, Q is the state weighting matrix, and R is the control weighting matrix.

IV. IMPLEMENTATION OF CONTROLLERS ON AVR

The step response of the basic AVR loop for an isolated power generating unit, whose parameters are given in TABLE I, is not satisfactory which is indicated by Figure 2.

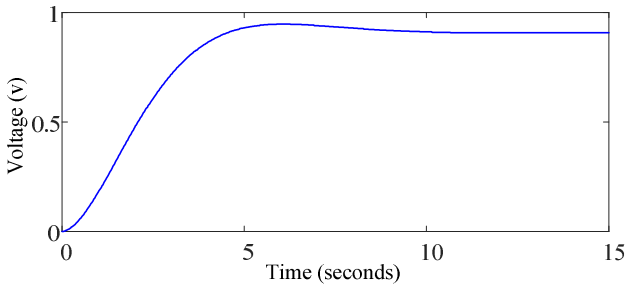


Fig. 7: Terminal voltage step response with rate feedback stabilizer.

Use of a rate feedback stabilizers can be a key to the problem. With proper adjustments of derivative gain, K_F and stabilizer time constant, τ_F better response as compared to conventional AVR loop can be obtained. When $\tau_F = 0.04$ s and $K_F = 2$, then the step response of the control loop is shown in Figure 7.

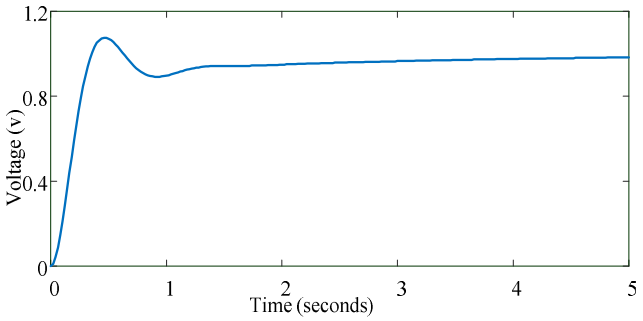


Fig. 8: Terminal voltage step response for PID controller.

The system gives about 4 % overshoot along with reduced steady-state error and a settling time of almost 8 s. When PID controller is added in the forward path of the AVR system then the step response of the combined system is given in Figure 8 which yields better transient response characteristics.

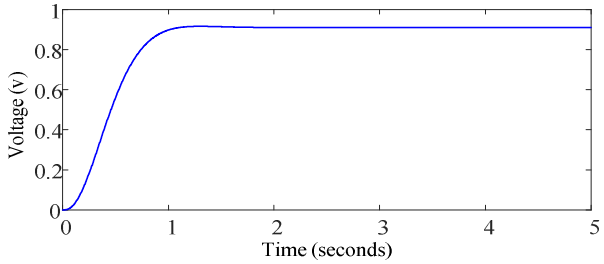


Fig. 9: Terminal voltage step response with LQR controller.

It shows about 4.5 s settling time with much improvement in the steady-state error but suffers from overshoot issue. Figure 9 shows the step response for the AVR system when LQR control technique is implemented. For the robust control of system dynamics and to improve transient response characteristics proposed control scheme based on an LQG controller augmented with an integral action

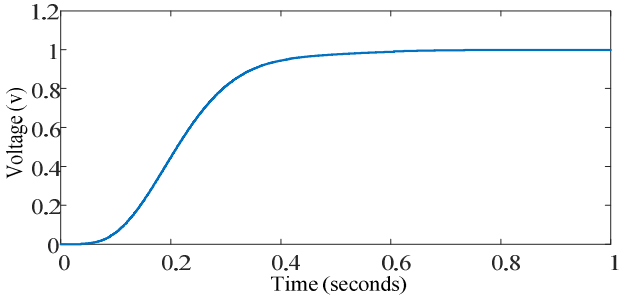


Fig. 10: Terminal voltage step response with proposed LQG controller.

is implemented on the same AVR model. Figure 10 shows the improved step response with around 0.5 s settling time accompanied with no overshoot issue and much reduced steady-state error.

V. COMPARISON AND RESULT

To investigate the robustness of the suggested LQG control technique for restoring the terminal voltage on load variation a comparison is demonstrated with existing controllers which is shown in Fig. 11. From figure 11 it can

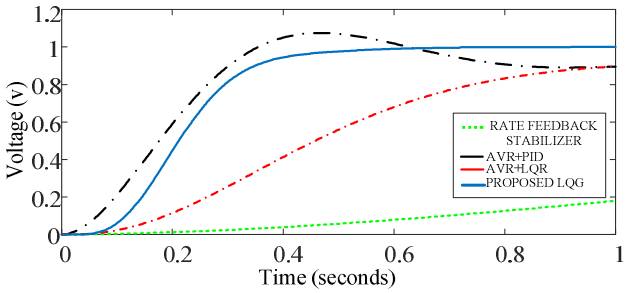


Fig. 11: Terminal voltage step response for different controllers.

be observed that after implementing proposed controller the closed loop system performance satisfies the requirement criteria more effectively. TABLE II illustrates the results for different control techniques implemented on AVR containing some important transient responses. With the proposed method, the steady-state error is relatively small with a zero overshoot and the settling time is approximately 0.5 s, which is a significant improvement.

TABLE II. Performance parameters of different control techniques

Controller	Settling Time (seconds)	Overshoot (%)	Steady-State Error (%)
Rate feedback stabilizer	8.1274	4.1812	0.0913
PID	4.5010	7.5763	0.0010
LQR	0.9697	0.6359	0.0910
Proposed LQG	0.5161	0	5.82e-04

VI. CONCLUSION

A broad analytical analysis of AVR control loop for regulating the terminal voltage is presented in this article. Different control techniques to control the basic AVR loop are inspected by taking their corresponding step responses of terminal voltage. The design and implementation of an optimal LQG controller augmented with an integral action for the robust control of AVR system dynamics is described in this article. Then some transient response characteristics of terminal voltage are investigated and compared with the existing controllers which justify the effectiveness of proposed control scheme. The control methodology is applicable for an isolated power generating unit but with some modifications it can be applied to the interconnected power grid as well, which may be considered as the future work to perform.

REFERENCES

[1] M. A. Rahman, M.S. Rana and M.S. Anower, "Indemnity for Frequency Disruption in a smart Grid during Cyber-Attack," in 2nd International Conference on Electrical & Electronic Engineering (ICEEE), Rajshahi, 2017, pp. 1-4.

[2] Jorge Calderon, "Stability Enhancement of an Industrial Power System by AVR Gain Readjustment", *IEEE Latin America Transactions,* vol. 15, no. 4, pp. 663-668, Apr. 2017.

[3] Rumi Rajbongshi, Lalit Chandra Saikia, "Combined Voltage and Frequency Control of a Multi-area Multisource System Incorporating Dish-Stirling Solar Thermal and HVDC Link", *IET Renewable Power Generation,* vol. 12, no. 3, pp. 323-334, Feb. 2018.

[4] Mauro Cimino, Prabhakar R. Pagilla, "Reactive power Control for Multiple Synchronous Generators Connected in Parallel" *IEEE Transactions on Power System,* vol. 31, no. 6, pp. 4371-4378, Nov. 2016.

[5] Kuntal Satpathi, Abhisek Ukil, Josep Pou and Michael Adam Zagrodnik, "Design, Analysis, and Comparison of Automatic Flux Regulator With Automatic Voltage Regulator-Based Generation System for DC Marine Vessels", *IEEE Transactions on Transportation Electrification,* vol. 4, no. 3, pp. 694-706, Apr. 2018.

[6] Hany M. Hasanien, "Design Optimization of PID Controller in Automatic Voltage Regulator System Using Taguchi Combined Genetic Algorithm Method", *IEEE Systems Journal,* vol. 7, no. 4, Nov. 2012.

[7] Abhinav Kumae Singh and Bikash C. Pal, "Decentralized Control of Oscillatory Dynamics in Power Systems Using an Extened LQR", *IEEE Transactions on Power System,* vol. 31, no. 3, pp. 1715-1728, May 2016.

[8] Xiangyu Wu and Chen Shen, "Distributed Optimal Control for Stability Enhancement of Microgrids With Multiple Distributed Generators", *IEEE Transactions on Power System,* vol. 32, no. 5, pp. 4045-4059, Sep. 2017.

[9] Mahendra Bhadu, Nilanjan Senroy, Indra Narayan Kar and Gayatri Nair Sudha, "Robust linear quadratic Gaussian-based discrete mode wide area pwer system damping controller", *IET Generation, Transmission & Distribution,* vol. 10, no. 6, pp. 1470-14278, Apr. 2016.

[10] Francisco Huerta, Jorge Perez, Santigo Cobreces and Mario Rizo, "Frequency-Adaptive Multiresonant LQG State-Fedback Current Controller for LCL-Filtered SCs Under Distorted Grid Voltages", *IEEE Transactions on Industrial Electronics,* vol. 65, no. 11, pp. 8433-8444, Nov. 2018.

Design & Implementation of a Microcontroller Based Automatic Power Factor Rectification System for Different Loads

Md. Mayen Uddin
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
mayeen.mail@gmail.com

Abdullah Al Mahmud
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
mahmudkhaled020845@gmail.com

Naeemul Islam
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
Naymnaeem08@yahoo.com

***Abstract*—In this work, we implement automatic power factor correction technique using a microcontroller to minimize penalty charges, cut the losses and save power. The chief objective is to improve the power quality by continuously monitoring the load power factor, when the load power factor falls below a specific value it results in the increase of line current, resulting in more line loss and greater voltage drop. Thus, the aim is to inject capacitances of required value when the power factor falls below the specified level. In this work, we use static capacitor method which will be controlled by PIC Microcontroller with very low cost although many existing systems are present which are expensive and difficult to manufacture. In this study, many small rating capacitors are connected in parallel. The suitable number of static capacitors is automatically connected according to the instruction of the Microcontroller to improve the power factor close to unity**

Keywords— microcontroller, detector, power factor, loads, correction.

I. INTRODUCTION

Electrical vitality portion is of fundamental significance to mechanical and business organizations working in the presently focused markets [1]. Ideal utilization of plant and hardware is one of the important thing that companies are attempting to harmonize with the vitality proportion for both financial and natural reasons. As society turns out to be continuously aware of its side effects on nature, limited vitality utilization turns out to be more ideal, which is a feasible objective for everybody. Using estimates, for example, readdressing of control factor, power utilization improvement, which at last direct to diminish of vitality utilization and diminished CO_2 greenhouse gas radiation [2]. Programmed control factor controller is needed to raise the control factor consequently at whatever point the power factor falls underneath an unequivocal level [3].

As the interest for electrical vitality is amplifying day by day, an ever increasing number of inductive burdens are being utilized in mechanical and local applications. Inductive loads are the primary purpose behind low power factor in the power system [4]. This way we have to build up a strategy to enhance control factor naturally. Programmed control controller can answer this issue. By rising force factor of intensity framework typically, control framework effectiveness can be improved [5]. In this work, microcontroller, relays, potential transformer, current transformer, and zero intersection circuit are utilized for creating power factor amendment model. The essential morals behind this programmed power factor enhancement are the power ideality in a power framework which has turned into a vital issue these days with a beautiful advancement of intensity gadgets innovation application and nonlinear load as of late prompts symphonious impedance in the power framework. Another inspiration is, in a request to guarantee most great condition for a supply framework from the building.

Furthermore, from financial point of view, it is imperative to have control factor as near solidarity as could be expected under the circumstances, that is the reason we require programmed control factor enhancement. In our nation, the more significant part of the heaps in the majority of the enterprises are inductive, which have slacking power factor. So on the off chance that we can enhance the power factor level closest to the 1, it will be savvy and productive. For instance, consider an asset with 1 KVA rating which is required to control apparatuses of 80 watts each [6]. A control factor of 0.61, the asset, best case scenario can oversee 8 machines every one of 80 watts.

If a similar power factor can by one way or another be enhanced to 0.91, the asset ought to be agreeable to control up to no less than 11 gadgets according to the dealings [6]. The inner voice behind this work which began from one of the fundamental issues looked by the nearby private area. So by enhancing power factor can prompt abatement in KVA request and thus the client can perform moderate charging. Along these lines, in light of the above contentions, it tends to be expressed that enhancement in power factor enhances framework productivity [6]. Chaitanya N. Jibhakate, Madhuri A. Chaudhari, Mohan M. Renge proposed a work where Power factor enhancement is accomplished by controlling DC interface capacitor voltage and PID controller is utilized to control capacitor voltage, however the work isn't ideal for various strategies aside from controlling capacitor-voltage [7].

II. SYSTEM DIAGRAM

A. Block DIagram

Here Fig. 1 stands for the block diagram representation of the total operating system which depicts its working process and how the proposed work manages to venture down the voltage [8]. The equation of electromagnetic enlistment is being used by the transformer to drive electrical vitality starting with one circuit then onto the next loop or inside a similar circle. Time shifting attractive fields in a transmitter

deliver compel for electromagnetic acceptance and Transformers can be utilized to venture up a voltage or venture down energy relying on its arrangement for the utilization in various electrical circuits[9]. The microcontroller at that point uses its competency and program to deduct the stage edge and stage fluctuation between two waveforms. The pins of the microcontroller are associated with the hand-off IC and afterward to transfers.

At the point when a low voltage circuit is utilized to drive a heap which is associated with high voltage an IC is used which is known as Relay driver IC. A hand-off loop requires more current as break down to different gadgets. Incorporated circuits, for example, Op-amp circuits can't supply such expansive flows to driver delays. Transfers are individual under lock and key and have replaced sturdy state switches and are more grounded than sturdy state gadgets. Transfers have high current limits. To drive circuit disconnection and to spare from ESD, transmissions can be utilized

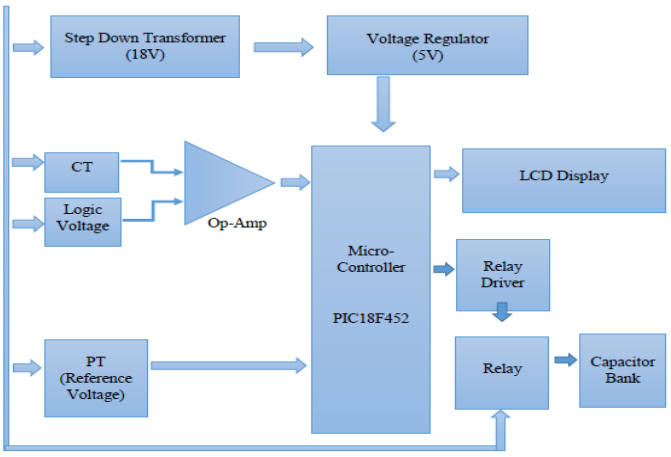


Fig. 1. Block diagram of the entire system.

III. Circuit Diagram

A. *Circuit Diagram for Auto PFC*

Fig. 2 demonstrates the circuit outline for Auto PFC(Power Factor Correction). In circuit graph of programmed control factor controller, the present transformer is used to get present waveform of load current, and current transformer likewise ventures down air conditioning current. The operational intensifier is utilized as a comparator in this circuit. Also, voltage transformer is utilized to get the present waveform and sustained this wave to Op-amp comparator. LM358 has been used as zero intersection indicators in this work after that both current and voltage waveforms are nourished to PIC18F452 microcontroller [10].

PIC18F452 microcontroller calipers zero intersection location and power factor by estimating time contrast among current and voltage wave frame. Microcontroller incorporates control factor and takes required activities dependent on power factor. Transfer driver is associated with a microcontroller which is issued to drive transfers. Microcontroller occupies large flag to hand-off driver IC at whatever point control factors fall under 0.9. Hand-off driver turns on transfers which consequently interfaces capacitor and keeps money with the heap. As a matter of first importance, the stage distinction among voltage and current wave shape are estimated, and after that, the power factor is figured. If there should be an occurrence of low power factor, capacitors are added to enhance it.

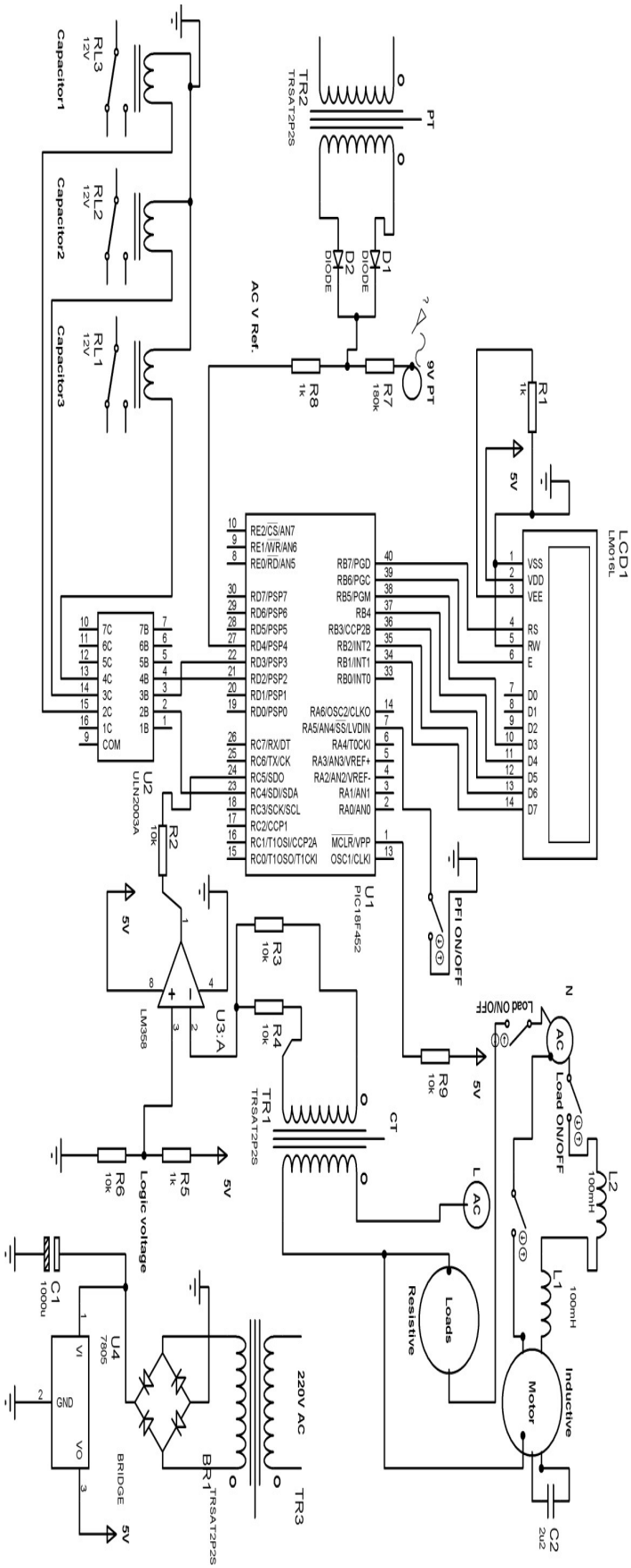


Fig. 2. Circuit diagram for auto PFC.

B. Hardware Implementation

The PCB was fabricated according to the diagram for the complete circuit and all the components are soldered on it. The connection for the potential transformer [11] and current transformer [11], capacitor bank was done at the respective places [12]. Continuity check using the multimeter was carried out to avoid any accidental short-circuits.

1) Prototype of Realized PF Correction Equipment:

In Fig.3 shows the prototype of realized power factor correction equipment where different types of components are set up on the PCB Board.

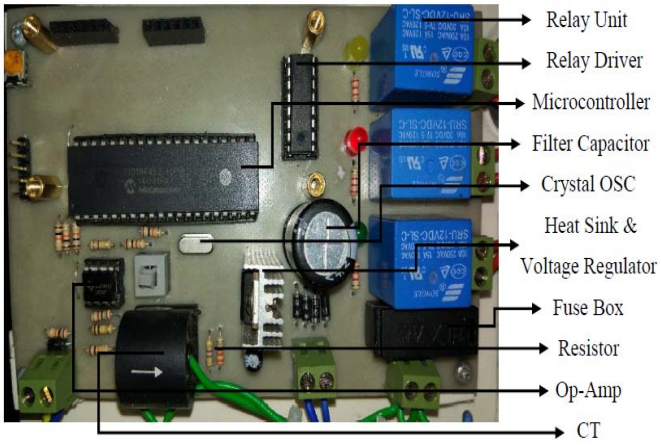


Fig. 3. The realized PF correction equipment.

2) The Complete Circuit Implementation: Fig. 4 shows the complete hardware model of APFC (Automatic Power Factor Correction).

C. Algorithm of PFC

- Step -1: Take input for voltage and current in the circuit.
- Step -2: Measure the phase lag and calculate the power factor.
- Step -3: Differentiate power factor from the targeted power factor.
- Step -4: Switch ON or OFF appropriate number of capacitors from capacitor bank.
- Step -5: Again compare the power factor with targeted PF and continue from step-1.

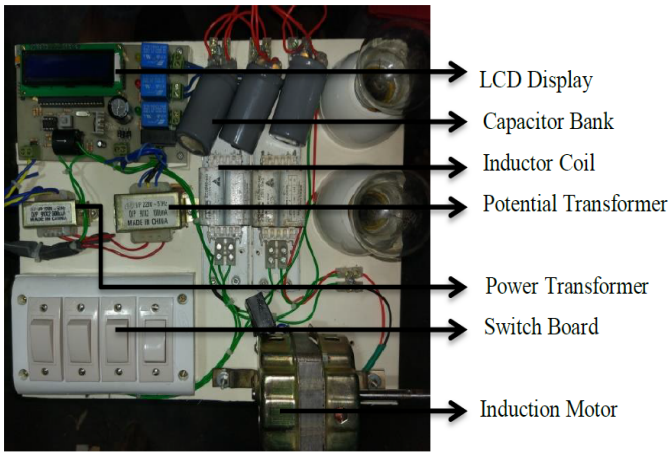


Fig. 4. The complete hardware model of APFC.

3) Flow Chart of PFC:

Fig. 5 represents the entire flow chart for power factor correction. In the flowchart, first, we start our whole system with power supply. The system will get initialized and read out the input voltage and current. After that, the operational amplifier will convert the sine wave to square wave for PFC angle translation. The microcontroller measures the power factor and displays it. If the power factor is one than Power Factor 1 is displayed on the LCD otherwise it will add capacitor bank by relay connection to remove the magnetizing current, and increasing power factor thus it will display power factor one on LCD, then the system will be stopped.

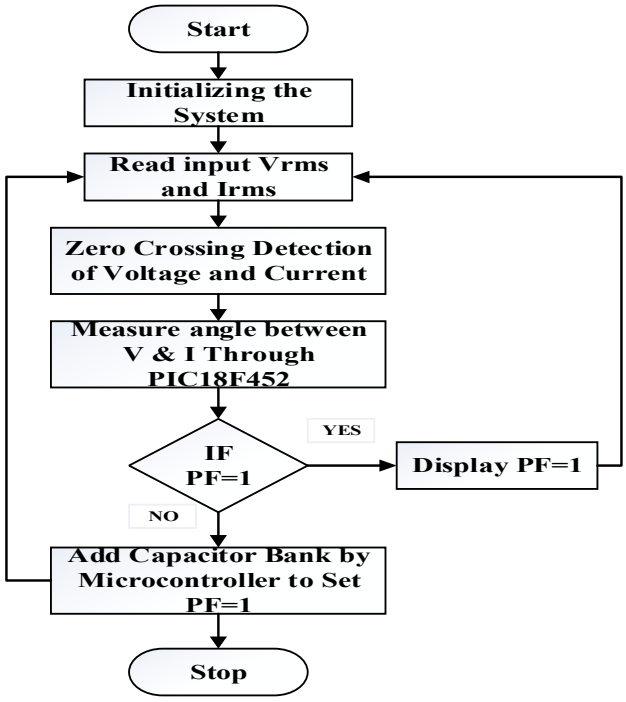


Fig. 5. Flow chart of PFC.

IV. Results and Calculations

A. Measurement of Power Factor

All the measurements need to be done for the three different types of load designed, separately with and without the correction equipment. The setup is shown below Fig. 6.

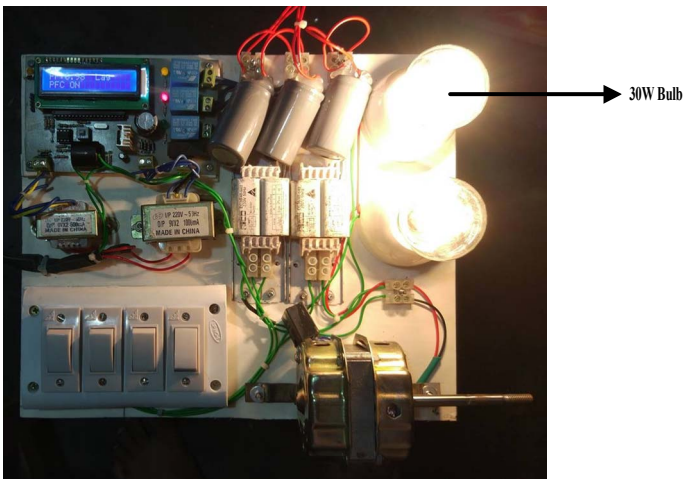


Fig. 6. Setup for measuring power factor.

B. Wave shape for Different Loads

1) Wave shape for Resistive Loads: When the resistive load is ON, there is no phase detain between current, and voltage signals and they are in phase. In this case, the power factor would be 0.9 as documented value, so there is no portion of capacitors as shown in Fig.7. In the fact of resistive load, the V and I are in phase, so there is no part of capacitors to improve power factor as shown in Fig.8.

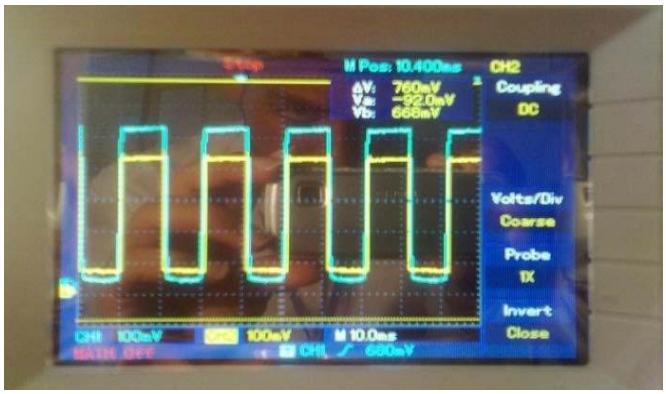


Fig. 7. Zero crossing detection for resistive load.

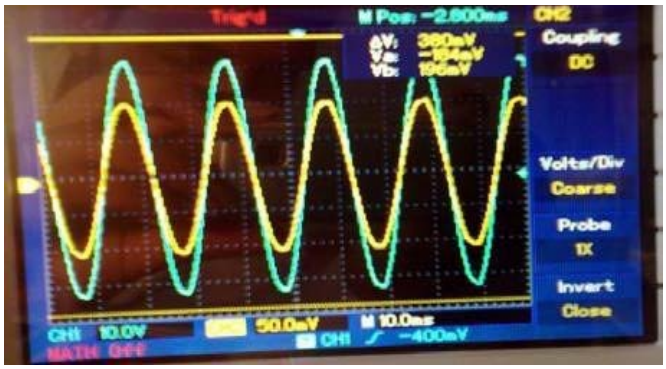


Fig. 8. V and I behavior for resistive load.

2) Wave Shape for Inductive Loads before Correction:

At the point when an inductive heap of an enlistment engine is ON, there is a stage confine among voltage and current signs. Microcontroller detects the confine delivered by the load, and as indicated by the keep, it enters the coveted estimation of the capacitor to enhance the power factor of the framework as appeared in Fig. 9. The zero-intersection identification of V and I motions for an acceptance engine has appeared in Fig. 10.

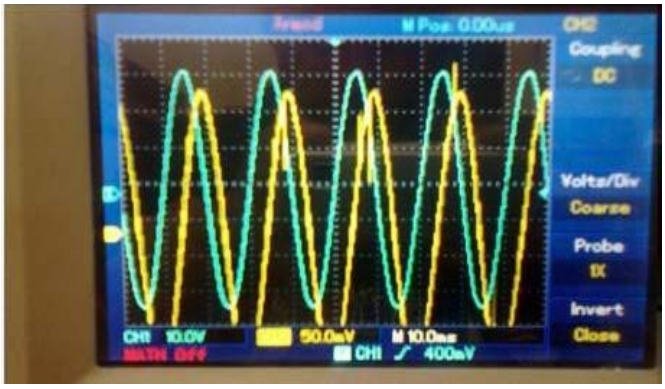


Fig. 9. V and I behavior before correction with inductive load.

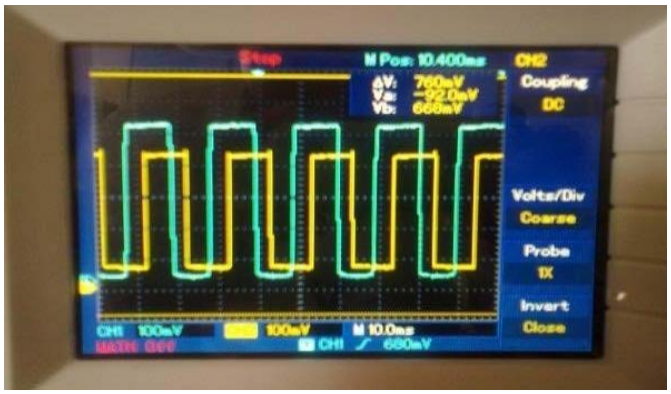


Fig. 10. Zero crossing detection before correction with inductive load.

3) Wave Shape for Inductive Loads after Correction: According to the phase delay in signals, microcontroller takes the intelligent decision and adds the desired value of the capacitor. When the desired amount of the capacitors added the required reactive power to the system, the current and voltage waveforms are in phase, as shown in Fig 11. After that, the insertion of the required value of the capacitor, the V and I zero cross detector signals are also in phase by the set referenced value of power factor (0.9) as shown in Fig. 12

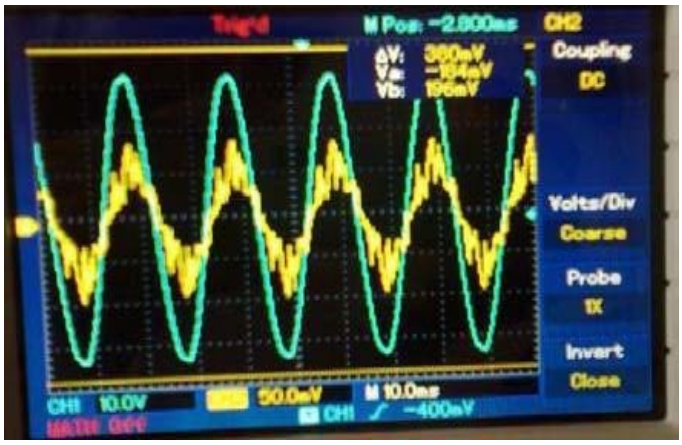


Fig. 11. V and I behavior after correction with inductive load.



Fig. 12. Zero crossing detection after correction with inductive load.

C. Output Voltage across Current Transformer

1) For No Load: When there is no load, the output voltage across the current transformer is 0.018V as shown below Fig. 13.



Fig. 13. Output voltage across current transformer (no load).

2) For Different Load: When there is Inductive, Non-Inductive and Resistive Load, the output voltage across the current transformer is 0.021V, 0.015V and 0.015V respectively as shown below Table 1. Data in Table 1 are shown in Fig.14, Fig.15, and Fig.16.

TABLE I

MEASUREMENTS OF CURRENT TRANSFORMER FOR DIFFERENT LOADS

No.	Load Type	Output Voltage (V)
1.	Resistive	0.021
2.	Inductive	0.015
3.	Capacitive	0.015



Fig. 14. Measurement of current transformer for resistive load.



Fig. 15. Measurement of current transformer for inductive load.



Fig. 16. Measurement of current transformer for capacitive load.

D. Analysis of Load before Correction: The analysis of the pure restive load (R Load), Inductive load (L) and Non Inductive load (C) was done without using the correction equipment. The designed loads were connected and the readings for different electrical parameters were recorded.

TABLE II

ANALYSIS OF LOAD BEFORE CORRECTION

No.	Load Type	Supply Voltage (V)	Load Current (A)	Power Drawn (W)	Power Factor
1.	Resistive	212	1.22	200	0.94
2.	Inductive	217	0.54	40	0.34
3.	Capacitive	219	0.36	30	0.38

E. Analysis of Load after Correction: As there is a demand for power factor improvement for load, they were connected to the supply along with the correction equipment designed. All the three loads designed were applied and the observed values were recorded.

TABLE III

ANALYSIS OF LOAD AFTER CORRECTION

No.	Load Type	Supply Voltage (V)	Load Current (A)	Power Drawn (W)	Power Factor
1.	Resistive	215	0.672	200	0.98
2.	Inductive	217	0.177	40	0.95
3.	Capacitive	219	0.142	30	0.68

TABLE IV

IMPROVEMENT OF POWER FACTOR

No.	Load Type	Percentage of improvement
1.	Resistive	4%
2.	Inductive	61%
3.	Capacitive	30%

F. Graphical Representation of Auto Power Factor Correction of Different Loads:

1) Resistive Load: Fig. 17 shows the graphical representation of auto power factor correction of the resistive load.

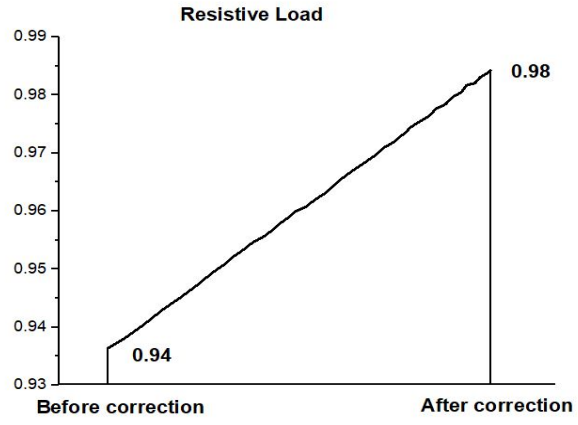


Fig. 17. Graphical representation of auto power factor correction for resistive load.

2) Inductive Load: Fig. 18 shows the graphical representation of auto power factor correction of the inductive load.

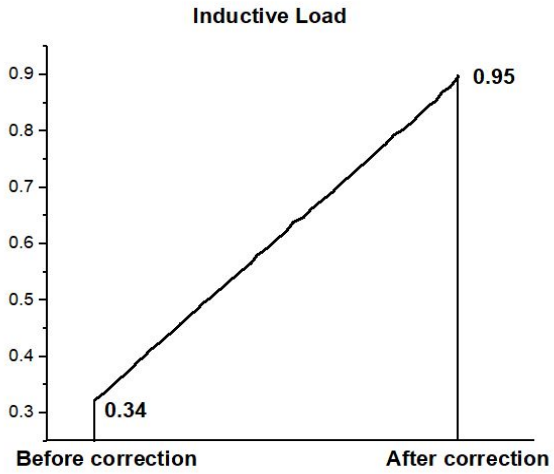


Fig. 18. Graphical representation of auto power factor correction for inductive load.

3) Capacitive Load: Fig. 19 shows the graphical representation of auto power factor correction of the capacitive load.

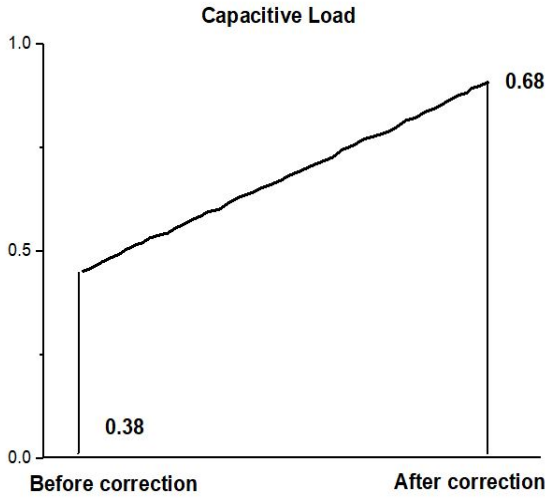


Fig. 19. Graphical representation of auto power factor correction for capacitive load.

It should be noted that when the value of power factor improves, the losses made because of low power factor also gets reduced.

V. Conclusion

It can be concluded that the power factor correction approach can be applied to the industries, power systems and also households. Thus an improved power factor reduces reactive power, copper loss, transmission loss and avoids poor voltage regulation, overloading also improve voltage control and efficiency of the supply system. However, this method comes with some limitations such as the compensation lower ratings, they have a short service life range, easily damaged if the voltages exceed the rated value. In future, PWM techniques can be employed in this scheme along with power factor correction .As it operates automatically, human resources are not required and the use of automated power factor rectification capacitor banks can help attain more efficient way of running industries in the future.

References

[1] G.J. Stigler, C. Friedland. “What can regulators regulate? The case of electricity.” The Journal of Law and Economics 5 (1962): 1-16.

[2] S. Billatos, Green technology and design for the environment, CRC Press, 1997.

[3] G.Escobar, D. Chevreau, R.Ortega, and E. Mendes. "An adaptive passivity-based controller for a unity power factor rectifier." IEEE Transactions on Control Systems Technology 9, no. 4 (2001): 637-644.

[4] H. Akagi. “New trends in active filters for power conditioning.” IEEE transactions on industry applications 32, no. 6 (1996): 1312-1322.

[5] B.Wollenberg, and A.Wood. “Power generation, operation and control.” John Wiley&Sons, Inc (1996): 264-327.

[6] M.B.Khan, and M. Owais. “Automatic power factor correction unit.” In Computing, Electronic and Electrical Engineering (ICE Cube), 2016 International Conference on, pp. 283-288. IEEE, 2016.

[7] C.N.Jibhakate, M. A. Chaudhari, and M. M. Renge. “Power factor improvement using nine switch AC-DC-AC converter.” In Power Systems (ICPS), 2016 IEEE 6th International Conference on, pp. 1-4. IEEE, 2016.

[8] A.Hoque, M. B. H. Shorif, S. Nuruzzaman, and M. E. Alam. “Arduino based battlefield assistive robot.” In Humanitarian Technology Conference (R10-HTC), 2017 IEEE Region 10, pp. 304-309. IEEE, 2017.

[9] W. M. A. B. W. Sabri, I. R. Zabidi, M. A. M. A. Musthalib, Z. A. Muhamad, M. I. M. Hazman, M. S. M. Yusof, L. Naing, and Z. Z. Abidin. “Low-cost remotely operated underwater vehicle for underwater observation purposes.” In Underwater System Technology: Theory and Applications (USYS), 2017 IEEE 7th International Conference on, pp. 1-5. IEEE, 2017.

[10] M.S.Islam,“|Home security system based on pic18f452 microcontroller.” In Electro/Information Technology (EiT), 2014 IEEE International Conference on, pp. 202-205. IEEE, 2014.

[11] S. S. Routray, “Microcontroller Based Automatic Power Factor Correction in Mines,” National Institute of Technology, 2015.

[12] I.Filipiak, and M. Wierzbowski. “Capacity market in Poland—Evalutaion of the proposed solution.” In European Energy Market (EEM), 2017 14th International Conference on the, pp. 1-5. IEEE,2017.

IMPROVED LOAD FREQUENCY CONTROL OF TWO-AREA POWER SYSTEM

Abdullah Al Mamun[1], Mohammad Mostafizur Rahman[1], Anowar Hosen[1], Sanjita Akter[1] and M.M. Rahman[1*]

[1]*Department of Electrical and Electronic Engineering,*
Dhaka University of Engineering and Technology,
Gagipur-1707
*mijan.duet@yahoo.com

Abstract-The power lines are connected in an interconnected power system through tie lines and Load Frequency Control (LFC) places the frequency and tie line power errors zero to keep the system in steady state condition. This paper represents an effective decentralized Load Frequency Control (LFC) by means of the concept of Active Disturbance Rejection Control (ADRC). The performance of ADRC based LFC is compared with PID based LFC in two area power system. Simulation is done by MATLAB Simulink and the results demonstrated the better robustness and dynamic performances with ADRC based LFC over PID based LFC.

Keywords—ADRC, LFC, ESO and two-area power system

I. INTRODUCTION

The main problems of modern interconnected power system are sustaining a normal working position and optimum scheduling of generation [1]. To keep the system steady, active- and reactive-power are the two operating parameters [2]. Moreover, these two equilibrium point are controlled as economically and reliably as possible to meet the predetermined value of frequency and voltage in an interconnected system [3]. Changing the frequency of the power system is primarily influenced by the uncertainty of the system and dynamic external disturbances. System frequency and voltage are directly correlated with change in real- and reactive-power respectively.

The LFC loops manage the system active and frequency which is termed as Automatic Generation Control (AGC) [4]. The AGC divides the loads among the generating units and correctly control the tie line power. In order to abate Area Control Error (ACE) to zero in the presence of system uncertainty and disturbances, it is need to introduce LFC in the interconnected power system [5]. For improving the performance and stability of the power system there are some conventional controllers proposed earlier[6].The traditional PI, PID and others controllers are robust and prolific in disturbance but deteriorate performance by the nonlinearities and untrustworthy load conditions [7].

Now, the promising and a new solution is Active Disturbance Rejection Control (ADRC), which is easily enumerated and mollifies uncertainties of the system and internal-external disturbances [8]. Estimation and cancellation of disturbances can be done by using an External State Observer (ESO) [9]. In this controller hardware-software are easily comprised and easily tuned by the factory personnel in the industry.

In comparison with traditional controllers, ADRC illustrated better dynamic performances i.e. lower settling time, overshoot and oscillations [10]. LFC of multi-area power system is recommended using ADRC based on second-order sliding mode control. Here, the load variations and tie-line power exchange deviations are considered as a lumped disturbance [11]. The ADRC controller is broadly used in chemical industry, electricity, aerospace, aviation and others field with many advantages [12]. In this progression, ADRC is applied LFC of two area power system and the performance of this controller is compared with PID based LFC. To the best of my knowledge, improved LFC of two-area power system using ADRC based LFC is not reported yet.

II. ACTIVE DISTURBANCE REJECTION CONTROL (ADRC)

The new control technology, ADRC, estimate the total internal uncertainties and external disturbances using an observer called Extended Estate Observer (ESO) and actively cancel the disturbance with the nonlinear feedback controller. The plant reformatting and ADRC structure are given in the following section. The designed parameters of this controller are considered form [13].

A. Plant Remodeling

The ADRC is designed using a common transfer function model of a physical system with primary loop. Both time-domain and frequency–domain representations is considered. The foremost endeavor

is to develop ADRC for a higher order system. Hence, an nth order system is introduced for the design of ADRC. By utilizing a minimum phase transfer function, a plant with finite zeros can be developed which is as follows:

$$\frac{Y(s)}{R(s)} = G_p(s) = \frac{b_m s^m + b_{m-1}\ s^{m-1} + \cdots + b_1\ s + b_0}{a_n\ s^n + a_{n-1}\ s^{n-1} + \cdots + a_1\ s + a_0} \quad (1)$$

Where, $Y(s)$ and $R(s)$ are input and output of the plant respectively. a_i and b_j *(i=1,..., n, j=1,..., m)* are the coefficient of the transfer function.

After performing the longhand division, the plant model may be shown as

$$s^{n-m}Y(s) = b_o U(s) + D(s) \quad (2)$$

with $b_o = \frac{b_{m+1}}{a_{n+1}}$

Where, *D(s)* includes both the internal and external disturbances.

A. Estimation of Generalized Disturbance

It is very important to cancel the generalized disturbance *d (t)* when the plant is remodeled. An extra state of the system is considered for handling the discrepancies called Extended State Observer (ESO).The state space model of plant represented as follows:

$$sX(s) = AX(s) + BU(s) + E(s)D(s) \quad (3)$$

$$Y(s) = CX(s) \quad (4)$$

Where

$$X(s) = \begin{bmatrix} X_1(s) \\ X_2(s) \\ \vdots \\ X_{n-m}(s) \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & \ddots & 1 & \ddots & \vdots \\ \vdots & \ddots & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & 0 & \cdots & \dots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ b \\ 0 \end{bmatrix}, E = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

$$C = [1 \quad 0 \quad \dots \quad 0]$$

in order to eliminate the error a state space model of the disturbed process can be represented as follows.

$$sZ(s) = AZ(s) + BU(s) + L\,(Y(s) - \hat{Y}(s) \quad (5)$$

$$\hat{Y}(s) = CZ(s) \quad (6)$$

$$Z(s) = [z_1(s) \quad z_2(s) \dots \quad z_{n-m}(s)] \quad (7)$$

where, $L = [\beta_1 \quad \beta_2 \dots \quad \beta_{n-m}]$

In order to trace all eigenvalues of the ESO to $-\omega_0$ observer bandwidth), the observer gains are chosen as

$$\beta_i = \binom{n-m}{i}.\ \omega_0^i, \quad i = 1,\dots,n\text{-}m \quad (8)$$

By proper designing of ESO, $Z_i(s)$ will be able to estimate the values of $X_i(s)$ closely where i= (1, 2, 3……n-m). Then,

$$Z_{n-m}(s) = \hat{D}(s) \approx D(s) \quad (9)$$

The generalized disturbance *d (t)* can be removed by the time domain estimated value x_{n+1}

$$U(s) = \frac{U_o(s) - Z_{n-m}(s)}{b_0} \quad (10)$$

Now the system is reduced to a pure integral plant by substituting

$$S^{n-m}Y(s) = b_0 \frac{U_o(s) - Z_{n-m}(s)}{b_0} + D(s) \quad (11)$$

The control law for the pure integral plants

$$U_0(s)\big(R(s) - Z_1(s)\big)K_1 Z_2(s) \dots \dots K_{n-m-1} Z_{n-m-1}(s) \quad (12)$$

To simplify the tuning process, all the closed loop poles of the PD controller are set to $-\omega_c$.Then the controller gain in have to be selected as

$$K_i = \binom{n-m}{n-m-i+1}.\ \omega_c^{\,n-m-i+1}, \quad i = 1,2,\dots,$$

$$n - m - 1 \quad (13)$$

Placing all the observer poles at one location is known as bandwidth parameterization ω_c represents the bandwidth of the controller. Increasing ω_c the tracking speed of the output of ADRC controlled system will increase. In other words the tracking error, overshoot and settling time of the output will decrease. Generally, ω_c varies from 3~10 rad/s.

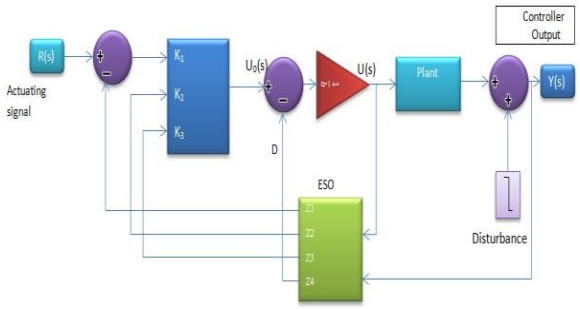


Fig. 1 ADRC structure with state feedback

The proposed ADRC control for decentralized system is shown in Fig.1. In this figure an ADRC is placed in each area acting as a local LFC under decentralized control strategy. These two decentralized areas are connected through tie line. For this LFC problem, ACE-1 and ACE-2 are the reference inputs and load changes ΔP_{L1} and ΔP_{L2} are the external disturbances in areas-1 and area-2 respectively. The parameters for both areas of the system are shown in Table 1.

TABLE-I PHYSICAL PARAMETERS OF GENERATING SYSTEM

Parameters	***Symbols(Units)***	***Value***
Inertia constant	H(p.u.sec)	10.00
Frequency sensitive load coefficient	D(p.u/Hz)	1.00
Turbine time constant	$T_i\,(second)$	0.30
Governor time constant	$T_g\,(second)$	0.10
Governor speed regulation	R(Hz/p.u)	0.05
Synchronizing coefficient	T(p.u/rad)	22.00
ADRC Bandwidth	$\omega_c\,(rad/sec)$	10.00
Observer Bandwidth	$\omega_o\,(rad/sec)$	20.00
Load change in area-1	$DP_{L1}\,(p.u)$	0.10
Load change in area-2	$DP_{L2}\,(p.u)$	0.08

The ESO is derived for observer bandwidth ωo = 20 rad/s. Next controller gains are computed for controller bandwidth ωc = 10 rad/s, as K1 = 1000, K2 = 300, K3 = 30.

III. PROPORTIONAL INTEGRAL DERIVATIVE (PID) CONTROL

In a closed loop system, the exact permutation is done among proportional, integral and derivative actions. This controller has universal applicability regardless of their low order. The conventional PID controller is represented by block diagram in Fig. 2. The tuning constraints of the controller are given in Annexure. A simplified implementation of the basic principle in error-based feedback control can be found in a classical PID controller. The control law is driven by the dissimilarities between the set point *v* and plant output *y,i.e.* $e = v - y$, as well as its differentiation de/dt and integration $\int e\,dt$ in Equ. 14.

$$u = k_0 \int_0^t e\ dt + k_1 e + k_2 \frac{d\,e}{d\,t} \qquad (14)$$

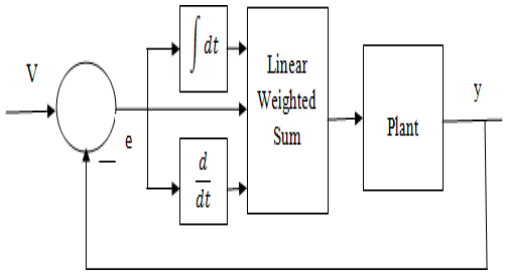


Fig. 2 Architecture of PID controller

TABLE-2 PARAMETERS OF PID CONTROLLER

Symbol	Name of the Parameter	Value of the Parameter
P	Proportional Gain	0
I	Integral Gain	-0.0298920
D	Derivative Gain	0

IV. PERFOMANCE COMPARISON OF ADRC AND PID BASED LFC FOR TWO AREA POWER SYSTEM

The two area power system consists of two different decentralized areas, which are connected to each other through tie-lines. Each area has three major components, which are turbine, governor and generator. The ADRC controlled interconnected power system is shown in Fig.3.The parameters of the system are obtained from [5, 6], and listed in TABLE 1.

For each decentralized area in Fig. 3, an ADRC is designed based on the transfer function $G_p\,(S)$.

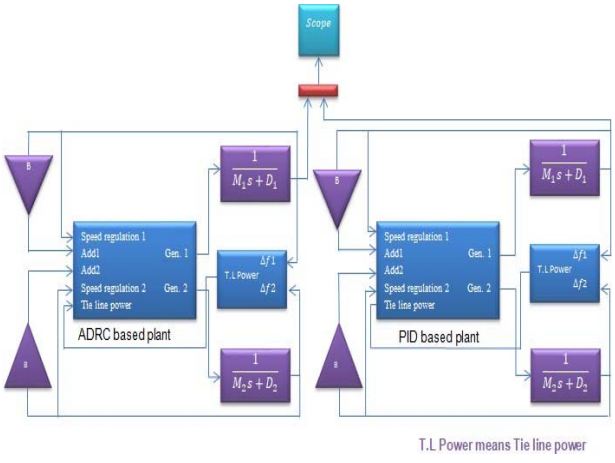
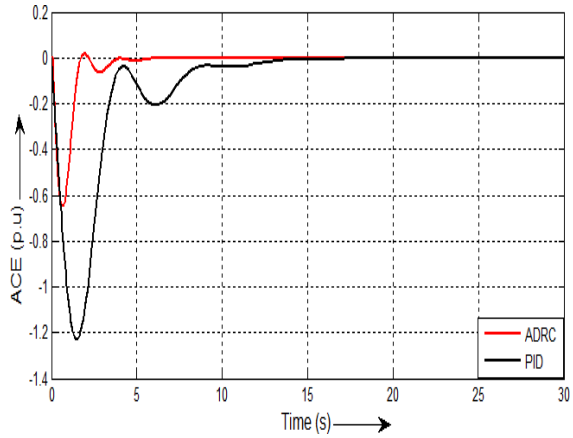


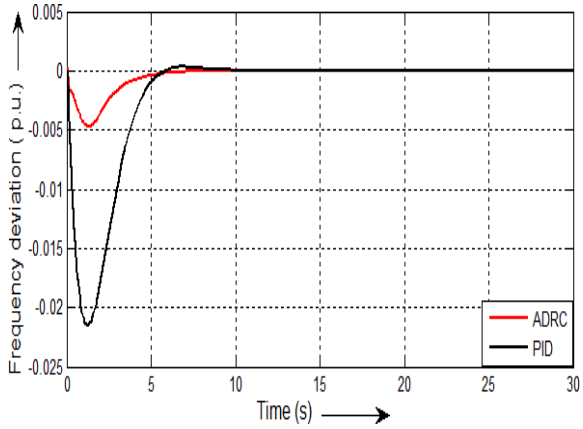
Fig. 3 ADRC and PID based LFC model of two-area interconnected power system

V. SIMULATION RESULTS

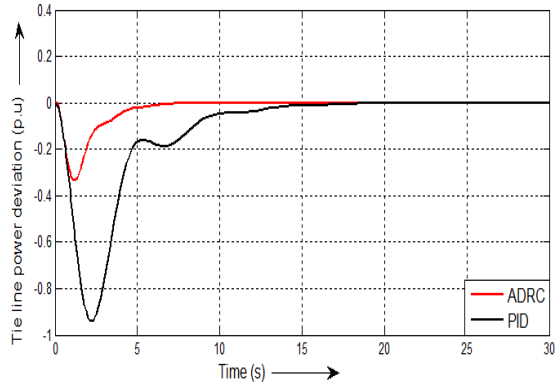
The effectiveness of ADRC is tested in this case by simulating the closed-loop control system in Fig.4.



(a)



(b)



(c)

Fig.4: Comparison of performance of PID and ADRC based LFC (a) ACE, (b) frequency error, (c) tie-line error

In order to verify the effectiveness of ADRC, a series of simulations are conducted comparing with PID controller in two-area interconnected power system. The system responses for two different areas are shown in Fig. 4. In case of LFC, area control error, frequency error and tie-line error are considered as output.

Comparison of ACE responses between ADRC and PID controller in both areas (Fig. 4) shows that, For ACE: ADRC has an over shoot, %OS = 0.63 and settling time, T_S = 5.10 sec. whereas the PID controller has an over shoot of %OS = 1.22 and settling time, T_S = 13.00 s. ADRC has lower settling time and less over shoot, hence the ADRC shows better performance than PID controller for ACEs in two area interconnected power system.

Similarly, the comparison of frequency deviation and tie line power deviation responses between ADRC and PID based LFC in both areas (Fig. 4). Better performance is illustrated in ADRC based LFC than PID based LFC.

All the responses of ADRC based LFC and PID based LFC are tabulated for better comparison in Table. III.

TABLE-III PERFORMANCE COMPARISON OF ADRC AND PID CONTROLLER BASED DOUBLE AREA POWER SYSTEM

SL. No.	***Controllers***	***Performances***	***PID Controller***	***ADRC Controller***
1.	Area control error (ACE)	Overshoot (p.u.)	1.22	0.63
		Settling time (Sec.)	13.00	5.10
2.	Frequency deviation error (Δf)	Overshoot (p.u.)	0.023	0.005
		Settling time (Sec.)	8.00	7.50
3.	Tieline power deviation (ΔPtie)	Overshoot (p.u.)	0.94	0.35
		Settling time (Sec.)	13	6.00

a. Dynamic Performance Analysis

To test the dynamic performance, a 0.01 p.u. step load change is applied in area-1. Observer bandwidth and controller bandwidth are considered 20 and10 respectively for ADRC. On the other hand, for PID controller, K_P = 0.5, K_i = 0.83 and K_d = 0.5. The simulation results are shown in Fig.4. From the simulation results, it can be demonstrated that the ACEs, the frequency errors and the tie-line power deviations have been driven to zero by ADRC in the presence of load power changes. Similarly, it is observed that due to the disturbance compensation by ADRC controller, the overshoot of frequency deviation in both areas and tie-line power deviations

are greatly reduced comparing with PID controller. Therefore, it can be vindicated that the ADRC controller provides a better dynamic performance.

b. Robustness Analysis

In order to test the robustness of ADRC, the variations of all of the parameters (*M, D, T,* T_i *,* T_g *, R*, and *B*) of the single unit can be changed. However, the controller constraints of ADRC are not changed with the variations of the system parameters. From the simulation results, it is observed that even though such large parameter variations, the system responses do not show remarkable differences from the results. Therefore, the simulation results demonstrated the toughness of ADRC against system parameter variations. If the system parameters are changed for reheat unit or non-reheat unit, the termination will be same since the model for each area is similar to the other.

The simulation results established the effectiveness of ADRC on the LFC problem in various aspects. First of all, ADRC is capable to cancel the ACE to satisfy the LFC requirements of maintaining the standard frequency and keeping the tie line power exchange according to the schedule. Secondly, ADRC is efficient to terminate both random and large load disturbances. Thirdly, ADRC can be used to resist the interferences to the controller intend of parameter uncertainties. Thus, ADRC is considered to be a suitable control technique for LFC solution.

VI. CONCLUSION

In this works, an improved LFC is proposed using ADRC for two-area interconnected power systems. Two area power systems is modeled and developed to examine the steadiness, steadfastness and toughness of ADRC controlled system in the presences of external disturbance, system constraints dissimilarities. The simulation results proved the usefulness of ADRC through successfully regulating the ACEs, frequency errors and tie-line power errors in the presences of external disturbance, parameter uncertainties. Eventually, it can be said that this scheme is much capable of handling anomalous stipulation in a power system produced by load variations. The ADRC presents the following satisfactory performance indices:

1. Overshoot: overshoot can be reduced by using ADRC

2. Rise Time: Which minimum as compare to the other control techniques

3. Steady state error: almost zero

REFERENCES

[1] N. S. Nise, Control System Engineering-Sixth Edition. ISBN-13:978-0470917695, Wiley, December 8, 2010

[2] H. Saddad, Power System Analysis Third edition, New York: McGraw-Hill, January 1, 2011.

[3] P. Kundur, Power System Stability and Control. New York: McGraw-Hill, 1994

[4] Y. Zhang, L. Dong, Z. Gao, "Load Frequency Control of Multiple-Area Power Systems,"American Control Conference, 10 July, 2009.

[5] Y. Li, L. Dong, Y.You, Y. Gao, and X. Liao. "Load frequency control based on ADRC for interconnected power system." In Control Conference (CCC), 2014 33rd Chinese, pp. 8158-8163. IEEE, 2014.

[6] Nagarjuna, K., and M. S. Krishnarayalu. "ADRC for Two Area-LFC." International Journal of Engineering Research & Technology (IJERT) Vol 3, 2014.

[7] A. J. Wood, and B.F. Wollenberg,"Power Generation, Operation, and Control," 2nd edition., Wiley India (P.) Ltd., 4435/7, Ansari Road, Daryaganj, New Delhi 110 002, India, Chapter 9, pp. 328-360, 2007.

[8] F. Liu,Y. Li, Y. Cao, J. She, and M. Wu. "A two-layer active disturbance rejection controller design for load frequency control of interconnected power system." IEEE Transactions on Power Systems 31, no. 4: 3320-332, 2016.

[9] M. V., Srikanth, and N. Yadaiah. "An AHP based optimized tuning of Modified Active Disturbance Rejection Control: An application to power system load frequency control problem." ISA transactions, 2018.

[10] Z. Gao, "Active disturbance rejection control: a paradigm shift in feedback control system design", in: Proceedings of the American Control Conference, pp. 2399–2405, 2006.

[11] K. Liao, and Y. Xu. "A robust load frequency control scheme for power systems based on second-order sliding mode and extended disturbance observer." IEEE Transactions on Industrial Informatics 14, no. 7: 3076-3086, 2018.

[11] J. Han "From PID to active disturbance rejection control." IEEE transactions on Industrial Electronics 56, no. 3 :900-906, 2009.

[12] S. Li, J. Yang, W. H. Chen, and X. Chen. "Disturbance observer-based control: methods and applications". CRC press, 2016.

[13] M.M. Rahman, A. H. Chowdhury, and M. A. Hossain. "Improved Load Frequency Control Using a Fast Acting Active Disturbance Rejection Controller." Energies 10, no. 11 1718, 2017.

A Comparative Study of Various Methods of Phasor Measurement Unit Algorithms

Zahidur Rahim Talukder
BUET
Dhaka, Bangladesh
zahidur.eee.buet@gmail.com

Kazi Nishat Tasnim
BUET
Dhaka, Bangladesh
tasnimkazinishat@gmail.com

Md. Shamim Reza
BUET
Dhaka, Bangladesh
shamim_reza@eee.buet.ac.bd

***Abstract*--Phasors are significant to power system analysis in both steady-state and dynamic conditions. IEEE Standard C37.118.1 2011 along with its recent amendment C37.118.1a2014 define synchrophasor measurements with the focus on the behavior of phasor measurement units (PMUs) under steady-state and dynamic operating conditions. The synchrophasors can find many applications in modern power systems including real-time status monitoring, precise state estimation, validating simulation models, protection, taking centralized control actions using wide-area measurements and integrating renewable and intermittent resources. In this paper, various PMU algorithms are compared to find out the best algorithm which gives best result in specific conditions. As different research papers suggested and focused on different criterions, it is very difficult and cumbersome task to accumulate all data in a common platform and compare them. After vigorous research a unique way is developed to present overall comparison result in a well-defined and more understandable manner.**

Keywords--Synchrophasor, Taylor, EPLL, Kalman, Clarke, DFT, FIR Filters.

I. INTRODUCTION

Phasor Measurement Unit (PMU) has become an important tool for power system monitoring and control. Being a new a concept, the main advantage of this synchrophasor measurement over the existing SCADA system [1] is that their measurements are time-stamped using a reference synchronization signal, generally provided by a Global Positioning System receiver, and are taken much faster. In this way, PMUs can yield phase information with much greater accuracy.

IEEE Standard C37.118.1-2011 [2] along with its amendment C37.118.1a-2014 [3] mention requirements for synchrophasor, frequency, and rate of change of frequency (ROCOF) particularly in the presence of amplitude and phase modulated signals, linear ramp of frequency, and step changes. The standard defines two performance classes of PMUs, namely P and M. The P-class is intended for applications requiring fast measurement response time, while the M-class is there for measurement accuracy. A standard compliant PMU should meet all the requirements at least for one class. The main differences between the two performance classes are the test conditions for steady state performance and the requirements for dynamic condition. On the other hand, the standard leaves to PMU manufacturers free choice on hardware, software architecture, and algorithm for phasor, frequency, and ROCOF computation.

Various algorithms have proposed so far and which can be broadly divided into two categories 1) DFT- based approach 2) Non-DFT- based approach. Due to its less computational burden, more tolerance in harmonics and easy implementation, DFT based approach is used in the standard [2] and many authors have tried to implement it after slight modifications. The classic DFT based method works well when the system frequency is close to nominal frequency but due to spectral leakage issue in DFT, a significant error will arise when deviated from the nominal value. A lot of efforts have been made to improve the accuracy of DFT-based estimation algorithms under off-nominal frequency conditions. The common DFT based approaches includes Interpolated DFT based approach [4]-[5], Adaptive phasor and Frequency algorithm [6], Taylor Fourier Multifrequency Model [7], Clarke Transformation based DFT algorithms [8].

The non-DFT based algorithms includes Kalman filter based approach [9], Enhanced phase-locked loop [11], recursive approach [12], Taylor Fourier transform [13], Taylor-Kalman filters [14], Adaptive Cascaded filters [16], fast and accurate PMU algorithms [17], Polynomial Phase-Locked-Loop Taylor–Fourier Filters [18], Space Vector-Based PMU Algorithms [19], Positive-Sequence Synchrophasor Estimation Algorithm [20] etc.

From this vast range of algorithms, it is really tough to select appropriate algorithms for particular purposes. This paper gives a general overview of various existing algorithms and a comparison of these algorithms in a common platform which certainly be helpful for the users to select required algorithms.

The organization of this paper is as follows: Section II describes various algorithms, Section III summarizes results, Section IV describes discussion and lastly Section V gives conclusion by discussing future scope.

II. VARIOUS METHODS DISCUSSION

In PMU standard [2], a sinusoidal signal can be represented as,

$$X(t) = A(t)\,cos(\psi(t)) \quad (1)$$

Where, $\psi(t) = 2\pi f t + \phi(t)$

Where, Amplitude (A), frequency (f), Phase (ψ) & Phase shift (ϕ). Frequency is defined as

$$f(t) = \frac{d\psi(t)}{2\tau dt} \quad (2)$$

The ROCOF is defined as,

$$ROCOF(t) = \frac{df(t)}{dt} \quad (3)$$

Various PMU algorithms are described below:

A. Enhanced Phase-Locked Loop (EPLL)

Conventional PLL method uses Phase Detection Scheme (PD), Low Pass Filter (LPF) and a Voltage Controlled Oscillator (VCO) which generates a signal whose frequency is proportional to the input signal.

Though PLL system works properly for three phase balanced system, it faces obstacles when operating with three phase unbalanced system as it generates double frequency ripple in the loop. To mitigate this problem EPLL [11] is proposed which continuously updates the parameters to reduce error. If (2) is the output of the system, then an error index is defined (3). Hat sign is used to indicate estimated values.

$$y(t,P) = \sqrt{2}\hat{A}\cos(\hat{\psi}(t)) = \sqrt{2}\hat{A}\cos(2\pi\hat{f}t + \hat{\phi}) \quad (4)$$

$$J(t,P) = e(t,P)^2 = [X(t) - Y(t,P)]^2 \quad (5)$$

Vector parameter P contains $\hat{A}, \hat{\omega}, \hat{\phi}$. Gradient Descend Method (GDM) is use for continuous adjustment of P vector for minimizing the error index. The following sets of differential equation are obtained from $\dot{P} = -K(\partial J / \partial P)$, here *K* is positive definite matrix.

$\dot{\hat{A}} = 2k_1 e\,\widehat{cos\,\psi}$, $\dot{\hat{\omega}} = -2k_2 e\,\widehat{sin\,\psi}$, $\dot{\hat{\psi}} = -2k_3 e\,\widehat{sin\,\psi} + \hat{\omega}$

Here,

$e(t) = X(t) - \sqrt{2}\hat{A}\,\widehat{cos\,\psi}$

B. Taylor^K-Kalman Filter

In this method [13], a bandpass signal model of a power oscillation is defined by the state transition matrix of the K-th order Taylor approximation to the dynamic phasor and its first derivatives.

The signal is simplified for complex exponential function as,

$$X(t) = Re\{p(t)e^{j2\pi f_1 t}\}; -T/2 \le t \le T/2 \quad (6)$$

Here, $p(t) = A(t)e^{j\phi(t)}$ is referred to as dynamic phasor. The complex dynamic phasor function *p(t)* can be approximated by Kth Taylor polynomial centered at t_o. If *p(t)* is a smooth function the approximation would be more exact as $t \rightarrow t_o$ and then the truncated signal (7) can be applied at any time instance as the time interval is very short.

$$X_k(t) = Re\{h^T P_k(t) e^{j2\pi f_1 t}\} = Re\{h^T r_k(t)\} \quad (7)$$

Here, h^T extracts the first component with K zeros and *r(t)* is the rotated vector. Finally, the truncated signal model is given by (8) that shows the instantaneous dependence of the signal model on the dynamic phasor.

$$X_k(n) = \frac{1}{2}(h^T h^T)\begin{pmatrix} r_k(n) \\ \bar{r}(n) \end{pmatrix} \quad (8)$$

Kalman recursive method will be used to minimize the error for *nth* cycle according to Kalman gain (9)

$$K(n) = P^-(n)H^T(HP^-(n)H^T + \sigma_\omega^2)^{-1} \quad (9)$$

in which, $H = (h^T h^T)$ and σ_ω^2 is the variance of the measurement unit. State of the signal model is updated according to the Kalman gain.

To avoid the tradeoff between P (protection) and M (measurement) classes, the center of the window is considered fixed at Δt_a equal to one nominal cycle before the reference time and the synchrophasor is obtained by [fast@accu REF]

$$A_c = A + A'\Delta t_a$$
$$\varphi_c = \varphi + f\Delta t_a$$

C. Interpolated DFT Algorithm

Interpolated Discrete Fourier Transform (IpDFT) [4] can estimate synchro phasor parameters both in steady state and dynamic condition when two or three cycle length observation intervals are used.

The relationship between $f_{in,}$ *and* f_s is (11) where, *l* and δ are the integer and fractional parts of number of acquired signal cycle λ_o.

$$\frac{f_{in}}{f_s} = \frac{\lambda_o}{M} = \frac{l+\delta}{M} \quad (11)$$

In IpDFT method the signal (10) is multiplied by a window sequence,

$$w(m) = \sum_{h=0}^{H-1} a_h \cos\left(\frac{2\pi h}{M}(m + \frac{1}{2})\right) \quad (12)$$

Here, $m = -\frac{M}{2}, \ldots, \frac{M}{2} - 1$ and *H* is the number of window terms and *h = 0, ..., H-1* indicates the window coefficient. DFT of the windowed signal, $x_w(m) = x(m).w(m)$ is calculated as,

$$X_w(k) = \frac{A}{\sqrt{2}}[W(k-\lambda_o)e^{j\phi_o} + W(k+\lambda_o)e^{-j\phi_o} \quad (13)$$

If, $|\lambda| << M$ then, the DTFT of the window signal (12) is expressed as,

$$W(\lambda) = \frac{M\sin(\pi\lambda)}{\pi}\sum_{h=0}^{H-1}(-1)^h a_h \frac{\lambda^2}{\lambda^2 - h^2} \quad (14)$$

After, calculating δ, phasor amplitude and phase are measured according to,

$$\hat{A} = \frac{\sqrt{2}|X_w(l)|}{|W(-\hat{\delta})|}$$
$$\hat{\phi} = phase\{X_w(l)\} \quad (15)$$

When, K cycles of signal (10) are observed then window of order H≤ K must be used for accurate phasor estimation.

D. Finite Impulse Response (FIR)

A complex analytical signal $X_p(t)$ of the input signal *x(t)* is obtained by passing the signal through K-tabs FIR bandpass filter [6] whose, complex exponential with respect to the center frequency is,

$$h_{BP}(k, f_c) = h_{\cos}(k, f_c) + jh_{\sin}(k, f) \; ; k = 0,1,\ldots,K-1$$

The center frequency f_c can be re-tuned as it's an adaptive filter. FIR prototype derived form a 4 cycle Taylor window is used for M class accuracy,

$h_M(n) = \frac{1}{2K}\sum_{h=0}^{8}(-1)^h A_h \cos(2\pi h \frac{n}{2K})$; *n = 0,1,..., 2K-1*

1.2 cycle of the Taylor window is used for P class accuracy,

$$h_p(n) = K + (1-K)W_h(n) \text{ ; } n = 0,1,..., K-1$$

Here, K = 0.5189 and W_h is Henning window.

E. DFT based Demodulation Approach

Demodulation system [6] is efficient for tracking phasor. Demodulation system (16) first brings the signal back to the bass band with N_1 point low pass FIR filter,

$$X[\mathrm{k}] = \sum_{n=k}^{k+N_1} h_1[n-k].\mathrm{e}^{-2\pi f_o(n-k)T}.x[n] \tag{16}$$

Here, k is the discrete time index and T is the sampling time period. The raw phasor is filtered again with N_2 point low pass FIR filter (17) and replaced on (16) to get modified demodulated signal (18).

$$h_1(n) = \frac{1}{N_1}, n = 0,1,...,\mathrm{N}-1 \tag{17}$$

$$X(k) = \frac{1}{N_1}\sum_{n=0}^{N_1} x[n-k].e^{-2\pi j f_o nT} \tag{18}$$

It is seen that the FIR filtering method can be implemented recursively as,

$$X[k] = \frac{1}{N_1}[x[k] - x[k-N_1]] + W_{fo}X[k-1]$$

Where, $W_{fo} = e^{-2\pi j f_o T}$.

This is a very fast process and cannot provide good attenuation at high frequency, but it is very efficient in rejecting the harmonics in the absence of frequency offset.

F. A Novel Modular Positive Sequence Synchrophasor Estimation Algorithm

In this method there are four steps are included for estimating positive sequence synchrophasor for both P and M classes. In the first step, the signal is demodulated to separate the positive, negative and zero sequence signal (19) in frequency domain to eliminate the zero-sequence signal.

$$x_{dq}[n] = a_1[n]\begin{bmatrix}\cos(\phi_1[n])\\ \sin(\phi_1[n])\end{bmatrix} + a_2[n]\begin{bmatrix}\cos(2\omega_o nT + \phi_2[n])\\ -\sin(2\omega_o nT + \phi_2[n])\end{bmatrix} \tag{19}$$

Then, prefiltering is used to mitigate the noise and interference according to,

$$y_{dq}[n] = \begin{bmatrix}y_d[n]\\ y_q[n]\end{bmatrix} = \begin{bmatrix}(h_{PF} * x_d)[n]\\ (h_{PF} * x_q)[n]\end{bmatrix} = a[n]\begin{bmatrix}\cos(\phi[n])\\ \sin(\phi[n])\end{bmatrix}$$

After that the synchrophasor parameters are extracted according to PLL algorithm. A phase predictor is used inside the loop for accuracy.

Forth step is to use a Narrow-band Compensation Algorithm (NCA) to compensate the effects of the filters on the main signal.

G. Clarke Transformation Based DFT Transformation

In this method [8], Clarke Transformation generates a second signal that is exactly 90-degree phase angle difference from the input single-phase signal that help to reduce the estimation error for a wide frequency range and then Weighted Least Square Taylor Fourier (WLS-TF) is used to make their amplitude identical.

At first the signal is passed through a Moving Average Filter (MAF) to get rid of the noise. Then the three phase signals (20) are passed through Clarke transformation unit.

$$\begin{aligned} X_A(k) &= \sqrt{2}A * \cos\left\{\frac{2(f_0+\Delta f)k\pi}{f_0 N} + \varphi\right\} \\ X_B(k) &= \sqrt{2}A * \cos\left\{\frac{2(f_0+\Delta f)(k-\frac{N}{3})\pi}{f_0 N} + \varphi\right\} \\ X_C(k) &= \sqrt{2}A * \cos\left\{\frac{2(f_0+\Delta f)(k+\frac{N}{3})\pi}{f_0 N} + \varphi\right\} \end{aligned} \tag{20}$$

Equation (20) can be expressed as,

$$X_{ABC}(k) = \begin{bmatrix}X_A(k)\\ X_B(k)\\ X_C(k)\end{bmatrix} \tag{21}$$

T is the Clarke transformation matrix (22).

$$T = \begin{bmatrix}\frac{2}{3} & \frac{-1}{3} & \frac{-1}{3}\\ 0 & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}}\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\end{bmatrix} \tag{22}$$

According to (23) V(k) is obtained, and from that Amplitude, Phase, Frequency and ROCOF are estimated.

$$V_{\alpha\beta\gamma}(k) = TX_{ABC}(k) \tag{23}$$

H. Other Algorithms:

Other algorithms like adaptive cascaded filter [16], fast and accurate method [17] which is basically Taylor Fourier Transform (TFT), Taylor Fourier multi-frequency method [7], Polynomial PLL Taylor Fourier Filter [18], Space Vector based algorithms [19] are the slight modification of above-mentioned algorithms. Due to space limitation and lack of uniqueness, description of these algorithms is not included in this paper. But the result of these algorithms is tabulated below, and users can easily go through these algorithms from reference if they need.

III. RESULTS

The parameter that is needed to be calculated is the Total Vector Error (TVE), Frequency Error (FE), and Rate OF change of Frequency Error (RFE). The standard formula for calculating this are,

$$TVE = \sqrt{\frac{(\hat{X}_r(n) - X_r(n))^2 + (\hat{X}_i(n) - X_i(n))^2}{(X_r(n))^2 + (X_i(n))^2}}$$

$$FE = |f_{true} - f_{measured}||\Delta f_{true} - \Delta f_{measured}|$$

$$RFE = |(df/dt)_{true} - (df/dt)_{measured}|$$

Now to be compliant with this standard, a PMU shall provide synchrophasor, frequency, and ROCOF measurements that meet the requirements of the standard. These requirements shall always be met and under all configurations whether the PMU function is a stand-alone physical unit or included as part of a multifunction unit. There are two types of compliances for testing one for steady state compliances and another for dynamic compliances under modulated test signal, ramp of system Frequency and input step change. The result based on the standard and from different algorithms are tabulated below to show which algorithms maintain the standard and to see which class perform how under different conditions.

TABLE I STEADY STATE COMPARISN OF 11 DIFFERENT ALGORITHMS

STEDY STATE TEST	**Off-nominal Frequency**			**Harmonic Distortion**			**Out of Band Interference**		
	TVE (%)	***FE***	***RFE***	***TVE (%)***	***FE***	***RFE***	***TVE (%)***	***FE***	***RFE***
IEEE Standard	1 (P) 1(M)	0.005 (P) 0.005 (M)	0.01 (P) 0.01 (M)	1 (P) 1 (M)	0.005 (P) 0.025 (M)	0.01(P) 6(M)	0(P) 1.3(M)	0.005(P) 0.005(M)	0.01(P) 6(M)
EPLL Method	0.7 (P) 0.7 (M)	- -	- -	- -	- -	- -	- -	- -	- -
Taylor^K Kalman Filters	1 (P) 1(M)	- -	- -	- -	- -	- -	- -	- -	- -
Adaptive Cascaded Filters	0.035 (P) 0.7(M)	0.001 (P) 0.001 (M)	0.02 (P) 0.02 (M)	0.02 (P) 0.02 (M)	0.001(P) 0.001(M)	0.001(P) 0.001(M)	10(P) 10(M)	1(P) 0.05(M)	50(P) 50(M)
Interpolated DFT Method	0.06 (P) 0.06 (M)	- -	- -	- -	- -	- -	- -	- -	- -
Fast & Accurate Method (Taylor)	0.22 (P) 0.22(M)	0.004 (P) 0.004 (M)	0.0045 (P) 0.0084 (M)	0.0023 (P) 0.0228 (M)	0.00011(P) 0.00113(M)	0.0045(P) 0.0084(M)	0.28(P) 0.28(M)	0.0062(P) 0.0062(M)	0.55(P) 0.55(M)
Adaptive Phasor for Wide Range	0.07 (P) 0.07(M)	- -	- -	- -	- -	- -	- -	- -	- -
Taylor Fourier (Multi-freq)	- -	- -	- -	0.011 (P) 0.011 (M)	0.00036(P) 0.00036(M)	- -	- -	- -	- -
PLL Taylor Fourier Filters	- -	- -	- -	0.015 (P) 0.015 (M)	8.6e-5(P) 8.6e-5(M)	4.9e-2(P) 4.9e-2(M)	- -	- -	- -
Novel Modular Method	- -	- -	- -	0.0072(P) 0.0027 (M)	0.0026(P) 1.79e-4(M)	0.2766(P) 5.0e-3(M)	0.0776(P) 0.0776(M)	0.0066(P) 0.0066(M)	0.1675(P) 0.1675(M)
Space Vector Method	0 (P) 0 (M)	0 (P) 0 (M)	0 (P) 0 (M)	6.74e-4 (P) 2.22e-3 (M)	1.1e-2 (P) 1.1e-2 (M)	4.6e-4(P) 4.6e-4(M)	- -	- -	- -
Clarke Trans-formation	0.006(P) 0.006(M)	0.25 (P) 0.25 (M)	0.014 (P) 0.014 (M)	3e-4 (P) 3e-4 (M)	9.7e-3(P) 9.7e-3(M)	5.9e-4 (P) 5.9e-4(M)	0.045(P) 0.045(M)	0.19(P) 0.19(M)	0.034(P) 0.034(M)

a. P= P class PMU, M= M class PMU

TABLE II DYNAMIC STATE COMPARISN OF 11 DIFFERENT ALORITMS

DYNA-MIC STATE TEST	**Modulated Signal**			**Frequency Ramp Test**			**Input Step Change**			
	TVE (%)	***FE***	***RFE***	***TVE (%)***	***FE***	***RFE***	***Response Time (ms)***	***Over-shoot (%)***	***Frequency Response Time (ms)***	***ROCOF Response Time (ms)***
IEEE Standard	3(P) 3(M)	0.06(P) 0.3(M)	3(P) 30(M)	1(P) 1(M)	0.01(P) 0.01(M)	0.4(P) 0.4(M)	- -	- -	- -	- -
EPLL Method	- -	- -	- -	0.025(P) 0.025(M)	0.05(P) 0.05(M)	1(P) 1(M)	- -	- -	- -	- -
Taylor^{K} Kalman Filters	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -
Adaptive Cascaded Filters	- -	- -	- -	0.02(P) 0.02(M)	1.2-4(P) 1.2e-4(M)	0.08(P) 0.08(M)	- -	- -	- -	- -
Interpolated DFT Method	0.52(P) 0.52(M)	- -	- -	0.29(P 0.29(M)	- -	- -	30(P) 30(M)	- -	65(P) 65(M)	45(P) 45(M)
Fast & Accurate Method (Taylor)	0.022(P) 0.022(M)	0.01(P) 0.031(M)	0.49(P) 2.69(M)	7e-4(P) 7e-4(M)	3.3e-4(P) 3.3e-4(M)	3.8e-3 (P) 3.8e-3(M)	22(P) 22(M)	4.6(P) 4.6(M)	69(P) 69(M)	76(P) 76(M)
Adaptive Phasor for Wide Range	8.478(P) 8.478(M)	0.098(P) 0.098(M)	- -	- -	- -	- -	65.1(P) 65.1 (M)	10.28(P) 10.28(M)	45(P) 45(M)	55(P) 55(M
Taylor Fourier (Multi-freq)	0.124(P) 0.124(M)	- -	- -	- -	- -	- -	- -	- -	- -	- -
PLL Taylor Fourier Filters	4e-3(P) 4e-3(M)	- -	- -	- -	- -	- -	- -	- -	- -	- -
Novel Modular Method	0.148(P) 1e-2(M)	0.001(P) 0.021(M)	0.274(P) 1.165(M)	- -	- -	- -	20(P) 33(M)	4.3(P) 7.5(M)	89(P) 82(M)	10(P) 83(M)
Space Vector Method	0.077(P) 0.077(M)	0.0017(P) 0.0017(M)	0.021(P) 0.021(M)	0.028(P) 0.028(M)	9.8e-8(P) 9.8e-8 (M)	0(P) 0(M)	42.5(P) 42.5 (M)	0(P) 0(M)	67.5(P) 67.5(M)	72.5(P) 72.5(M)
Clarke Trans-formation	1.408(P) 1.408(M)	0.0978(P) 0.0978(M)	3.14(P) 3.14(M)	0.086(P) 0.086(M)	3e-5(P) 3e-5(M)	0.001 (P) 0.001 (M)	- -	- -	- -	- -

a. P= P class PMU, M= M class PMU

IV. DISCUSSION

There have been several comparison-based works in other domains [10][15]. In PMU domain, most of the algorithms tried to maintain the IEEE standards in most of the cases but few results outperform others. Some of the algorithms are concerned with few cases only, this is might be due to time shortage or the limitation of the algorithms. But we didn't find any single algorithm which is okay with the standard and maintained all the requirements. Some algorithms have calculation burdens, and some are easy but less accurate. So, it is tough to say which algorithms are the best and which one is the worst. From the result we found for off nominal frequency condition space vector method out performs other in the case of TVE (%), FE and RFE. Again, in presence of harmonic distortion, space vector method is better even though dynamic phasor-based algorithms do well. For out of band interference condition CT-DFT based algorithm performs better than other. For modulated signal and frequency ramp test space vector-based method and fast and accurate P+M method perform better than other algorithms. For input step change fast and accurate P+M, space vector-based method, wide frequency range method and IpDFT based algorithm perform better.

V. CONCLUSION

There are numerous methods proposed in the literature for the estimation of the phasor, frequency and rate of change of frequency. All of the algorithms try to maintain all the requirements as mentioned in the standard. But from the literature we found some limitations as most of the algorithms don't test all the cases and take different conditions during the test. In general, it is not so straightforward to identify which approach possesses the superior performance. For steady state compliances the Space Vector based algorithm performs well whereas for Dynamic State compliances no single algorithms do remarkably well. So, there are lot of work to do to increase the performance and accuracy of phasor measurement unit. Our aim was not to find novel algorithms rather we worked on existing algorithms to find out the limitations of the various proposed algorithms and give a proper guideline for the users as well as for the scholars. To provide the proper guidelines for the given applications, several commonly used methods have been reviewed and compared in this paper, according to the viewpoint of the IEEE Standard.

REFERENCES

[1] S. Azizi, G. B. Gharehpetian and A. S. Dobakhshari, "Optimal Integration of Phasor Measurement Units in Power Systems Considering Conventional Measurements", in IEEE Transactions on Smart Grid, vol. 4, no. 2, pp. 1113-1121, June 2013.

[2] IEEE Standard for Synchrophasor Data Transfer for Power Systems", in IEEE Std C37.118.2-2011 (Revision of IEEE Std C37.118-2005), vol., no., pp.1-53, 28 Dec. 2011.

[3] K. Martin and G. Brunello, "An overview of the IEEE Standard C37.118.2 — Synchrophasor Data Transfer for Power Systems," 2014 *IEEE PES General Meeting | Conference & Exposition, National Harbor, MD*, 2014, pp. 1-1.

[4] D. Belega and D. Petri, "Accuracy Analysis of the Multicycle Synchrophasor Estimator Provided by the Interpolated DFT Algorithm", in *IEEE Transactions on Instrumentation and Measurement,* vol. 62, no. 5, pp. 942-953, May 2013.

[5] B. Khettaoui and M. Boudour, "Synchrophsor Estimation Using the Three-point Interpolated DFT Based on Polynomial Approximation," *2018 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM)*, Algiers, 2018, pp. 1-5.

[6] I. Kamwa, S. R. Samantaray and G. Joos, "Wide Frequency Range Adaptive Phasor and Frequency PMU Algorithms", in *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 569-579, March 2014.

[7] M. Bertocco, G. Frigo, C. Narduzzi, C. Muscas and P. A. Pegoraro, "Compressive Sensing of a Taylor-Fourier Multifrequency Model for Synchrophasor Estimation", in *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 12, pp. 3274-3283, Dec. 2015.

[8] L. Zhan, Y. Liu and Y. Liu, "A Clarke Transformation-Based DFT Phasor and Frequency Algorithm for Wide Frequency Range", in *IEEE Transactions on Smart Grid*, vol. 9, no. 1, pp. 67-77, Jan. 2018.

[9] Sajedul Talukder, Md. Iftekharul Islam Sakib, Md. Faruk Hossen, Zahidur Rahim Talukder, Md. Shohrab Hossain, "Attacks and Defenses in Mobile IP: Modeling with Stochastic Game Petri Net", In IEEE International Conference on Current Trends in Computer, Electrical, Electronics and Communication, Sep. 2017.

[10] J. A. de la O Serna and J. Rodriguez-Maldonado, "Instantaneous Oscillating Phasor Estimates with Taylor^K-Kalman Filters", in *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2336-2344, Nov. 2011.

[11] M. Karimi-Ghartemani, B. Ooi and A. Bakhshai, "Application of Enhanced Phase-Locked Loop System to the Computation of Synchrophasors", in *IEEE Transactions on Power Delivery*, vol. 26, no. 1, pp. 22-32, Jan. 2011.

[12] P. K. Dash and S. Hasan, "A fast recursive algorithm for the estimation of frequency, amplitude, and phase of noisy sinusoid," *IEEE Trans. Ind. Electron*., vol. 58, no. 10, pp. 4847–4856, Oct. 2011.

[13] M. A. Platas-Garza and J. A. de la O Serna, "Dynamic phasor and frequency estimates through maximally flat differentiators*," IEEE Trans. Instrum. Meas*., vol. 59, no. 7, pp. 1803–1811, Jul. 2010.

[14] Sajedul Talukder, Md. Iftekharul Islam Sakib, Zahidur Rahim Talukder, Upoma Das, Arnob Saha and Nur Sultan Nazar Bayev, "USenSewer: Ultrasonic Sensor and GSM-Arduino Based Automated Sewerage Management", In IEEE International Conference on Current Trends in Computer, Electrical, Electronics and Communication, Sep. 2017.

[15] J. A. de la O Serna and J. Rodriguez-Maldonado, "Instantaneous oscillating phasor estimates with Taylor k-Kalman filters," *IEEE Trans. Power Syst*., vol. 26, no. 4, pp. 2336–2344, Nov. 2011.

[16] A. J. Roscoe, I. F. Abdulhadi and G. M. Burt, "P and M Class Phasor Measurement Unit Algorithms Using Adaptive Cascaded Filters", in *IEEE Transactions on Power Delivery*, vol. 28, no. 3, pp. 1447-1459, July 2013.

[17] P. Castello, J. Liu, C. Muscas, P. A. Pegoraro, F. Ponci and A. Monti, "A Fast and Accurate PMU Algorithm for P+M Class Measurement of Synchrophasor and Frequency", in *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 12, pp. 2837-2845, Dec. 2014.

[18] J. A. de la O Serna, "Synchrophasor Measurement with Polynomial Phase-Locked-Loop Taylor–Fourier Filters", in *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 2, pp. 328-337, Feb. 2015.

[19] S. Toscani, C. Muscas and P. A. Pegoraro, "Design and Performance Prediction of Space Vector-Based PMU Algorithms", in *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 3, pp. 394-404, March 2017.

[20] F. Messina, P. Marchi, L. Rey Vega, C. G. Galarza and H. Laiz, "A Novel Modular Positive-Sequence Synchrophasor Estimation Algorithm for PMUs", in *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 6, pp. 1164-1175, June 2017.

Water Cooled Chiller Based HVAC System Used in a Linear Generator for Oceanic Wave Energy Conversion

Selim Molla and Omar Farrok
Department of Electrical and Electronic Engineering
Ahsanullah University of Science & Technology
Dhaka 1208, Bangladesh
eee.n_selim@yahoo.com and omarruet@gmail.com

***Abstract*—The existing linear generator are mostly made of conventional iron core and permanent magnet (PM). During the operation of this generator, it becomes heated because of the magnetic and mechanical losses. When the temperature rises in the linear generator, then the output power becomes low because of degradation of the remanence magnetism of the PMs. In this paper, a cooling system is proposed for the linear generator for preventing temperature rise. Simulation is executed by ANSYS Maxwell software to perform finite element analysis. To observe the temperature effect, a permanent magnet linear generator along with a heating, ventilation, and air conditioning (HVAC) are designed. The generator is placed inside a water cooled wind chiller based HVAC system. Simulation results show that, the linear generator using the proposed HVAC system produces 23.9 % more output power than that of the same without using HVAC.**

Keywords—cooling, energy conversion, generators, HVAC, magnetic materials, permanent magnet machines, renewable energy sources, temperature control, wave power.

I. INTRODUCTION

The generation of electricity is very important in order to minimize the global energy crisis [1]. Most of the power generation system uses fossil fuels which are harmful for not only environment but also closely involves to the greenhouse effect. To avoid this situation, the importance of renewable energy is increased day by day. At present, wind, solar, and oceanic wave are popularly used to extract the electrical energy. Out of these energies, the sea wave energy contains some effective advantages such as high energy density, much predictability, and availability [2], [3].

The basic concept of the generation of electricity from the oceanic wave is considered in 1799 [4]. But this energy conversion system becomes much popular in 1970s after occurring the oil crisis [5]. According to the installed location, two types of converters are found which are onshore and offshore. In general, offshore is suitable place for the large scale generation than onshore installed converter [6]. At present, there are two types of converter applied for the conversion of oceanic wave energy viz. rotational generator and direct drive linear generator (DDLG). It is found that, the topology of DDLG is much attractive than rotational type generator because of having no complexity for mechanical gear box design [1].

A DDLG based wave energy converter is presented in Fig. 1. It is consisted of a linear generator, inverter, and rectifier. The main elements of a linear generator (LG) are stator and translator. This translator is connected to a floating buoy by a connecting rod. When the translator moves with that buoy with respect to stator, the electricity is generated according to the Faraday's law of electromagnetic induction.

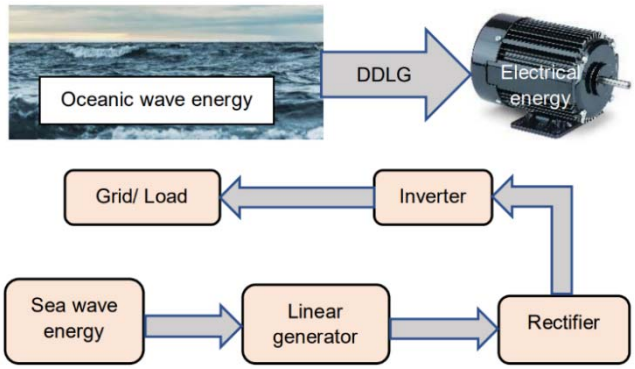


Fig. 1. Block diagram of the working principle of WEC.

The stator is made of magnetic core and coil winding and the translator is made of magnetic core and permanent magnet (PM). Recently, a 21.21W DDLG is developed where the magnetic attraction between the iron core and PM is applied in order to overcome the alternating magnetic field. It utilizes only 0.36 square meter area [7]. An LG is numerically simulated by using MATLAB which supplies 300W output power [8]. Another 1kW LG is validated both in experimentally as well as numerically where detent force is reduced [9]. An electromechanical design of a MW scale high temperature superconductor based LG is presented in [10] in which, the mechanical structure and cryogenic cooling system are highlighted.

Recently, some linear generators are designed by using high grade NdFeB PMs and magnetic cores in order to achieve better performance, higher output power, lower core loss, and higher efficiency. Such as, two types of high grade magnetic cores are applied in an LG in which Supermendur core based generator exhibits better performance [11]. During the operation of linear generator, it becomes heated. For this reason, the LG designed with an ordinary neodymium iron boron (NdFeB) PM is demagnetized. In order to avoid this problem, a high grade NdFeB is proposed in [12] which has higher operating temperature than that of the conventional NdFeB. The application of superconductors in a DDLG are presented in [13] and [14]. The effect of temperature in linear generator is presented in [15], in which NdFeB PM is used to achieve high performance at high temperature. This permanent magnet's maximum operating temperature is 200°C.

The conventional NdFeB PM cannot generate much output power because of decreasing the remanence magnetism at high temperature. This problem is solved by applying high grade NdFeB PM N28EH in an LG [12]. But, high grade PMs add an extra cost to the DDLG and has lower magnetic saturation point. As the generator becomes heated during

operation, for this reason, the PM becomes demagnetized for its property. As a result, the generator cannot produce much output power.

From the review of some research works on DDLG, it is found that, cooling system is not familiar with this machine [1], [5]–[15]. In this paper, a cooling system is included with an LG to avoid the problem of demagnetization due to heat generation and is compared with the same without cooling. The load currents, instantaneous, and rms powers are compared for both cases. As per simulation results, it is found that, the proposed LG with HVAC system generates much output power than that of the LG with no HVAC.

II. Mathematical Model

The temperature effect on conventional NdFeB permanent magnet is presented in Fig. 2. It is found that, the remanence magnetism and coercive force of the magnet at 100°C are lower than that of at 60°C. The highest coercive force, remanence, and energy product of this PM at 20°C are greater than other temperatures.

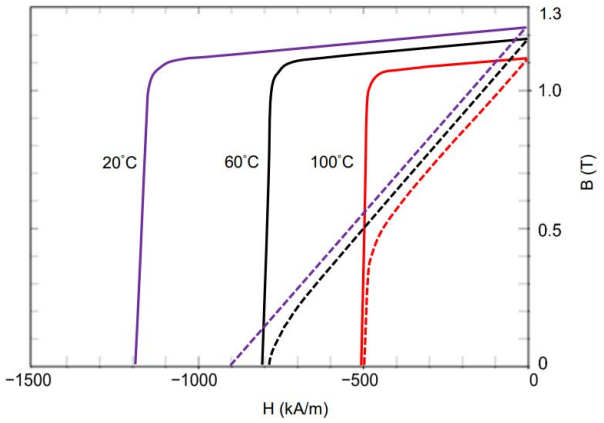


Fig. 2. Temperature effect of conventional NdFeB PM.

Since, the energy product of permanent magnet is high at low temperature and vice-versa, the low temperature is required to achieve high output power from DDLG placed near the shore of the ocean. The model of oceanic wave energy considers shallow water wave. In general, sine wave is considered for description of the energy of sea wave for the simulation of LG [16] where the nominal vertical speed is 0–2 m/s and time period is 4–6s. The model of shallow water is described in Fig. 3.

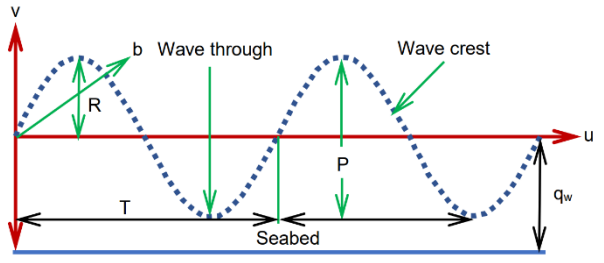


Fig. 3. Sinusoid form of the oceanic wave.

The amplitude, vertical distance, water depth, time period, and angular velocity are denoted by R, P, q_w, T, and ω, respectively. We know that, b is considered as zero as the sea wave propagates along u direction where the component u is horizontal and v is vertical and can be explained as

$$U = R\omega \frac{\cosh\{k(v+q_w)\}}{\sinh(kq_w)}\cos(ku-\omega t) \quad (1)$$

$$V = M\omega \frac{\sinh\{k(v+q_w)\}}{\sinh(kq_w)}\sin(ku-\omega t) \quad (2)$$

where k is the number of wave. In this case, T is much greater than wave height and q_w. That is why, $T >> q_w$ and $T >> R$. Hence, (1) and (2) can be described as

$$U_s = M\omega \frac{e^{k(v+q_w)}+e^{-k(v+q_w)}}{e^{kq_w}-e^{-kq_w}}\cos(ku-\omega t) \quad (3)$$

$$V_s = M\omega \frac{e^{k(v+q_w)}+e^{-k(v+q_w)}}{e^{kq_w}-e^{-kq_w}}\sin(ku-\omega t)\,. \quad (4)$$

Now, (3) and (4) becomes

$$U_s \cong \frac{M\omega}{kq_w}\cos(ku-\omega t) \quad (5)$$

$$V_s = M\omega\left(1+\frac{v}{q_w}\right)\sin(ku-\omega t)\,. \quad (6)$$

When $kq_w << 1$, then $\omega = k\sqrt{gq_w}$. If, $v = \frac{\omega}{k}$ and g is the gravitational acceleration, the speed of phase of the oceanic wave can be found as $s = \sqrt{gq_w}$.

III. Construction of the Linear Generator

The basic construction of a DDLG is presented in Fig. 4. It has mainly two parts viz. translator and stator. The generator contains three stators which are shown at left hand side.

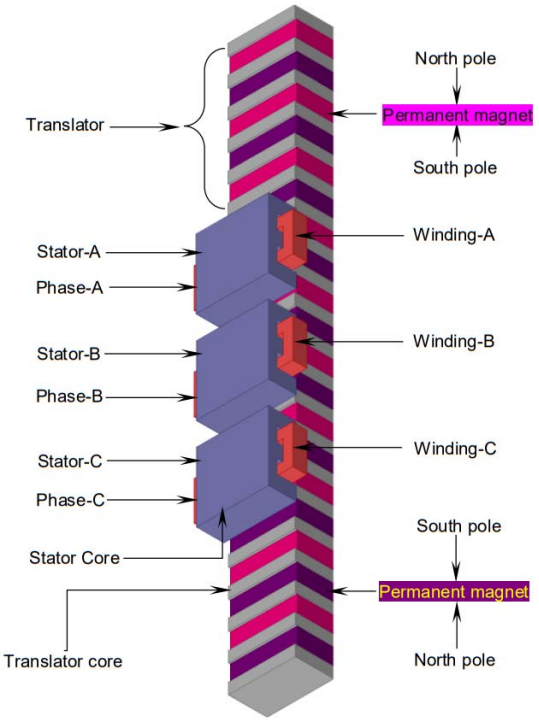


Fig. 4. Basic structure of proposed DDLG.

Each stator is made of "E" shaped magnetic core and copper windings. Winding-A, winding-B, and winding-C are placed inside the stator slots which are also shown in Fig. 4. The translator is shown beside the stator which is made of magnetic core and permanent magnets. The orientations of the permanent magnets are presented in the right-hand side of the translator. Air gap between the translator and stator is considered 3 mm. Table I shows the physical dimensions of the proposed DDLG. This translator is connected with a buoy that floats on oceanic wave. It moves upward and downward with the wave. Due to having connection with buoy, the translator also moves upward and downward with respect to stators. At this stage, emf is induced in each winding by following the Faraday's law of electromagnetic induction.

TABLE I. DIMENSIONS OF THE PROPOSED DDLG

Name of The Parameters	Value
Width of DDLG (m)	0.36
Width of translator (m)	0.25
Width of winding (m)	0.14
With of stator's core (m)	0.1
Height of translator (m)	2.04
Height of winding (m)	0.05
Height of magnetic core (m)	0.25

IV. THE PROPOSED DDLG WITH COOLING SYSTEM

A. Cooling system

The proposed generator with cooling system is shown in Fig. 5. Cooling system is consisted of a chiller, air handling unit (AHU), dehumidifier, supply chilled water pipe (SCWP), return chilled water pipe (RCWP), and control panel.

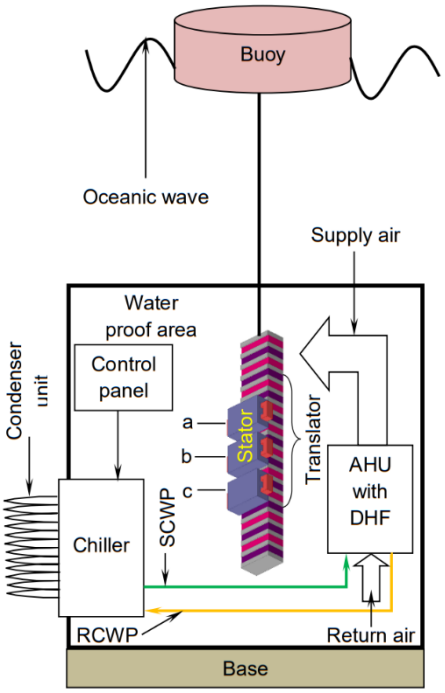


Fig. 5. Proposed cooling system with DDLG.

There are three different sensors used in this system. One is required to collect existing temperature of chilled water which is fitted on the RCWP and the other is required to measure the actual temperature and humidity in the linear generator. According to the signal of sensors, the chiller, AHU, and dehumidifier are operated automatically with the aid of control panel to maintain the desired temperature and humidity. In Fig. 5, a, b, and c indicates phase a, phase b, and phase c, respectively.

The basic block diagram of the proposed chiller is depicted in Fig. 6. Chiller is an electromechanical arrangement that reduces the temperature of water around 8°C. The main parts of this system are compressor, condenser, evaporator, filter core, solenoid valve, and expansion valve. The complete operation is maintained by a control panel according to the signal of chilled water sensor.

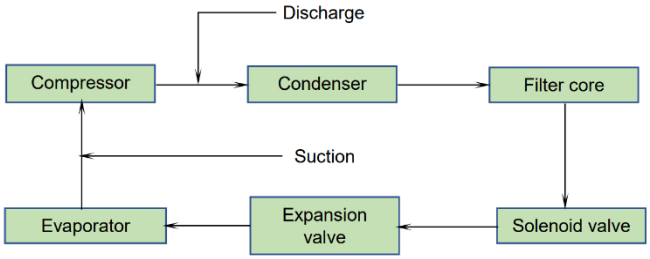


Fig. 6. Basic architecture of chiller.

The refrigerant R22 is filled in the whole chiller unit. In general, the suction pressure of refrigerant is kept 80 PSI at rest mode and 50– 65 PSI during the operation mode. On the other hand, the discharge pressure is considered around 250 PSI at operation mode and same as suction pressure at rest mode. When the compressor is started, the pressure of refrigerant gas increases from 80 PSI to 250 PSI and fed it to condenser unit. At this time, the discharge pipe becomes much heated approximately up to 70–80°C. This condenser unit is made of some copper coil and aluminum fins.

When the gaseous refrigerant enters into the chiller unit, its heat and pressure decreases for which phase transition occurs and its state changes into liquid. This liquid refrigerant goes to expansion valve through the filter core and solenoid valve. The filtered refrigerant's flow is controlled by a solenoid valve which is operated by control panel so that it reach a desired value. The expansion valve is a key element of the chiller which converts the liquid refrigerant into vapor according to the signal of a gas operated valve sensor. Further, the chilled refrigerant flows through the coils in evaporator which is also called cooling tower. When the water flows through the cooling coil, it also becomes chill. Finally, this refrigerant returns in the compressor through the suction pipe. Thus, it completes a cycle and the whole process runs continuously.

B. The Basic Operation of AHU Along With Dehumidifier

The network between evaporator of chiller and cooling coil of AHU is made by using stainless steel pipe. A chilled water pump is connected between these. The main components of AHU are cooling coil or fins coil, mixing box, temperature sensor in both supply and return duct, humidity sensor on return duct, motorized valve on chilled water return (CW-R) or chilled water supply (CW-S) pipe, and supply fan motor. Another return duct is fed to the dehumidifier which supplies the moisture free air to the mixing box of AHU. The chilled water is continuously flowed through the cooling coil by the pump. A controller is applied for receiving the signal of temperature and humidity sensor which are fitted on the

return duct as well as controls the motorized valve, dehumidifier, and speed of supply fan motor in order to maintain the desired temperature and humidity. The network of this system is depicted in Fig. 7.

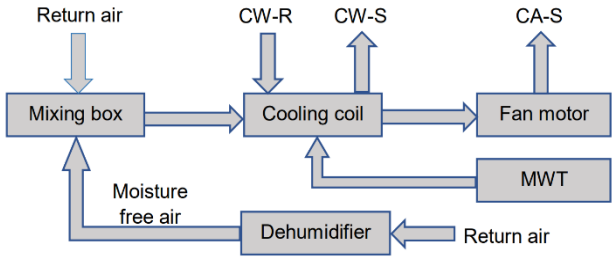


Fig. 7. Connection between AHU and dehumidifier.

When the fan motor is started, the return air flows through the fins coil and enter into the DDLG through supply air duct. The supplied air becomes cool because of passing through the cooling coil. The working principle of dehumidifier (DHF) is illustrated in Fig. 8. It is mainly made of a silica gel disc, electric heater, a blower fan, and a disc driving induction motor.

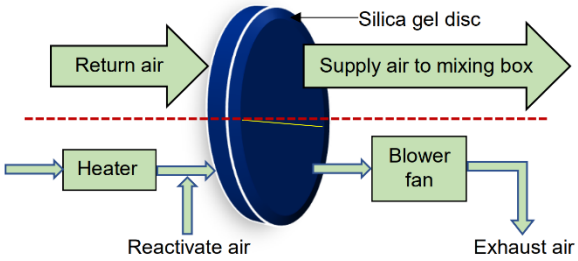


Fig. 8. Basic structure of dehumidifier.

The signal of dehumidifier sensor plays a vital role for controlling the whole system. The motor makes the disc rotating slowly. When the return wet air passes through the disc, the moisture is isoloted from the air and this air is supplied to the mixing box of AHU. At this time, the air passes a portion of the disc reaches to near saturation state. In order to avoid saturation, reactivation arrangement of silica gel bed are ensured. The whole system is seperated into two sections namely, active and reactive. At first, reactivation air is made by passing the fresh air through the heater. When this heated air passes through the wet disc by blower fan, the disc gets reactivation which is also a continous process.

V. Simulation Results

Simulation is performed for the DDLG considering both conditions, with and without using HVAC system. The velocity of translator is set to 1 m/s. Resistive load is connected to the output terminals of the proposed DDLG. The magnetic field intensity and flux density are shown in Fig. 9 and Fig. 10, respectively to observe magnetic properties. The output voltage of propsed linear generator is plotted in Fig. 11. The peak value of votages found for Phase-A, Phase-B, and Phase-C are 263V, 264V, and 256V, respectively.

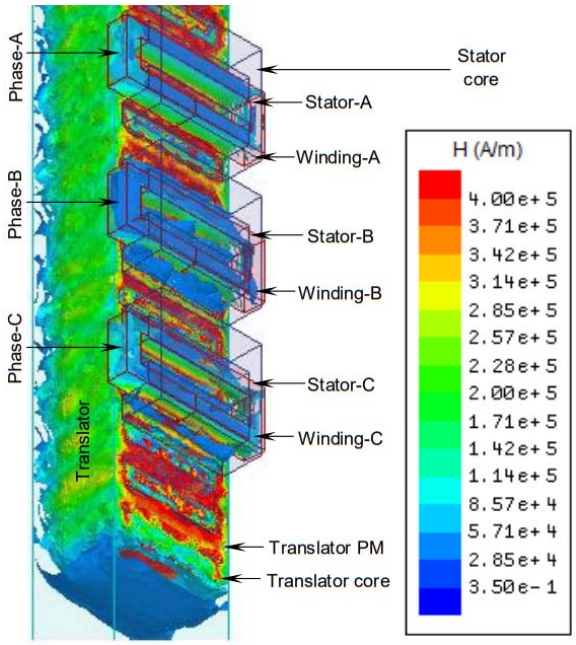


Fig. 9. Field intensity of the proposed linear generator.

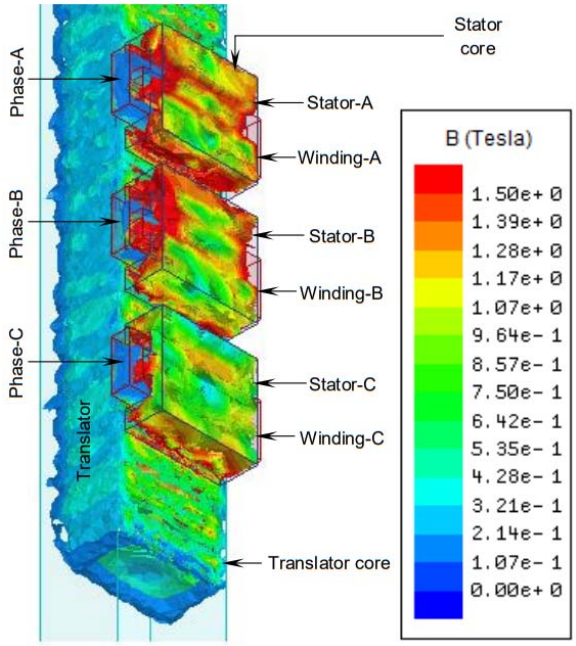


Fig. 10. Flux density of proposed linear generator.

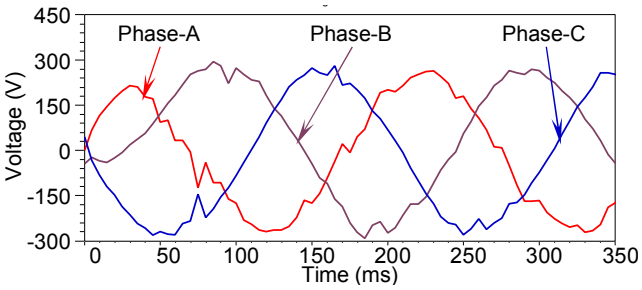


Fig. 11. Output voltages of the proposed DDLG.

The load current of the proposed linear generator is plotted in Fig. 12. The rms value of current found for Phase-A, Phase-B, and Phase-C are 15.56A, 16.3A, and 16.3A, respectively.

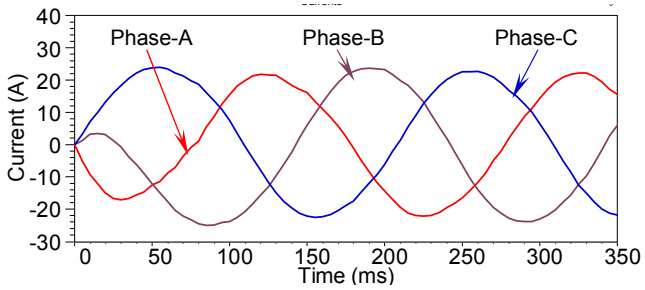


Fig. 12. Load currents of the proposed DDLG.

The flux linkage of proposed linear generator for three phase windings are shown in Fig. 13 and the vertically moving force applied to the translator by the oceanic wave is also plotted in Fig. 14.

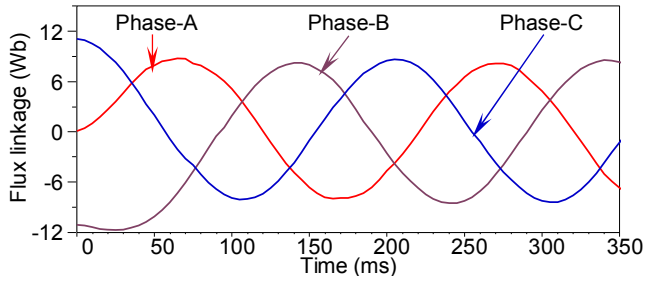


Fig. 13. Flux linkages of the proposed DDLG.

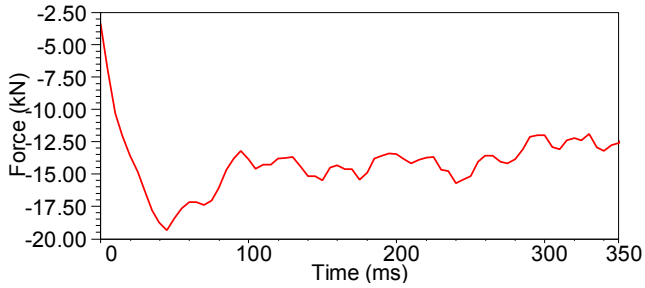


Fig. 14. Applied force to the translator of the proposed DDLG.

The comparison of load currents is found for the two cases as shown in Fig. 15. Out of these, one is conventional and another is proposed cooling system based DDLG. It is seen that, the cooling based generator produces higher current than that of the conventional one.

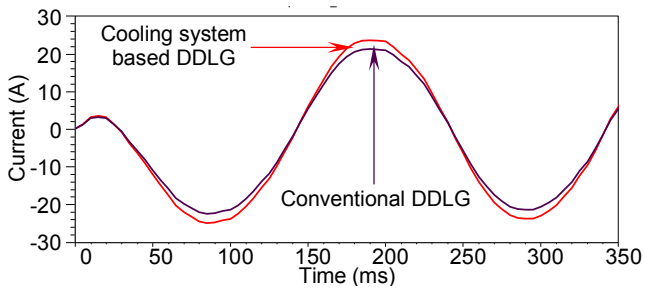


Fig. 15. Comparison of load currents between coventional and cooling system based DDLG.

The comparison of rms powers between the DDLG with and wihout cooling system are plotted in Fig. 16. The generator which uses cooling system produces more electrical power than that of the conventional counterpart.

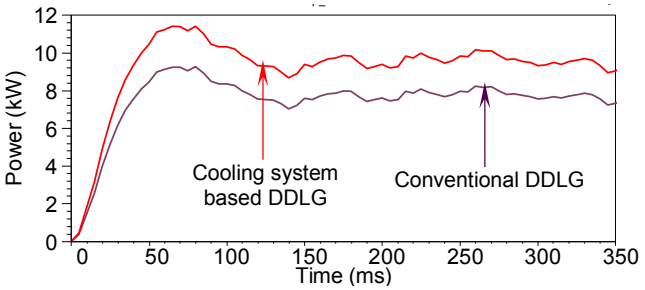


Fig. 16. Comparison of rms powers between the coventional and cooling system based DDLG.

Another comparison of output powers for the two cases is presented in Fig. 17 where the coventional DDLG produces 4755W and the proposed one produces 5870.4W which is more than the conventional one.

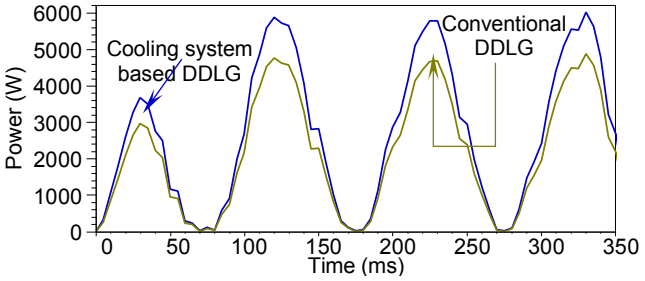


Fig. 17. Comparison of instantaneous output powers between the coventional and cooling system based DDLG.

The comparison of different parameters of the DDLG for two different cases are shown in Table II.

TABLE II. COMPARISONS OF DIFFERENT PARAMETERS

Name of the parameters	Conventional DDLG		Cooling system based DDLG	
	Peak	*RMS*	*Peak*	*RMS*
Current (A)	21.2	15	23.5	16.62
Power (W)	9200	6500	11400	8061

VI. CONCLUSION

The paper introduces an effective cooling system for the DDLG placed in the ocean for electrical power generation. The proposed HVAC system based DDLG is compared with the conventional DDLG. The simulation results show that the cooling system based linear generator produces 2.2kW more electrical power in the output than that of the conventional one. Since heat and humidity inside the linear generator is controlled, it is expected that, the life span and performance would be better for the proposed system.

REFERENCES

[1] L. Huang, H. Yu, M. Hu, J. Zhao, and Z. Cheng, "A novel flux-switching permanent-magnet linear generator for wave energy extraction application," IEEE Trans. Magn., vol. 47, no. 5, pp. 1034–1037, May 2011.

[2] M. A. Kowser, M. T. Islam, M. G. Uddin, T. B. Chakma, and M. Z. R. Chowdhury, "Feasibility study of ocean wave of the Bay of Bengal to generate electricity as a renewable energy with a proposed design of

energy conversion system," Int. J. Renewable Energy Res., vol. 4, no. 2, pp. 445–452, May 2014.

[3] A. I. Chowdhury, S. M. Safiuddin, M. N. Rahman, and A. H. M. Shatil, "Design & analysis of vertical movement ocean wave generator," American J. Eng. Natural Sci., vol. 1, no. 2, pp. 8–13, January 2017.

[4] P. C. Binh *et al.*, "Analysis, design and experiment investigation of a novel wave energy converter," IET Generation, Transmission Distribution, vol. 10, pp. 460–469, February 2016.

[5] D. Elwood *et al.*, "Design, construction, and ocean testing of a taut-moored dual-body wave energy converter with a linear generator power take-off," Renewable Energy, vol. 35, pp. 348–354, February 2010.

[6] R. Krishna, O. Svensson, M. Rahm, S. K. Kottayil, and R. Waters, "Analysis of linear wave power generator model with real sea experimental results," IET Renewable Power Generation, vol. 7, pp. 574–581, September 2013.

[7] S. R. Huang *et al.*, "Multivariable direct-drive linear generators for wave energy," Appl. Energy, vol. 100, pp. 112–117, December 2012.

[8] V. N. Phung, The Mich Nguyen, The Ba Dang, and D. T. Phan, "Numerical simulation of a wave energy converter using linear generator," Vietnam Symposium on Advances in Offshore Engineering, doi. 10.1007/978-981-13-2306-5_44.

[9] M. Trapanese *et al.*, "Minimization of detent force in a 1 kW linear permanent magnet generator for the conversion of sea waves energy: Numerical and experimental validation," IEEE Int. Magn. Conf., 11–15 May 2015, Beijing, China, pp. 1–1.

[10] H. Jing, N. Maki, T. Ida, and M. Izumi, "Electromechanical design of an MW class wave energy converter with an HTS tubular linear generator," IEEE Trans. Appl. Supercond., vol. 28, no. 4, Art. ID 4902504, June 2018.

[11] S. Molla and O. Farrok, "Vitroperm 500F and Supermendur ferromagnetic cores used in a linear generator for oceanic wave energy conversion," *Int. Conf. Robot., Elect. Signal Process. Techn.*, 10–12 Jan. 2019, Dhaka, Bangladesh, pp. 602–605.

[12] M. S. Bashir and O. Farrok, "Harvesting oceanic wave energy by a linear generator using high graded N28EH permanent magnets," Int. Conf. Elect. Eng. Inform. Commun. Technol., 13-15 September 2018, Dhaka, Bangladesh, pp. 514–518.

[13] S. Molla, O. Farrok, M. R. Islam, and K. M. Muttaqi, "Analysis and design of a high performance linear generator with high grade magnetic cores and high temperature superconducting coils for oceanic wave energy conversion," IEEE Trans. Appl. Supercond., vol. 29, no. 2, Art. ID 5201105, March 2019.

[14] M. S. Bashir and O. Farrok, "Yttrium barium copper oxide superconductor used in a linear generator for high power generation from the oceanic wave," *Int. Conf. Elect., Comput. Commun. Eng.*, 7–9 Feb. 2019, Coxbazar, Bangladesh, pp. 1–5.

[15] W. N. L. W. Mahadi, S. R. Adi, and K. M. Nor, "Thermal analysis of a tubular permanent magnet linear generator using multiphysics solver," IEEE Region 10 Conf., 21–24 November 2005, Melbourne, Qld., Australia, pp. 1–6.

Generation of Electrical Power by using High Graded Permanent Magnet Linear Generator in Wave Energy Conversion

Md. Samiul Bashir and Omar Farrok
Department of Electrical and Electronic Engineering
Ahsanullah University of Science & Technology
Dhaka 1208, Bangladesh
afcl.engr@gmail.com and omarruet@gmail.com

***Abstract*—The conventional permanent magnet linear generator (PMLG) used for oceanic wave energy extraction are made of ordinary graded neodymium iron boron (NdFeB) permanent magnets (PMs). As a result, while the temperature of the generator gradually increases because of continuous operation, the generation of electrical power would be low because of demagnetization. In this paper, it is analyzed that, the conventional PMLG can not produce much electricity at high temperature because of using the ordinary graded NdFeB PMs. On the other hand, NdFeB that can maintain its remanence magnetism at high temperature saturates at low magnetic field. Therefore, high graded N48H NdFeB is proposed to be used in the PMLG to overcome this dilemma. Simulation results show that, because of using N48H PMs in the PMLG, the generated voltage, current, flux linkage and output power reaches to higher value than that of using the conventional N35 NdFeB PMs. The proposed PMLG with N48H is able to produce 28% more electrical power output than that of using the conventional N35. Finite element method is used for the simulation platform. During the simulation, same operating temperature and parameters of the PMLG is considered for both the high graded and ordinary one for fare comparisons.**

Keywords— linear generator; magnetic materials; permanent magnet machines; renewable energy sources; temperature; wave power conversion.

I. INTRODUCTION

With the enhanced market of electricity in the world, environment is polluted by carbon dioxide that is produced in fossil fuel-based power generating stations. On the other hand, renewable energy sources such as solar, thermal, wind, and wave energy minimizes the production of carbon dioxide. Wave energy is one of the most promising resources among the others. It conceives with great potentiality of producing clean energy and provides major contribution in alleviating carbon dioxide [1]. The suitable system for the extraction of wave energy is direct-drive wave energy conversion (WEC).

The renewable energy includes conventional solar, wind, or hydro, but, ocean is one of the most promising [2] sources. It is estimated that, 70% oceans in the world are capable of providing 1 TW of electrical power [3]. In another estimation, the wave energy resource is calculated to be 2 TW which is presented in [4]. Wave energy resources are substantial and the total world's resources can provide 10 TW power in the open sea as reported in [5].

The design and testing procedures of wave energy converters using nine concepts are presented in [6] that also calculated energy production of each system. This is on the basis of measured model performances, scaled up to the wave conditions at a specific site in the North Sea. The direct drive linear generator is developed experimentally based on the electromagnetic and electric circuit models for getting higher efficiency of the transverse flux machine in a WEC system as presented in [7]. A direct drive linear generator with superconducting coils are used in [8] for WEC so that, better performance and production of more electrical energy can be achieved. The difference between the conventional and improved transverse flux permanent magnet linear generator (PMLG) for WEC is presented in [9]. The oceanic wave converts mechanical to electrical energy by several steps that is shown in Fig. 1. It depends on different factors such as wave energy, linear generator, converter, inverter, and load etc.

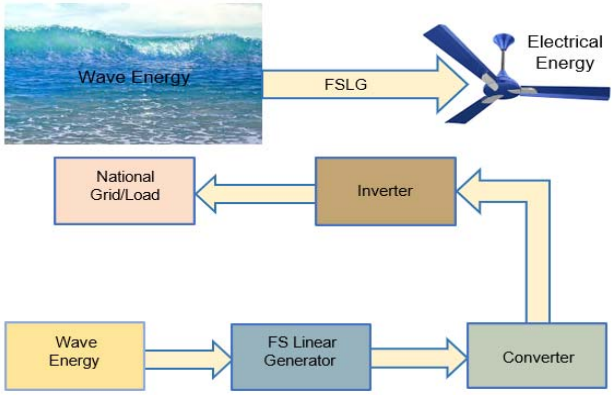


Fig. 1. Steps of the wave energy conversion.

The oceanic wave force is used as the input of a permanent magnet flux switching linear generator (FSLG). The FSLG generates ac output power from the oceanic wave. The generator output is connected to the converter which converts ac power to dc power. The dc power is connected to the inverter which is producing sinusoidal ac power to be transferred to the load or grid. This procedure is called wave energy conversion system.

Two high graded magnetic cores are used in a linear generator with comparisons for oceanic WEC in which Supermendur core exhibits better performance [10]. High graded core used in the PMLG produces less heat energy due to its low core loss profile. The internal resistance of the armature winding is also minimized by using superconductor. A piston driven PMLG is optimized using an automated method of optimization considering the constant or sinusoidal speed of the piston [11]. Due to the simplicity of linear generators, these are ideally suited for the power take off systems for completely submerged WEC system.

The Archimedes wave swing model is used to convert wave energy into electrical energy using a PMLG in addition to a current source inverter so that, about 18% output can be

improved yearly [9]. During electrical power generation by conventional PMLG from the oceanic wave, a temperature rise occurs in the translator and stator cores [12]. The Sm_2TM_{17} type permanent magnet (PM) is improved to meet the infliction at a high temperature which is discussed in [13]. Another linear generator designed by neodymium iron boron (NdFeB) N28EH PM that is proposed in [14] for mitigating the temperature effects on the PMs. A non-oriented DI-MAX HF-10X electrical steel is proposed in [15] as the magnetic of a linear generator.

The direct drive PMLGs are generally connected to a substation and the temperature metering is performed in the offshore location of the WEC array. The generated voltage of the linear generator is rectified and then connected to a common DC-link. With the aid of power inverter, 50 Hz ac voltage is generated and transmitted to shore to facilitate the grid-connection. During the operation, the temperature of the machine and cable increase to 70°C [16]. It is analyzed in [14] that, the performance of permanent magnets depend on temperature and N28EH PM is proposed to be used in PMLGs. The high temperature N28EH NdFeB permanent magnet has a major limitation which is its low remanence even at room temperature [14].

Fig. 2 shows the remnant magnetic flux density (B_{rem}) of N48H which is 1.33 T while operated at 60°C, coercive field strength is −930 kA/m, and knee point is 1.2 T. Again, B_{rem} of N35 is 1.14 T for 60°C, coercive field strength is −690 kA/m, and knee point is 1.06 T. Therefore, at the same temperature of the FSLG, B_{rem} and coercive field strength for N48H is more than that of the conventional N35 NdFeB PM.

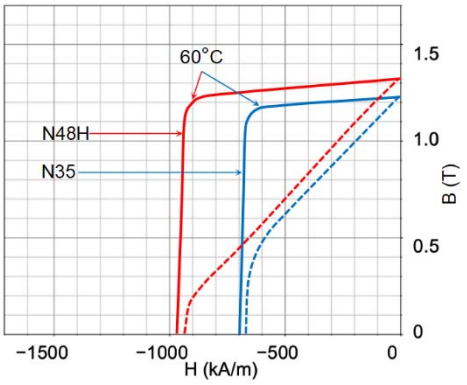


Fig. 2. Demagnetization curve of N48H and N35 PM.

A permanent magnet FSLG is designed and simulated individually for testing the performance of two different PMs N35 and N48H. From the simulation result, it is observed that, using N48H in the FSLG exhibits better result than that of using N35 for the same temperature. Therefore, N48H NdFeB PM is proposed to be used in the FSLG in order to obtain more output power at high temperature as well as prevention of demagnetization.

II. Mathematical Model

Fig. 3 illustrates the operating point for magnetization with the armature for the proposed FSLG where the magnetic permeance is denoted by G_t. The F-coordinate represents the reverse magnetic field intensity. The point K of the secondary hysteresis known as recoil line is realized by the resultant G_t drawn from the point $F'_{ad\max}$. The magnetomotive force (MMF), $F'_{ad\max}$ resembles to the highest demagnetizing d-axis field working directly on the permanent magnet that can arrive during the operation of FSLG. The magnetic permeance can be expressed in the following.

$$G_t = \frac{\phi_K}{F_K - F'_{ad\max}} \tag{1}$$

where, ϕ_k is magnetic flux at k point, $F'_{ad\max}$ is the highest demagnetizing MMF.

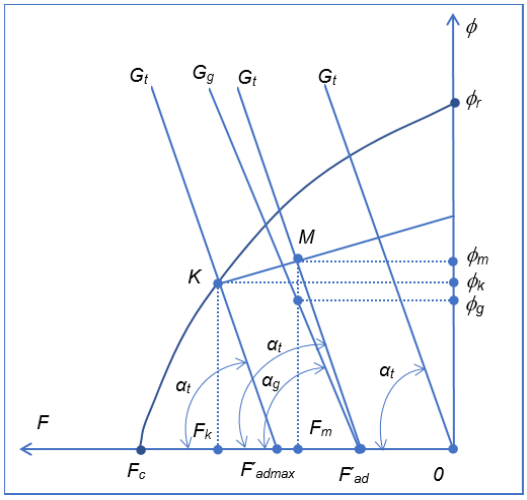


Fig. 3. Operating point for magnetization with the armature for FSLG.

The starting of the recoil line is determined by G_t of the permanent magnet which is mounted in the armature.

If $a_0 > 0$, the MMF corresponding to the point K is

$$F_k = b_0 \pm \sqrt{b_0^2 - c_0} \tag{2}$$

where, F_k is MMF, b_0 and c_0 are constant coefficients

$$\begin{aligned} b_0 &= 0.5\left(\frac{F_c}{a_0} + F'_{ad\max} + \frac{\phi_r}{a_0 G_t}\right) \\ c_0 &= \frac{(G_t F'_{ad\max} + \phi_r)F_c}{a_0 G_t} \end{aligned} \tag{3}$$

If $a_0 = 0$ for rare-earth PMs, the MMF F_k can be expressed as

$$F_k = \frac{\phi_r + G_t F'_{ad\max}}{G_t + \phi_r / F_c} \tag{4}$$

The magnetic flux ϕ_k can be found from (1).

III. Equivalent Magnetic Circuit

Fig. 4 depicts the equivalent magnetic circuit of the proposed permanent magnet FSLG where the open circuit MMF is working along the internal G_t denoted as G_M which can also be expressed by inversed internal reluctance $R_{\mu M}$. In Fig. 4, F_{ad} is the d-axis armature reaction MMF, ϕ_g is the useful magnetic flux in the air gap length, ϕ_{la} is the leakage flux of the external armature winding, ϕ_{ad} is the magnetic flux

produced by the armature winding, G_{lM} is the permeance for leakage flux because of the permanent magnet, $G_{\mu M}$ is the internal permeance, $G_{\mu g}$ is the air gap permeance, G_{la} is the external armature leakage permeance.

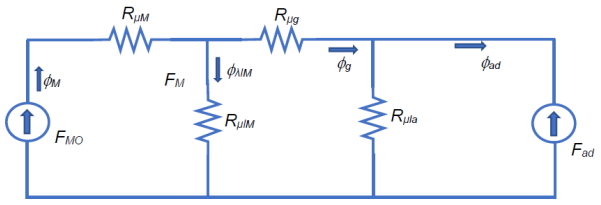


Fig. 4. Equivalent magnetic circuit of proposed FSLG.

According to the equivalent magnetic circuit shown in Fig. 4, the following equations can be written.

$$\begin{aligned} \phi_M &= \phi_{lM} + \phi_g \\ \phi_{la} &= \frac{\pm F_{ad}}{R_{\mu la}} \end{aligned} \tag{5}$$

where, ϕ_M is the total flux, ϕ_{lM} is the leakage flux, ϕ_{la} is the leakage flux of external armature, and $R_{\mu la}$ is armature leakage reluctance. The magnetic flux in the air gap length ϕ_g of the permanent magnet can be expressed in the following.

$$\phi_g = \left[F_{mo} \pm F_{ad} \frac{G_g (G_g + G_{lM})(G_M + G_{lM})}{(G_g + G_{lM})(G_g G_M)} \right] \left[\frac{G_g G_M}{G_g + G_{lM} + G_M} \right] \tag{6}$$

where, G_g is useful permeance and G_{lM} is leakage permeance. The full equivalent circuit of the proposed FSLG is shown in Fig. 5 where the field resistance R_F is in series with the adjustable resistor R_{adj} for controlling the field current. As the proposed FSLG is a permanent magnet machine, the equivalent circuit for field winding is replaced with a permanent magnet for production of the magnetic flux. The equivalent circuit at the right hand margin represents the model for three phase.

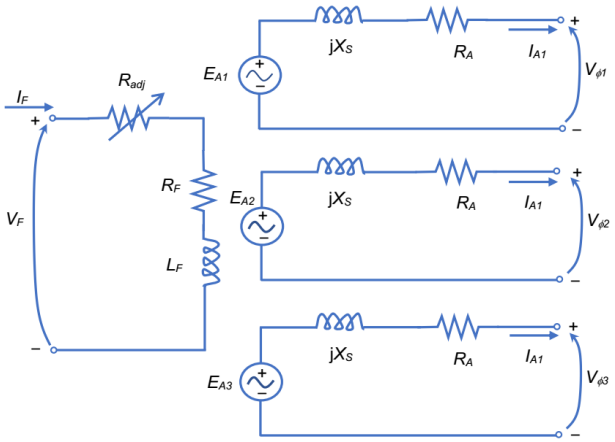


Fig. 5. Equivalent circuit of the proposed FSLG.

IV. Construction and Magnetic Properties

The construction of the proposed FSLG is shown in Fig. 6 including its different elements. The main parts of the machine are translator which is made of magnetic core and a stator with copper coil. Another essential component is NdFeB permanent magnet. The magnetic property of NdFeB used in the FSLG for different temperatures are shown in Fig. 2. As the buoy moves up and down with the oceanic wave, the translator also moves vertically with respect to stators. There is a relative motion between the translator and stators. The permanent magnet is installed inside the copper winding and stator pair. Because of relative motion by Farady's law, electromotive force is induced in the copper winding due to occur electromagnetic induction.

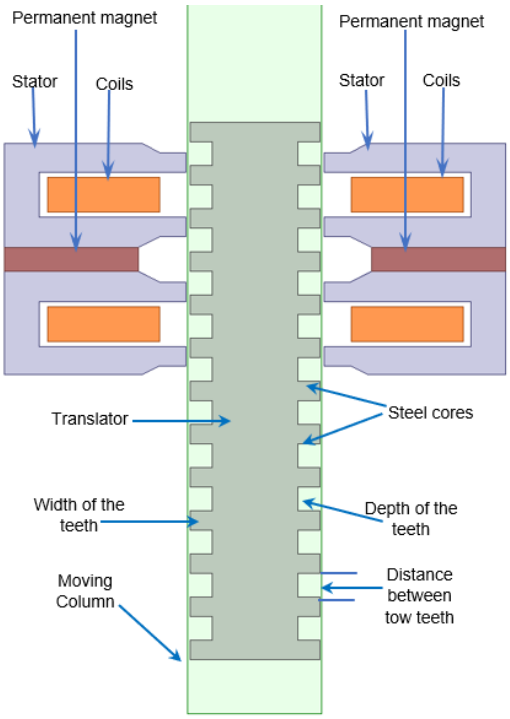


Fig. 6. Construction of the proposed FSLG.

Fig. 7 illustrates the magnetic field distribution of the FSLG for using the proposed N48H. The blue color represents zero magnetic field intensity, H(A/m), green represents the average, and red represents the maximum magnetic field intensity. Also, the scale of H appears in Fig. 7 as rainbow spectra.

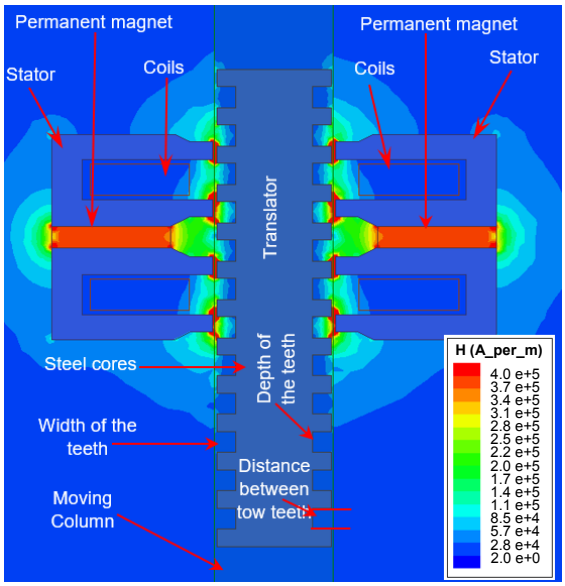


Fig. 7. Magnetic field intensity of FSLG using N48H.

The permanent magnet is placed on both sides between the stator pares in the FSLG. The translator is encompassed with the stators maintaining air gap. The moving column is situated along the transalator which represents the way of translator motion. Mesh plot of the FSLG is shown in Fig. 8 whereas Fig. 9 illustrates magnetic flux lines of the proposed FSLG. The rainbow spectra present the scale of magnetic flux lines in the surrounding area. The red and blue colors represent the maximum and minimum values of the magnetic flux.

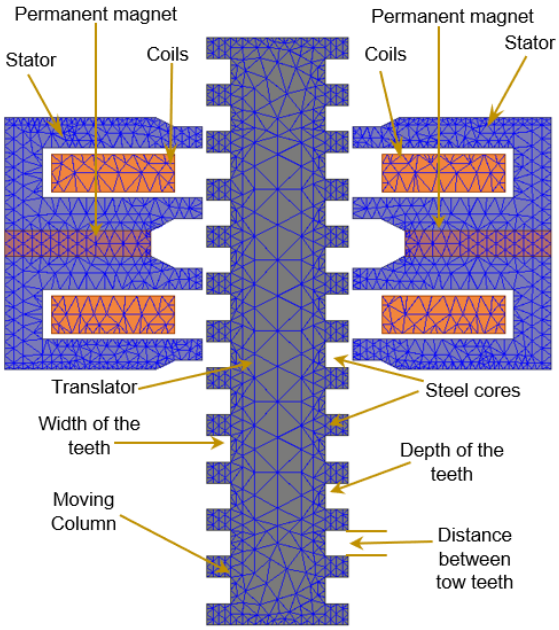


Fig. 8. Mesh plot of the of the FSLG using N48H.

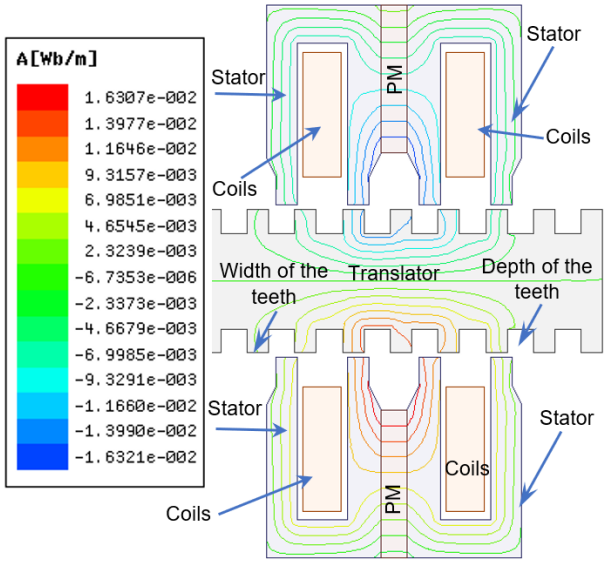


Fig. 9. Magnetic flux lines of the proposed FSLG.

Fig. 10 shows magnetic flux density, $B(T)$ of the FSLG using N48H where the zero magnetic flux density is represented by blue color. Green color represents average magnetic flux density. The scale of B is shown by rainbow spectra in Fig. 10.

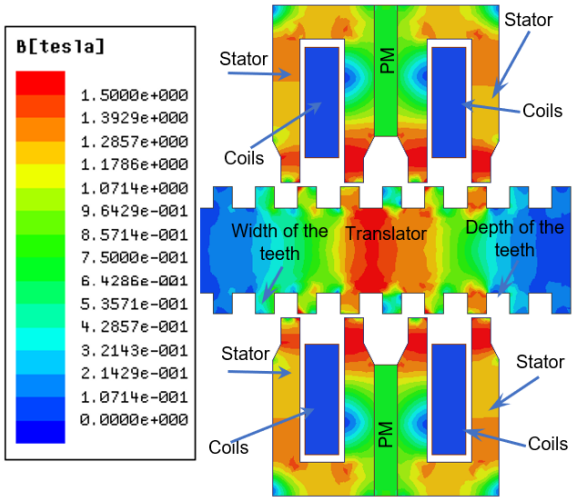


Fig. 10. Magnetic flux density of the FSLG using N48H.

V. Simulation Results

The translator speed is selected 1m/s with an air gap length of 1 mm. The thickness of the permanent magnet is 5.6 mm turns of the coils is 210. The load (resistive) of this proposed FSLG 10 Ω. Fig. 11 shows the voltage comparison for using the proposed N48H and conventional N35 permanent magnets.

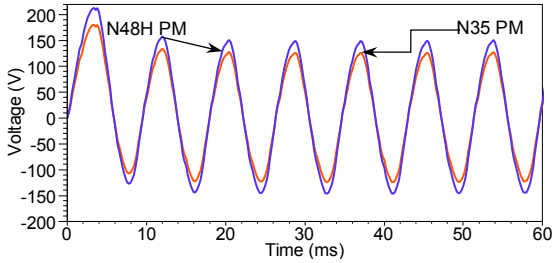


Fig. 11. Voltage waveform of the proposed FSLG.

In Fig. 12, current comparison of the FSLG with the proposed and conventional permanent magnet is shown. The load currents are plotted with respect to time. The red color presents load current for using N48H permanent magnet and green color represents the same for using N35. The peak value of current is near about 1 A for using N48H which is more than that of using N35. Fig. 13 shows the force (N) working on the moving translator. The purple and pink colors represent the horizontal and vertical force components of the translator, respectively.

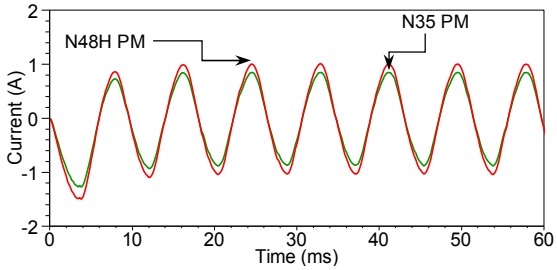


Fig. 12. Current waveform of the FSLG.

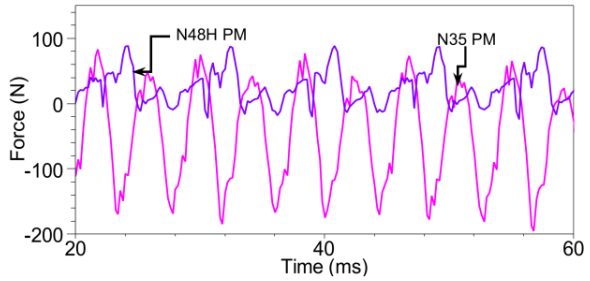


Fig. 13. Moving force of the proposed FSLG.

In Fig. 14, the core loss profile of the proposed FSLG is shown with respect to time. Fig. 15 plots the flux linkage of the proposed FSLG. Fig. 16 shows the comparisons between the instantaneous output power of the FSLG for using the proposed N48H and the conventional N35 permanent magnets. The output electrical power of FSLG for using N48H is represented by red color and blue color indicates the same for using N35. During operation of the FSLG, the temperature gradually increases and the output powers are plotted for the same temperature. Comparison shows that, the FSLG with N48H permanent magnet generates more electrical power than that of using the conventional one. In Fig. 16, the peak value of the power output is 112 W for using the conventional permanent magnets as indicated by blue color. On the other hand, the peak value of the power output is 156 W for using the proposed one as sketched by red color.

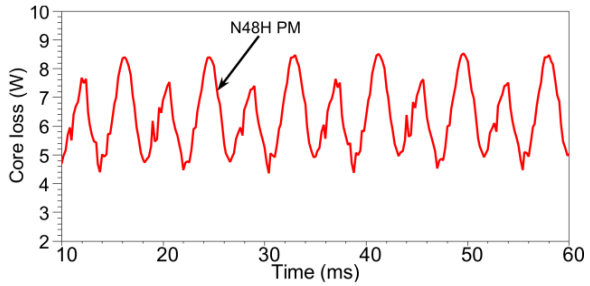


Fig. 14. Core loss of the proposed FSLG using N48H PM.

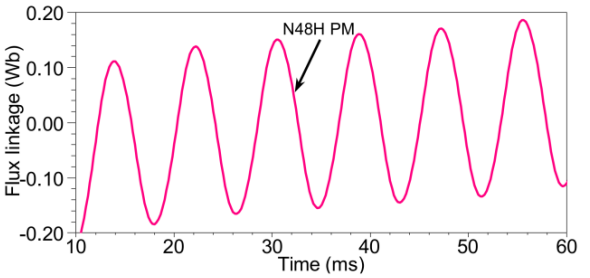


Fig. 15. Flux linkage of the proposed FSLG.

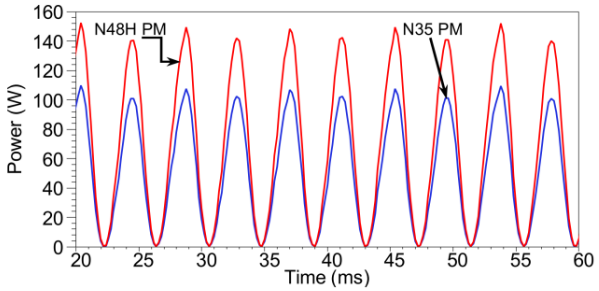


Fig. 16. Comparisons of power outputs of the proposed FSLG.

VI. Conclusion

From the simulation results it is shown that, the permanent magnet FSLG using the conventional grade N35 NdFeB permanent magnet produces less amount of electrical power compared to that of using the high graded N48H. Because, N48H can maintain the high magnetic remanence and large coercive force even at high temperature. The same FSLG can generate 44W more electrical power because of using N48H instead of N35. The operating temperature of the FSLG is considered 60° C for using both the conventional and the proposed NdFeB PMs. Simulation work is executed by using the software ANSYS/Maxwell.

References

[1] J. Falnes, "Ocean waves and oscillating systems," *Cambridge, U. K.: Cambridge Univ. Press*, 2002, pp. 196–260.

[2] A. Ramadan, *et al*, "Analytical investigation and experimental validation of an inverted cup float used for wave energy conversion," *Energy*, vol.70, pp. 539−546. Jun 2014.

[3] T. W. Thorpe, "The wave energy programme in the UK and the European wave energy network," *in Proc. Int. Conf. Wave Energy, Aalborg, Denmark*, 2000, pp. 19–27.

[4] T. W. Thorpe, "A brief review of wave energy," London, U. K.: Harwell Laboratory, *Energy Technology Support Unit*, 1999.

[5] D. Ross, "Power from the waves," in Oxford, U. K.: *Oxford Univ. Press*, 1995.

[6] N. I. Meyer and K. Nielsen, "The danish wave energy programme second year status," *in Proc. 4th Wave Energy Conf.*, Aalborg, 2000, pp. 10–18.

[7] M. A. Mueller, "Electrical generators for direct drive wave energy converters," *in Proc. Inst. Elect. Eng. Generation, Transmission and Distribution*, vol. 149, 2002, pp. 446–456.

[8] M. S. Bashir and O. Farrok, "Yttrium barium copper oxide superconductor used in a linear generator for high power generation from the oceanic wave," *Int. Conf. Elect., Comput. Commun. Eng.*, 7–9 Feb. 2019, Coxbazar, Bangladesh, pp. 1–5.

[9] H. Polinder, M. Damen, and F. Gardner, "Linear PM generator system for wave energy conversion in the AWS," *IEEE Trans. Energy Convers.*, vol. 19, no. 3, pp. 583–589, Sep. 2004.

[10] S. Molla and O. Farrok, "Vitroperm 500F and Supermendur ferromagnetic cores used in a linear generator for oceanic wave energy conversion," *Int. Conf. Robot., Elect. Signal Process. Techn.*, 10–12 Jan. 2019, Dhaka, Bangladesh, pp. 602 – 605.

[11] A. Wolfbrandt, "Automated design of a linear generator for wave energy converters—a simplified model," *IEEE Trans. Magnetics*, vol. 42, no. 7, pp. 1812−1819, Jul. 2006.

[12] W. N. L. W. Mahadi, S. R. Adi, and K. M. Nor, "Thermal analysis of neodymium iron born (NdFeB) magnet in the linear generator design," *Australasian Universities Power Eng. Conf.* (AUPEC 2004), Brisbane, Australia, Sep. 2004.

[13] C. H. Chen, *et al*, "New series of Sm_2TM_{17} magnet materials for applications at temperature up to 550°C," MRS Symposia Proceedings, *Advanced Hard Soft Magn. Materials*, vol. 577, pp. 277–287, 1999.

[14] M. S. Bashir and O. Farrok, "Harvesting oceanic wave energy by a linear generator using high graded N28EH permanent magnets," *Int. Conf. Elect. Eng. Inform. Commun. Technol.*, 13–15 Sept. 2018, Dhaka, Bangladesh, pp. 514–518.

[15] S. Molla and O. Farrok, "Use of high flux density ferromagnetic cores in linear generators for oceanic wave energy conversion," *Int. Conf. Elect., Comput. Commun. Eng.*, 7–9 Feb. 2019, Coxbazar, Bangladesh, pp. 1–5.

[16] C. Boström, *et al*, "Temperature measurement a linear generator and marine substation for wave power", *29th Int. Conf. Ocean, Offshore Arctic Eng*. OMAE2010 June 6-11, 2010, Shanghai, China.

Improved Bangla Language Modeling with Convolution

Shuvendu Roy
dept. of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
bikash.shuvendu@gmail.com

***Abstract*—Language modeling is a fundamental task for building any natural language processing application or language understandable intelligent system. Language modeling becomes difficult with the increase of input sequence length. It is even more difficult for Bangla as it is more complex than many other languages and contains more character in its alphabet. Few previous works that were done on Bangla language modeling used statistical language modeling or simple neural network very recently. With the neural network, the RNN is the fundamental choice to deal with sequential data but we follow the recent development in convolutional neural network in sequential modeling. In this work, we used dilated convolution to solve the issues of long sequence and improved the performance of Bangla language modeling.**

***Index Terms*—Language-Modeling, CNN, Dilated-Convolution**

I. Introduction

Language modeling is central and fundamental to many other natural language application. So, it is important to improve the language modeling first to get better performance in many other applications. Language modeling is the task of predicting next word or character or any token in a sequence. Bangla being a complex language compared to other pauses some challenges to model. Recently, neural-network-based language models have shown better performance compared to statistical language modeling. The input to the language model being sequential, the main choice for language modeling is Recurrent Neural Network. But the RNN and its variances have limitations dealing with very long sequences. In this work, we have proposed the use of Convolution to improve the Bangla language modeling and the experiments compare the performance with other previous approaches.

Recurrent neural network(RNN) is a kind of neural network that is used with sequential data. For language modeling task the input data is sequential and variable length. That is way RNN and some other variations of RNN were used for language modeling task. One problem with vanilla RNN is vanishing gradient. When the sequence of the input becomes larger it becomes hard for the model to propagate the gradient from the last time step to the initial time step. As a result, the model does not learn well for the earlier time step. To reduce the vanishing gradient problem few variations of RNN was proposed. Among them LSTM [1] and GRU [2] are the most popular. Even with these models input with very long sequences suffers from the same problem.

One other model called Convolutional Neural Network(CNN) [3] is mainly used for fixed size input like an image. This model does not have issued with vanishing gradient but it is not proper for sequential data. However recent research has found that certain varieties of CNN can reach state-of-the-art-result with sequential data. For our application of language modeling, we are not allowed to see the input after time step $(t + 1)$ while outputting for time step t. For this, we have used casual convolutions [4] which does not see the input after time step t. To see the input from very far and still avoiding the vanishing gradient problem van den Oord et. al. [5] proposed Dilated Convolution architecture. In this word, we used the dilated convolution architecture for improving the Bangla Language Modeling task.

II. Related work

Language modeling with the neural network was first proposed by Bengio et. al. [6] which was very successful compared to classical language modeling. later with the improvement of neural network and deep learning, the performance of language modeling with neural network kept improving. Not many works are available on Bangla language modeling in particular. Recently a word level stochastic language modeling for Bangla was proposed by Haque et. al. [7].

The concept of using regular CNN for sequential application is not new but it is showing good result only recently because of some variations of CNN. The use of diluted convolution in wavenet [5] gave convolution the way to look much more time step back than the result convolution. They used it for wave synthesis purpose. The convolution was used for language modeling by Dauphin et. al. [8]. Later Bai et. al. [4] introduced causal convolution with diluted convolution.

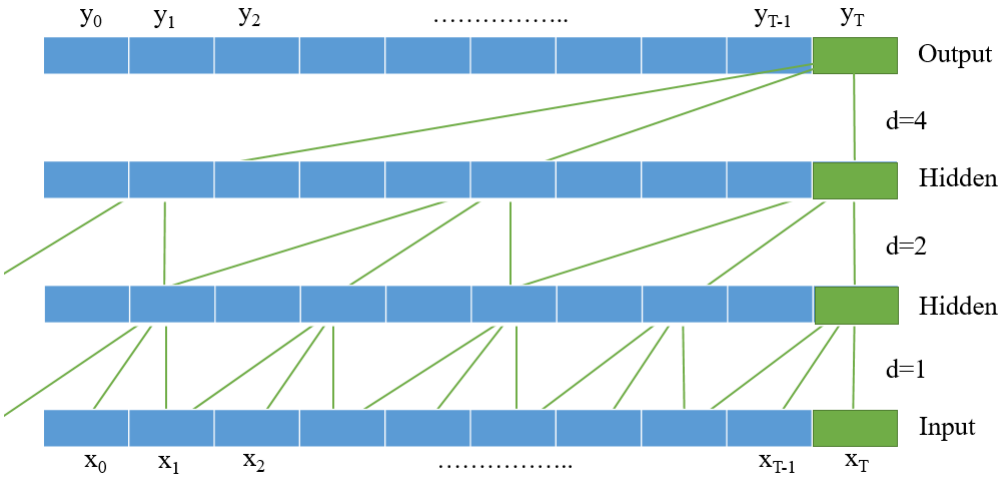


Fig. 1. A dilated causal convolution that has dilation factors of d = 1, 2, 4 with kernel size k=3. The entire receptive field of the dilated convolution covers all the input values in the sequence.

III. Methods

We used dilated convolution with causal convolution for modeling the sequence. The two main feature of the model is- 1) there is no information leakage from future to present or past time step. This is important for language modeling. In the case of regular RNN, we process one input from the sequence at a time so there is no scope for getting the input of future time step. 2) The architecture can take input sequence of any length just like the regular RNN with the help of dilated convolution operation.

A. Sequence modeling

Before defining the model architecture of formulating the language modeling task. Suppose the input sequence is $x_0, x_1, ..., x_T$, and we want to predict the corresponding output $y_1, y_2, ..., y_T$. For one time step, we are given all previous input and we want to predict y_T. The restriction here is that we are not allowed to see the input after the current step, which is this case is $x_0, x_1, ..., x_T$. Formally, we can define the task as $f : X^{T+1} \rightarrow Y^{T+1}$ which can be represented as

$$y_1, y_2, ..., y_T = f(x_0, x_1, ..., x_T) \tag{1}$$

For a language model by setisfying the constrain the current output y_T depends on only $x_0, x_1, ..., x_T$ and not $x_{T+1}, x_{T+2}, ...$ Then the task of laguage modeling is to find the function f to minimize the loss between the predicted and the actual output with a given loss L. The loss can be represented as $L(y_1, y_2, ..., y_T, f(x_0, x_1, ..., x_T))$.

B. Network Architecture

Our network has 3 main components. 1)causal Convolutions, 2) Dilated Convolution, and 3) Residual Connection. Following sections describe these components in details.

1) causal Convolutions: As mentioned above, our model is based on causal convolution. The is the model is not allowed to see the future input while processing the current output. And the other properties of the model is that the input and output size of the model is the same. For this, we used a 1D fully-convolutional neural network with zero padding of size $K - 1$, where K is the kernel size. Here the length of the hidden layer and the input layer is also the same. This part of the network is enough to generate the desired output. But the major disadvantage of this model is that for modeling such big input it will require a massive neural network which again will be hard to train. To solve this issue we used dilated Convolution instead of regular convolution.

2) Dilated Convolutions: A regular causal convolution can look back in history only with a size linear to the depth of the network. This makes it harder for such network to process the input data long back into the sequence. To solve this issue we follow the architecture of van den Oord et. al. [5] on diluted convolution. The idea to employ exponentially large receptive field. Mathematically, let x be a 1-D sequence of $x \in \mathbb{R}^n$ and a filter $f : \{0, 1, ...(k-1)\} \rightarrow \mathbb{R}$, s be a sequence on which the dilated convolution F is defined as

$$F(s) = (x *_d f)(S) = \sum_{n=1}^{k-1} f(i).X_{s-d.i} \tag{2}$$

where k is filter size, d is the dilation factor which is equivalent same as adding a fixed step between two adjacent taps. When d=1, the modified dilated conv. is the same as regular convolution. Here, $s - d.i$ directs to the past. The output in the top layer can include information from a wider range of input just by increasing the receptive field.

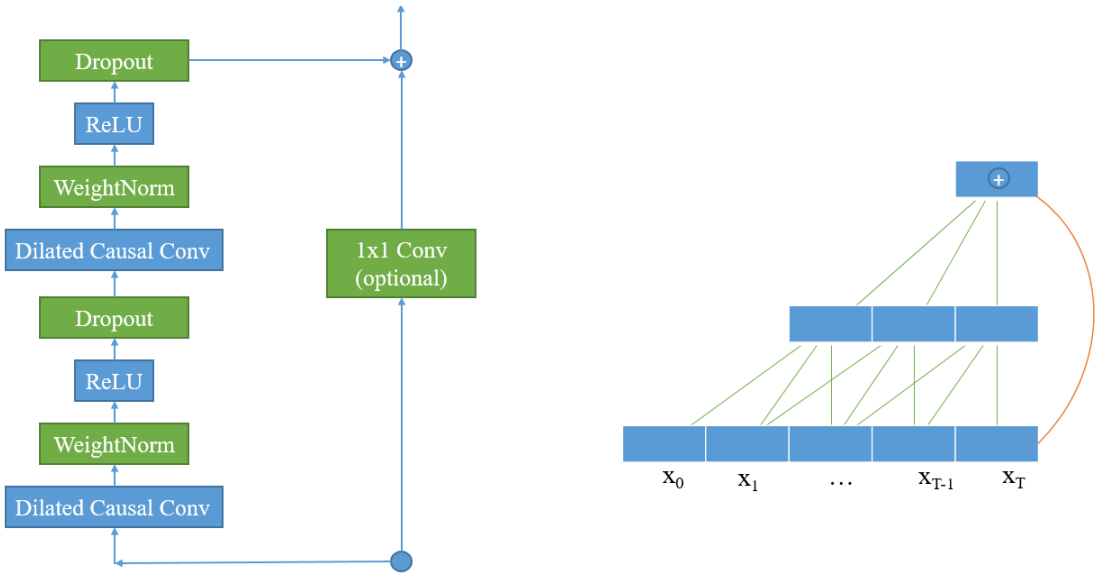


Fig. 2. A residual block.

We can increase the receptive field in 2 ways in this model. 1) by increasing the filter size *k*, 2) by subsequently increasing the dilation factor *d*. The effective history the output can see in one layer is *(k-1)d*. An efficient strategy to increase the dilation factor is to increase the value of d exponentially(i.e. d = O(2^i), i is the level of the network). Fig. 1 shows an illustration of dilated convolution operation.

3) Residual Connection: The concept of the residual block was introduced by He et. al. [9]. This is an important concept in deep learning and triggered a new era of training even deeper models possible. The residual block made many very deep CNN model for various computer vision applications possible. As our model is derived from the concepts of CNN it is used as an important concept for training our model.

The idea is to pass the input to a layer of the model and add with the output of that layer or some sequence of layers. If x is the input to a sequence of layers F, a residual operation will be

$$output = Activation(x + F(x)) \tag{3}$$

Fig. 2 shows one block of residual network.

Such residual connection helps the layer to learn the change in the transformation. Also, it helps to prevent over fittings and train bigger network.

IV. Experiments

We have implemented the model with pytorch [10] int python. All experiments were done in NVIDIA 1060 GPU.

A. Dataset

To compare the effectiveness of the model we trained the model on both English and Bangla dataset. Language modeling is a popular application and many research has used some common dataset for this task. We used text8 dataset [11] for character level language modeling. It contains about 100M characters from Wikipedia (90M in the train, 5M in the test, and 5M in validation set). It is a cleaned 27 unique character dataset.

As there is not much research done on Bangla language modeling and no common dataset was used we collected our own dataset following the source of text8. Training the language modeling requires a lot of plain data. So, we collected wiki-dump, which is XML format data of many wiki page combined. We used the total of 350MB raw data. To process the data we first removed the XML tags to collect the raw text then find out all the unique characters. The corpus had more than 11k unique characters and symbol. We removed all characters except Bangla alphabet. The resulting cleaned dataset is more than 100MB and contains 130M characters.

B. Experimental Setup

To train this model we used SGD optimizer with an initial learning rate of 4.0 and decayed it overtime with a decay rate of 0.95 and trained the model for 100 epochs. We used a kernel size of 3 in convolution. Used total 3 levels of convolution.

For the input data, we used the maximum sequence length of length 400 and the input character was embedded into the embedding dimension of 100. The number of hidden dimensions was 450.

All the training parameters of the model are summarized in table I.

TABLE I
VALUES OF TRAINING PARAMETERS

Parameters	Values
Kernel size	3
Levels	3
Input sequence length	400
Embedding dimension	100
Optimizer	SGD
Learning rate	4.0
LR Decay rate	0.95
Epochs	100
Batch size	32

C. Result Analysis

We trained our model with English and Bangla dataset and reported the performance in Bits per characters(BPC), which is a common measurement for language modeling and reported by many papers on this topic. First, we conducted the experiment of Vanilla RNN with English dataset. This is the most common and bottom lime method to compare the result. We got the bpc score of 1.69 with this. We got some improvement with the popular variations of RNN called LSTM and GRU with 1.50 and 1.43 bpc. But our dilated convolution method outperforms all this RNN method with a good margin of 1.45 bpc.

The table below summarizes the experimental result on English(text8) dataset-

TABLE II
COMPARISON OF THE PROPOSED METHOD WITH PREVIOUS APPROACHES ON ENGLISH

Experiment	bpc
Vanilla RNN	1.69
GRU [2]	1.53
LSTM [1]	1.50
Proposed method	1.45

Previously known best performing Bangla language modeling used LSTM. We trained our model on LSTM and dilated convolutional neural network.

TABLE III
COMPARISON OF THE PROPOSED METHOD WITH PREVIOUS APPROACHES ON BANGLA

Experiment	bpc
LSTM	1.80
Proposed method	1.56

As we can see from the table above our proposed method outperformed the LSTM model in a very big margin for Bangla dataset.

V. CONCLUSION

Language modeling is a fundamental part of any further development of a model for any language. This is also a benchmark for any sequence modeling architecture. In this work, we have used the dilated convolution to improve the Bangla language modeling task. The experiment has shown a big improvement from the previous benchmark for Bangla language modeling. This work will help further development of Bangla language in other applications.

REFERENCES

[1] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[2] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[3] Y. LeCun, Y. Bengio *et al.*, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.

[4] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv preprint arXiv:1803.01271*, 2018.

[5] A. Van Den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. W. Senior, and K. Kavukcuoglu, "Wavenet: A generative model for raw audio." in *SSW*, 2016, p. 125.

[6] Y. Bengio, R. Ducharme, P. Vincent, and C. Jauvin, "A neural probabilistic language model," *Journal of machine learning research*, vol. 3, no. Feb, pp. 1137–1155, 2003.

[7] M. Haque, M. Habib, M. Rahman *et al.*, "Automated word prediction in bangla language using stochastic language models," *arXiv preprint arXiv:1602.07803*, 2016.

[8] Y. N. Dauphin, A. Fan, M. Auli, and D. Grangier, "Language modeling with gated convolutional networks," *arXiv preprint arXiv:1612.08083*, 2016.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[10] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," 2017.

[11] T. Mikolov, I. Sutskever, A. Deoras, H.-S. Le, S. Kombrink, and J. Cernocky, "Subword language modeling with neural networks," *preprint (http://www. fit. vutbr. cz/imikolov/rnnlm/char. pdf)*, vol. 8, 2012.

Denoising Sequence-to-Sequence Modeling for Removing Spelling Mistakes

Shuvendu Roy
dept. of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
bikash.shuvendu@gmail.com

***Abstract*—Rule-based spelling correction system focused on finding the most matched word with the misspelled word. But this approach does not work well inside a sentence with multiple errors that has a combination of possible correct words to replace but only one current sentence. Replacing each word individually will result in errors. So, the spelling corrector system must understand the context of the sentence including the tense and gender of the subject and so on. The most popular example of typing mistake correction is the one Google provides in their search engine. It was introduced quite a while ago but no such good performing system is developed by anyone else. In this work, we have proposed a spelling correction system using deep learning. The basic intuition of our approach is taken from denoising autoencoder. Here we have trained the model with noisy input generated by changing, removing or adding extra character at random position inside the sequence. The job of the model is to model this noisy input to output the original errorless sequence. We have experimented with large English dataset and reported the performance in terms of character level accuracy. The proposed model has shown impressive results in correcting the spelling mistakes.**

***Index Terms*—Sequence modeling, Attention, Spelling correction**

I. Introduction

Typing mistake is very common for most people. Google introduced their famous algorithm for suggesting the right sentence in spelling mistake is there in the search query. Their spelling correction feature was unique and many search engine still does not have such facilities. The ”Did you mean?” algorithm was never fully disclosed. Not many research paper has solved this issue or produced any such result. The time it was introduced the deep learning was not that powerful and the algorithm used classical machine learning and natural language processing for this.

The recent success of deep learning has achieved the state-of-the-art result in many natural language processing application. Most of the application of NLP works with sequential data. In deep learning, the model of choice for working with sequential data is the recurrent neural network(RNN). The RNN module takes one token of the sequence of inputs at a time. Each of this is called one time-step. The model unfolds overtime to take the whole input sequence as input. At each time step, the RNN generates an output and a hidden state. The hidden state is passed to the next time step. This hidden state is responsible for encoding all the information it has received in all previous step. This operation is found to be inefficient and prone to losing information after few time-step. This problem is known as the vanishing gradient. To reduce this problem few variations of RNN were proposed. The most common are LSTM [1] and GRU [2].

In this work, we have proposed a neural network based method making a spelling corrector system. The model is trained on plain English text. The input to the model is noisy text, e.i. deliberately introduced errors into the sentence. The output is the original errorless sentence. Both input and the output of the model are sequences. The type of model that deals with this kind of input is commonly known as sequence-to-sequence model in the literature. These models are encoder-decoder based. The encoder takes all the input in the sequence and generates a bottleneck vector. The decoder takes this vector as input and generates one token of output at each time step. The main disadvantage of this kind of model is the bottleneck vector. One single vector often does not contain enough information to generate the entire output. Recently a new concept introduced in computer vision called ”Attention” has changed the approach. The attention operation is responsible for looking over the entire input sequence and give importance to the important section for the current output. This is deployed with the previous LSTM based encoder-decoder system. Later Vaswani et. al. [3] showed that attention alone with a simple feed-forward connection is enough to model complex long sequence.

We have taken the intuition of denoising sequence-to-sequence modeling from denoising auto encode in computer vision. Here the autoencoder is trained with a noisy image to reconstruct the original image. The model is consist of encoder-decoder architecture. The image is transformed into a low dimensional matrix which is upsampled by the decoder to reconstruct the original image. This method can reconstruct good quality images. Following this intuition, we have trained a sequence-to-sequence model with noisy input to reconstruct the original sentence. The noisy input is generated randomly to resemble all types of mistakes someone can make during spelling.

The central objective of the model is to use a sequence to sequence model for learning the sentence representation in character level. Until recently, for any sequence modeling task, the popular choice model involves recurrent neural network(RNN). But a recently proposed method called Annotated Transfer [3] has introduced attention based linear modules for sequence-to-sequence modeling and demonstrated on machine translation getting a significant improvement over the previous state-of-the-art result on machine translation. In this work, we have taken the idea from Annotated Transformer model and modified the objective to create the spelling corrector.

II. Methods

Sequence-to-sequence models are consist of encoder-decoder structure [4]. The task of the encoder is to encode the input sequence $(x_0, x_1, ..., x_n)$ into a intermediate sequence $(z_0, z_1, ..., z_n)$ or in some case a single vector z'. Given z as input to the decoder, the decoder then generates the output sequence $(y_0, y_1, ..., y_n)$, one symbol at each time step. Our model is also consist of encoder and decoder but none of those use any RNN block. The main building blocks of our model are stacked self-attention and point-wise fully connected layer. Fig. 1 shows the entire architecture of the model.

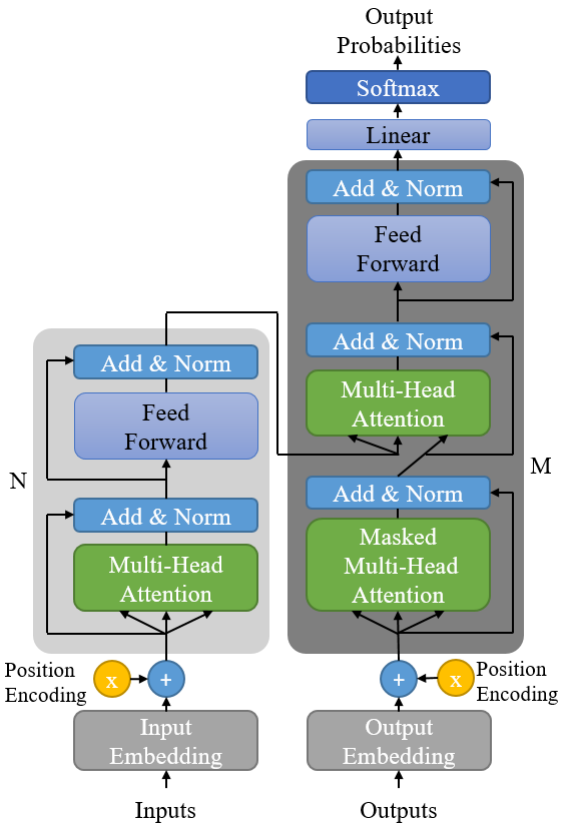


Fig. 1. Full Model architecture.

The entire encoder-decoder structure is auto-regressive at each step, i.e. it uses the previously generated output as additional input while generating current output.

$$y_i = x_i + y_{i-1} \tag{1}$$

As shown in Fig. 1 *N* and *M* is the number of identical layers of encoder and decoder. For encoder *N=6*, which means the encoder has 6 similar blocks each of which has two components- a multi-head attention and a position-wise feed-forward network. Each of this sub-layer contains a residual connction [5] followed by a layer normalization [6]. The concept of residual block is to add the output of a layer with the input of that layer. On top of that, it has a normalization layer. If x is the input to the sublayer that the output will be as followed.

$$y = LayerNormalization(x + sublayer(x)) \tag{2}$$

The decoder has also 6 identical blocks. Along with Multi-head attention and feedforward layer, it has a masked multi-head attention layer. The purpose of the musk is to prevent the model from seeing the subsequent positions in the output sequence. This makes sure the prediction of position i can depend only on the known outputs of position less than *i*. It also has the residual connection and the layer normalization.

Now we will go over the components of the encoder-decoder sublayers.

Attention. Attention is a concept of focusing on a particular portion of data while generating the current output. It is an important part of the current state of the art neural machine translation model. When translating the current word it worth giving more attention to current word and last few words and less attention in the very past sequences. But it is not a hardcoded rule. It might be important to give attention to where the last gender was introduced which might be important for the current word, e.g. his/her. So, attention is also a learnable parameter of the network. In our case, to correct the current word it is most important to see the last few words. The attention concept was popularized in computer vision application of image caption generation [7].

Both the encoder and the decoder use multihead attention. The input to the attention function is a set of key-values and a queue. The attention maps the query and key-values into the output. Here, all the inputs are vectors.

Fig. 2 shows two types of attention function. The left one is simple scaled dot-product attention, the right one is the modification to create the multi-head attention.

Scaled Dot-Product Attention. Two commonly used attention functions are dot-product attention and additive attention. The additive attention a single layer neural network to compute the output. Whereas the dot-product attention are simple matrix multiplication and faster than the additive

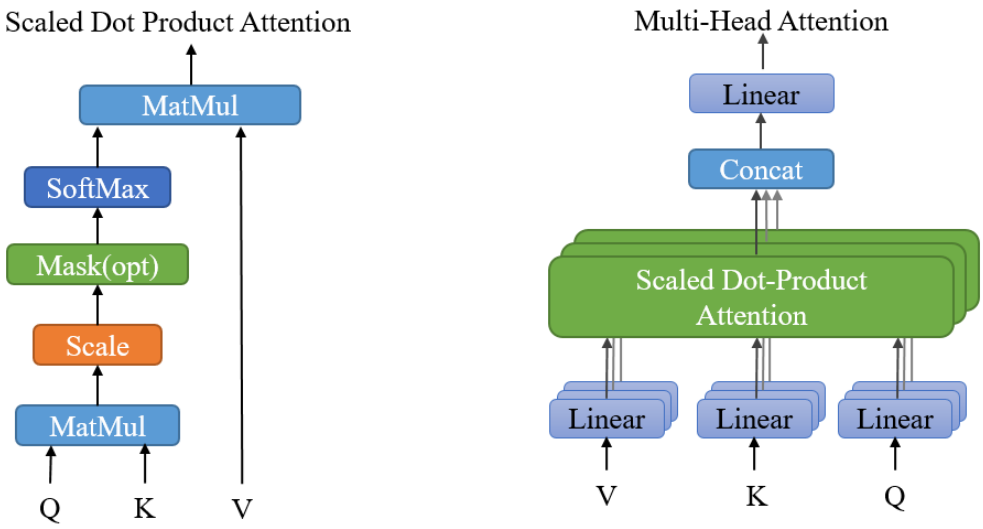


Fig. 2. (left) Sacled Dot-product attention. (right) Multi-Head Attention of several layers of attention.

attention.

Let, K is the keys and V is the values and Q is the queries as the input to the attention function. Where, queries and keys of dimension d_k and values have dimension d_v. The output is the softmax over the dot product of keys with queries and the output multiplied by values. The transformer [3] model introduces scaled dot-product attention by scaling the dot product with the square root of the dimension of keys. The output of the attention is represented as:

$$Attention(K, Q, V) = Softmax(\frac{QK^T}{\sqrt{d_k}})V \quad (3)$$

Multi-Head Attention. Instead of using simple one attention function it is beneficial to linearly project parallel sets of keys, values, and queries. We can calculate the parallel sets of attention outputs and concatenate them input single output. This multi-head approach gives the model an opportunity of attending more than one place at a time. The formulation can be represented as:

$$MultiHead(K, Q, V) = Concat(h_1, h_2, ..., h_n)W^O \quad (4)$$

$$h_i = Attention(KW_i^K, QW_i^Q, VW_i^V) \quad (5)$$

Here, n is the number of parallel queries. We have used n=8. W^Q, W^V, W^K, W^O are weight or projection parameters for queries, values, keys, and outputs.

In the attention function of the encoder-decoder layer, the encoder provides keys and values to the function and the decoder provides queries. As a result, the decoder has information overall input and output sequences got so far. This gives the model similar capabilities that a sequential attention-based model has.

The encoder attention is self-attention. If the queries, keys, and values all comes from the same place, the attention is called self-attention. In this case, the encoder can attend all position in its past sequence. Similarly, the decoder also contains a self-attention block before the attention over encoder. The decoder self-attention is masked. The purpose of the masked attention is to restrict the decoder attending or seeing the future position output in the current position. On top of that, the outputs are shifted one position as the current position must see up to one last position output at generating current output.

Position Encoding. As the model does not have any recurrent or convolution unit, the model has no way to know the position information. To let the model understand the position information we feed the input with a position data. There are different ways of assigning fixed-length position data. We followed the sine-cosine function introduced in the annotated transformer model.

$$PositionEncode(p, 2i) = sin(p/10000^{2i/d}) \quad (6a)$$

$$PositionEncode(p, 2i+1) = cos(p/10000^{2i/d}) \quad (6b)$$

where p is the position and i is the dimension. d is the dimension of the intermediate output of the model.

III. Experiments

We have implemented the model with pytorch [8] int python. All experiments were done in NVIDIA 1060 GPU. Here we first describe the dataset and the training procedures. We have provides the experimental setup and the analysis on the result.

A. Data

To learn the pattern of English language and correct the spelling errors we need a large dataset with lots of plain text. There are a few datasets of plain English text which are used for language modeling purposed mainly. We chose to use the Wikitext-103 [9] dataset for our application. This is a dataset collected from wiki dumps.

The dataset contains a vocabulary size of around 268k collected over 28k Wikipedia articles. It contains a total of over 100 million words. The dataset is divided into train, test and validation set. The test and validation dataset contains 60 articles each having around 246k and 218k respectively. This is a more realistic and representative dataset compared to other.

B. Experimental Setup

To train the spelling corrector we have taken the denoising approach. For that, we have built a sequence to sequence model. The input to the model is noise added English test and the output is the corrected sentence. The noise comes in form of changing, removing or adding a few characters. The model is trained with the original sequence for learning to remove the spelling errors.

Instead of regular stochastic gradient descent optimizer, we have used Adam [10] optimizer. It has two momentum parameters for controlling the learning rate. These are called beta1 and beta2.

We have applied dropout [11] to the output of all the sub-layers before it is added to the sublayer input residual connection and normalizes. Dropout is an efficient method for handling overfittings. We have used the $P_{drop} = 0.1$. We trained the model for 50 epochs with an initial learning rate of 0.001 and a batch size of 64.

All the training parameters are summarized in table I.

TABLE I
VALUES OF TRAINING PARAMETERS

Parameters	Values
Optimizer	Adam
Adam: beta1	0.5
Adam: beta2	0.999
d_{model}	512
d_k	64
d_v	64
n_{head}	8
Epochs	50
Batch size	64
Learning rate	0.001
LR Decay rate	0.99

C. Result Analysis

We have done the experiment with different types and variations of errors. First, we have done the experiment with a different number of errors in a sequence. The errors are introduced in character level. For this experiment we have takes the total length of the input sequence as 50. We reported the result in accuracy metric. We accuracy is calculated at the character level, where the accuracy is the percentage of character correct compared to the original sequence.

Table II summmarizes the result.

TABLE II
NUMBER OF ERRORS VS ACCURACY COMPARISON

Number of errors	Accuracies(in percentage)
2	99.36
3	99.06
4	98.18
5	97.88

We have introduces 2 to 5 characters errors in the sentences. Having more than 5 error in a 50 character sequences changes the meaning in a very unrecoverable way in most cases. From the table, we can see the increase in accuracy with less number of errors. The model produces enough good result with even 5 errors in the sentence.

Then we conducted experiments with the total length of the sentence. Regular RNN based sequence to sequence model has a tendency of generating a bad result with increased in the length of the sequence. Our attention-based model is more rigid with the length of the sequence. For this experiment, we introduced 1-4 character errors randomly.

Table III summarizes the result.

TABLE III
SEQUENCE LENGTH VS ACCURACY COMPARISON

Number of errors	Accuracies(in percentage)
30	98.12
50	98.18
75	98.02
100	97.93

From the table above we can see that the accuracy of the model fluctuates very little with the length increment. Even that is an increase in accuracy from an increase of 30 to 50 in length. The reason here is the fact of the number of errors. Too many errors in a short length of sentence introduce too much information loss to reconstruct the original sequence.

IV. CONCLUSION

The spelling mistake is a natural phenomenon in typing and it is important to have a system capable of correcting those. The simple dictionary-based spelling corrector system focuses on correcting each error separately without understanding the meaning or the context of the sentence which result in an error in sentence-level meaning or grammar. In this work, we have used deep learning to build a context-aware spelling corrector system that uses an attention mechanism to focus on the entire sequence to generate the output. The experiment shows impressive results on a wide range of errors. Our model is robust to change in the length and number of errors in the sentence and achieved close to perfect accuracy. This work

will help to lead a new way of solving the spelling correction problem.

REFERENCES

[1] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[2] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems*, 2017, pp. 5998–6008.

[4] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in neural information processing systems*, 2014, pp. 3104–3112.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[6] J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer normalization," *arXiv preprint arXiv:1607.06450*, 2016.

[7] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio, "Show, attend and tell: Neural image caption generation with visual attention," in *International Conference on Machine Learning*, 2015, pp. 2048–2057.

[8] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," 2017.

[9] S. Merity, C. Xiong, J. Bradbury, and R. Socher, "Pointer sentinel mixture models," *arXiv preprint arXiv:1609.07843*, 2016.

[10] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[11] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

Extracting RDF Triples from Raw Text

Yeasmin Ara Akter
Dept. of Computer Science and Engineering
East Delta University
Chittagong, Bangladesh
yeasmin.a@eastdelta.edu.bd

Md. Ataur Rahman
Dept. of Language Science and Technology
University of Saarland
Saarbrücken, Germany
arahman@coli.uni-saarland.de

***Abstract*—This manuscript manifests the results of our work on extracting RDF triples from raw text data. We took a corpus of news articles and applied several methods for extracting "subject - verb - object" relationships from texts. The first method that we used is syntactic parsing and constructing triples from *nsubj* and *obj* syntactic relations. The second approach incorporates extracting semantic roles with the help of frame parser and constructing triples from *agent*, *patient* and *predicate* relationships. After applying the aforementioned methods, a manual evaluation was done over a small number of hand labeled samples. Our system achieved a Precision score of *0.34*, Recall of *0.46* and a combined score of *0.39* for F-measure as the final result.**

***Keywords*—Knowledge and Data Engineering, Semantic Information Processing, RDF, Text Analysis and Mining, Natural Language Processing.**

I. Introduction

Dealing with unstructured data is a common problem in the field of Semantic Web and Linked Data. Probably the most prominent project in the field – DBpedia – has mostly processed data in Wikipedia tables, thus discarding the huge amount of information available in the article texts. The system we developed is thus dedicated to extracting *'Resource Description Framework'* (RDF) triples from raw texts. We aim at extracting specifically RDF data and not just some named entities because the extracted RDF triples then can be used for constructing knowledge graph and creating intelligent *Question Answering* (QA) systems and chatbots.

The goal of our project is to test simple methods that can be used to extract RDF data from natural language texts. The existing research has shown that carefully extracting triples from the text is a very complex task, but we can explore what we can achieve using shallow parsing and simple rules. Even a small amount of extracted data can be useful in various applications.

The full paper has the following structure: in Section II, we will describe the background and existing research, in Section III we will be describing our methodology and in Section IV, the results of our system will be depicted. Section V will present the evaluation of our results. In Section VI we will report what other methods apart from syntactic and semantic parsing we also tried to use. Finally, in Section VII we will conclude our work.

II. Background And Motivation

Several previous works have dealt with the topic of extracting RDF triples from unstructured text. In particular, we are inspired by the work by [1]. They have used a pipeline of semantic processing tools to extract DBpedia RDF triples from Wikipedia articles. Among others, they extracted existing Wikipedia links from the articles, perform semantic parsing based on PropBank [2] and coreference resolution on the article texts and used the Wikifier [3] tool to recognize further entities for which DBpedia entries exist. After the above steps, they obtained triples of the form *subject - predicate - object*, which they subsequently mapped onto the DBpedia namespace. Thus, the central contribution of their work is the extension of the DBpedia knowledge base with the facts which they extracted from Wikipedia articles in the form of RDF triples.

Other applications in the same vein include LODiffier by Augenstein et al. [4] and KnowledgeStore by Cattoni et al. [5]. The former also incorporated Wikifier and used deep semantic analysis and discourse representation structures to obtain relations between the entities recognized. The latter creates a knowledge base from multimedia resources with the aid of an existing background knowledge base; among others, they resolved coreference and used clustering to detect entities about which attributes can be extracted using so-called *triggers*. Gerber and Ngomo [6], on the other hand, used a pattern-based approach; instead of performing semantic parsing, they obtained triples by extracting recurrent patterns which represent natural language predicate-argument constellations.

In our project, we largely proceed along the lines of [1]. From the raw text, we intend to extract facts in the form of *subject-predicate-object* RDF triples with the help of semantic processing which will be compatible to be mapped into the DBpedia namespace where possible. In contrast to [1], we will not be using Wikipedia articles but rather plain running texts. This has the advantage that we will not need to clean our data of annotations in Wikipedia's markup language. On the other hand, Wikipedia articles will already contain useful links for entities which are covered by DBpedia, while our plain text data will not.

III. Methodology And Experimental Setup

A. Dataset

For the purpose of our experiment, we decided to use a corpus consisting of *News Articles*[1] provided by Signal Media [7]. It consists of texts from several news agencies, websites and blogs. The overall size of the dataset is over 100 million tokens, but we decided to use a portion of it that consists of 500000 tokens.

B. Pipeline

We intended for our system to be a pipeline of the following component:

[1] http://research.signalmedia.co/newsir16/signal-dataset.html

a. The Plain raw text has been obtained from the corpus data.
b. Tokenization and POS-tagging is performed on the data.
c. Both the dependency parsing and semantic role-labeling techniques have been explored.
d. From the processing output of step c, triples are extracted in the form: *subject argument - verb - object argument*
e. Finally, triples of the aforementioned structure are converted into RDF-triple (JSON-LD format).

Both the extraction of triples based on dependency parsing (hence syntactic processing) and on semantic role-labeling (hence semantic processing) were explored. At a later stage, we chose the latter.

C. Triple Extraction Based on Dependency Parsing

We performed syntactic parsing using UDPipe [8], which tokenizes the input sentences and returns for each sentence, among others, the word lemmas and the universal dependency relations which hold between the individual words in the sentence. We then extracted the following three lemmas to form a triple:

- The lemma with the dependency relation *root* (i.e. the root word of the sentence).
- The lemma with the dependency relation *nsubj* to the sentence root word.
- The lemma with the dependency relation *obj*, to the sentence root word.

The above three lemmas would respectively correspond to the *verb*, the *subject argument* and the *object argument* of a triple. To illustrate, Fig. 1 shows the relevant parts of the UDPipe output for the sentence '*The Texas State football team scored 63 points in its home opener*', from which we extracted the triple *(team - score - point)* (in red).

1	The	the	DET	DT	Definite=Def|PronTy	5	det
2	Texas	Texas	PROPN	NNP	Number=Sing	3	compound
3	State	State	PROPN	NNP	Number=Sing	4	compound
4	football	football	NOUN	NN	Number=Sing	5	compound
5	team	team	NOUN	NN	Number=Sing	6	nsubj
6	scored	score	VERB	VBD	Mood=Ind|Tense=Pa	0	root
7	63	63	NUM	CD	NumType=Card	8	nummod
8	points	point	NOUN	NNS	Number=Plur	6	obj
9	in	in	ADP	IN	_	12	case
10	its	its	PRON	PRP$	Gender=Neut|Numb	12	nmod:poss
11	home	home	NOUN	NN	Number=Sing	12	compound
12	opener	opener	NOUN	NN	Number=Sing	8	nmod
13	.	.	PUNCT	.	_	6	punct

Fig.1. Example UDPipe output with tokens forming the triple marked in red

This approach, based on dependency parsing, was characterized by the fact that each triple element would consist of exactly one single token. But this easily gave rise to errors of the following kind: The alleged triple *(Williams - get - America)* was falsely extracted from the sentence *Vanessa Williams Finally Got Her Miss America Apology*, which is at least in part due to the fact that the object argument, *(Her) Miss America Apology* consists of multiple tokens, while the dependency parser will only assign to one of the dependency relation *obj*. For this reason, we chose not to base further processing steps on this approach but turn instead to triple extraction on the basis of semantic processing described in the next section.

D. Triple Extraction Based on Semantic Role-labeling

Semantic role-labeling was conducted using the Senna system [9], which performs, among others, tokenization, chunking and semantic role-labeling on plain texts. A sample output of Senna for the sentence '*VETERANs saluted Worcester's first ever breakfast club for ex-soldiers which won over hearts, minds and bellies*' is given in Fig. 2 (with the columns with information on POS-tags and semantic roles marked in red).

VETERANS saluted Worcester's first ever breakfast club for ex-soldiers which won over hearts, minds and bellies.

VETERANS	NNP	S-NP	0	-	S-A0	0	(S1(S(NP*)
saluted	VBD	S-VP	0	saluted	S-V	0	(VP*
Worcester	NNP	S-NP	S-ORG	-	B-A1	0	(NP(NP*
's	POS	B-NP	0	-	I-A1	0	*)
first	JJ	I-NP	0	-	I-A1	0	*
ever	RB	I-NP	0	-	I-A1	0	*
breakfast	NN	I-NP	0	-	I-A1	0	*
club	NN	E-NP	0	-	E-A1	0	*)
for	IN	S-PP	0	-	B-A2	0	(PP*
ex-soldiers	NNS	S-NP	0	-	I-A2	S-A0	(NP(NP*)
which	WDT	S-NP	0	-	I-A2	S-R-A0	(SBAR(WHNP*)
won	VBD	S-VP	0	won	I-A2	B-V	(S(VP*
over	IN	S-PP	0	-	I-A2	E-V	(PP*
hearts	NNS	S-NP	0	-	I-A2	B-A1	(NP*
,	,	0	0	-	I-A2	I-A1	*
minds	NNS	B-NP	0	-	I-A2	I-A1	*
and	CC	I-NP	0	-	I-A2	I-A1	*
bellies	NNS	E-NP	0	-	E-A2	E-A1	*))))))))
.	.	0	0	-	0	0	*))

Fig.2. Example Senna output with POS-tags and semantic roles marked in red

In accordance with the PropBank framework [2], adopted by [9] for Senna - *V*, *A(rgument)0* and *A(rgument)1* respectively denote the predicate, the agent and the agent arguments. For each sentence, we therefore, extracted from the Senna output those words tagged as *A0* as the subject argument, those tagged as *V* as the verb, and those tagged as *A1* as the object argument. In the example shown in Fig. 2, we would output the triples *(VETERANS - saluted - Worcester's first ever breakfast club)* and *(ex-soldiers which - won over - hearts, minds and bellies)*.

However, without further cleaning, our approach would include alleged triples in which the subject is a personal pronoun, or in which some elements (e.g. the verb) length is longer than usual length for example the chunk '*obtain amid an expected record turnout, in what would be a big boost to a secession campaign which has been losing support over the last two years*' as Senna would tag the entire phrase as *V*. Therefore, we included the following constraints into the extraction process, which relate to the length of each element of a triple as well as the POS categories of the words:

- Words forming the verb in the triple must have as their POS-category one of the following tags: VB, VBD, VBG, VBN, VBP, VBZ or IN. These are the verbal POS-categories with the addition of IN for

prepositions, which is intended to include phrasal verbs such as *win over*.

- Words forming the subject and object arguments must *not* be of the POS-categories PRP, PRP$, WDT or WP, which represent personal and possessive pronouns as well as WH-determiners and pronouns.
- Alleged triple is discarded if the *verb* consists of more than three words.
- Alleged triple is discarded if the *subject* or *object* argument consists of more than eight words.

The above constraints ensured that, for instance, we would in fact extract *(ex-soldiers - won over - hearts, minds and bellies)* from the example in Fig. 2, excluding the word *which* from the subject argument.

E. Conversion of Triples to RDF

After extracting the triples from the Senna output we needed to convert it to a suitable format so that these can be read as RDF. We used JSON-LD format as an intermediate form because it is easy to convert text into JSON programmatically. Since JSON-LD is a *Linked Data* type, it is easily readable by any program such a Protégé. We thus used Protégé to convert the JSON-LD file to RDF/XML.

The entire pipeline of our implementations is given in Fig.3., with the UDPipe module greyed out as we did not continue onwards with its output.

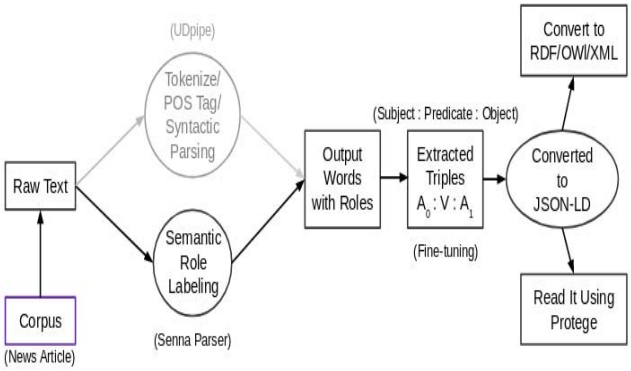


Fig.3. The Pipeline of our system with UDPipe module greyed out

IV. Result

As the result of our project, we extracted 2361 triple from our data[2]. The triples look in the following way:

Subject	Predicate	Object
A migrant	holds	A banner

The extracted triples were organized in an ontology with different subjects and objects serving as instances of individual classes, and predicates serving as object properties. The overall organization of the ontology can be seen in Fig. 4.

[2] Results of our system can be downloaded from the following link under the folder *Final RDF* - https://goo.gl/z4Yuiz

The ideal organization of the ontology would be to create different general classes like *Person*, *Organization* etc, but as we do not have a gold standard which we could use as a training data for classifying our instances, this step remains to be added in the future.

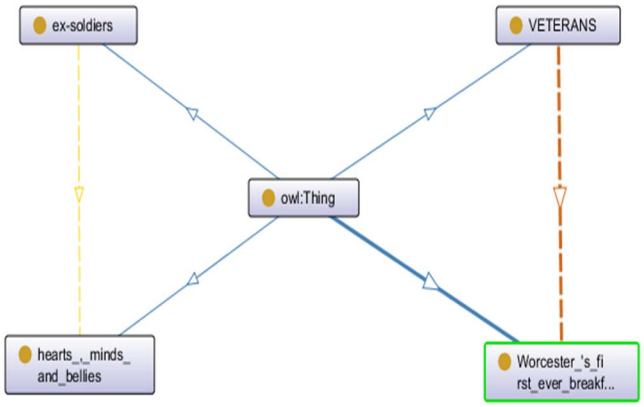


Fig.4. Organization of triples in the ontology.

V. Evaluation

Evaluation poses a challenge to our proposed project since we cannot compare the triples extracted by our system to an existing gold-standard. One possibility is simply to select a small sample of data and to manually extract triples from these sentences. The results of this manual extraction can function as gold-standard against which the system's performance could be evaluated in terms of precision, recall and F-measure.

To perform this kind of evaluation, we manually annotated a sample of 100 sentences. The main difficulty is to define what exactly we want to count as a part of a triple. It is obvious that triples should not include only named entities. At the same time, many cases are not interesting or useful for constructing a knowledge graph. For example, in a sentence '*mother talks with her daughter*' we wouldn't want to extract anything, though the triple can be extracted with the means of semantic parsing. In the end, we decided to analyze as triples those entities which are not generic, for which a Wikipedia page exists or might exist, or which can be considered named entities.

In the selected sample our system has found 81 triples. Based on the annotated sample, we calculated recall and precision measures. They can be found in Table I. The most common problem that can be noticed while annotating the triples is that borders of an extracted and annotated triple do not coincide. The system tends to take larger chunks, whereas only two or three words are enough for manual annotation. Also, our system has clear rules by which it selects certain phrases as parts of a triple: the chunks should have such semantic roles as agent, patient or predicate. Meanwhile, the data that can be used in RDF triples is much more varied, and many other semantic roles might be included in the triple. For example, time and place such as in the phrase '*The Worcester Breakfast Club ... met at the Postal Order in Foregate Street at 10 am on Saturday*'. are not detected by our system, whereas they should be included in a proper knowledge graph.

TABLE I. RESULTS FOR THE EXTRACTION OF RDF TRIPLES

Precision	*Recall*	*F-measure*
0.34	0.46	0.39

VI. DISCUSSIONS AND FUTURE WORK

At present, our triples have been mapped to our own example URL and are therefore not linked to any existing knowledge base. Ideally, we would like to enrich an existing knowledge graph with the facts we express in our extracted triples to the effect that these facts can be queried in a question answering system. This would resemble the final step in Exner and Nugues's work [1], in which they map their extracted triples to the DBpedia namespace.

Hence, we also attempted to link our triples to DBpedia entities using DBpedia Spotlight [10] or Wikifier [3], the latter of which have been used both by Exner and Nugues [1] and by Augenstein et al. [4]. The idea was to apply either of the tools to our raw text data. Ideally, some of the entities recognized (i.e. entities covered by DBPedia) would correspond to the subject or object arguments in our triples, and the fact expressed by our triples would be added to those entities as novel knowledge.

However, we ran into a similar problem as mentioned in the previous section with regard to evaluation: Senna outputs large chunks of text as subject and object arguments. They often have a complex internal structure and, for our purpose, ought to be further broken down into more "atomic" units. For instance, the object argument '*Worcester's first ever breakfast club*' of our aforementioned triple (see section III.D) is in fact internally complex, from which we would ideally extract, at a minimum, the triples *(unique_id – isA - breakfast_club)* and *(unique_id - hasLocation - Worcester)*. Both DBpedia Spotlight and Wikifier were able to recognise and link *Worcester* and at times also *breakfast_club* (depending on the pre-selected confidence value); however, '*Worcester's first ever breakfast club*' as a whole, which functions as the object argument in our triple, could, plausibly, not be recognized by either tool. Therefore, in order to connect our triples to the DBpedia knowledge base in the future, more fine-grained processing of the elements of each triple is needed, possibly using further semantic tools.

VII. CONCLUSION

In our system, we attempted to apply simple methods for extracting RDF triples from raw text data. It turned out that this task can be solved to some extent using semantic parsing and converting the extracted semantic roles to RDF triples. The resulting F-measure was 0.39, which can be improved in many ways, but still is decent for the beginning.

The process of extracting RDF triples can be improved both on the side of semantic parsing and on the side of conversion. In terms of semantic parsing, we can try to extract more complex semantic relations like date and time and also do anaphora resolution. Apart from that, our data needs more linguistic preprocessing like filtering out non-English words and special symbols, lemmatizing all the triple instances and removing possible duplicates. While in the case of conversion, we need to better work through the structure of our ontology and develop macro-classes for our individual instances so that the ontology becomes less sparse and more concise. For doing that, we need to collect gold standard data which can be used for applying machine learning to classify individual instances.

REFERENCES

[1] P. Exner and P. Nugues, "Entity extraction: From unstructured text to dbpedia rdf triples," in *The Web of Linked Entities Workshop (WoLE 2012)*, 2012.

[2] M. Palmer, D. Gildea and P. Kingsbury, "The proposition bank: An annotated corpus of semantic roles," *Computational linguistics,* vol. 31, pp. 71-106, 2005.

[3] L. Ratinov, D. Roth, D. Downey and M. Anderson, "Local and global algorithms for disambiguation to wikipedia," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, 2011.

[4] I. Augenstein, S. Padó and S. Rudolph, "Lodifier: Generating linked data from unstructured text," in *Extended Semantic Web Conference*, 2012.

[5] R. Cattoni, F. Corcoglioniti, C. Girardi, B. Magnini, L. Serafini and R. Zanoli, "The KnowledgeStore: an Entity-Based Storage System.," in *LREC*, 2012.

[6] D. Gerber and A.-C. N. Ngomo, "Extracting multilingual natural-language patterns for rdf predicates," in *International Conference on Knowledge Engineering and Knowledge Management*, 2012.

[7] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu and P. Kuksa, "Natural language processing (almost) from scratch," *Journal of machine learning research,* vol. 12, pp. 2493-2537, 2011.

[8] D. Corney, D. Albakour, M. Martinez-Alvarez and S. Moussa, "What do a million news articles look like?," in *NewsIR@ ECIR*, 2016.

[9] J. Daiber, M. Jakob, C. Hokamp and P. N. Mendes, "Improving efficiency and accuracy in multilingual entity extraction," in *Proceedings of the 9th International Conference on Semantic Systems*, 2013.

[10] M. Straka, J. Hajic and J. Straková, "UDPipe: Trainable Pipeline for Processing CoNLL-U Files Performing Tokenization, Morphological Analysis, POS Tagging and Parsing.," in *LREC*, 2016.

Bangla Spell Checking and Correction Using Edit Distance

Muhammad Ifte Khairul Islam
Daffodil International University
khairul3182@diu.edu.bd

Rahnuma Islam Meem
Daffodil International University
raahnuma1711@gmail.com

Faisal Bin Abul Kasem
Daffodil International University
mfaisalpasha@gmail.com

Aniruddha Rakshit
Daffodil International University
aniruddha.cse@diu.edu.bd

Md. Tarek Habib
Daffodil International University
md.tarekhabib@yahoo.com

***Abstract* - Automatic spelling correction for Bangla language is a very important thing as the modern world is almost completely dependent on digital devices where electronic keyboards are used. Bangla is a largely used language (3.05% of world population), worldwide securing the 7th most spoken language. But concerning the natural language processing it is still in its initial stage. Automatic spelling correction of Bangla language is just crossing its primitive stage. A few works have been done on automated Bangla spelling correction but those numbers are not still adequate. This research work is based on correct word prediction of typed Bangla sentences by using edit distance method and fractional accuracy. The results are promising. So, there is a high chance that the proposed model will help and have a great impact on automated Bangla spelling correction.**

Keywords - Bangla Spelling, Spelling Correction, Non-word error, Fractional Accuracy.

I. INTRODUCTION

Written form of a language is a very important factor in communication; whether it is handwritten or typed. When speaking it is not important to know proper spelling of a word and even without proper maintenance of grammar communications can be conducted, but when writing it is very important to have correct spelling. Spelling mistake is a common issue especially if it is a complex language such as Bangla. Bangla is one of the most used languages (3.05% of world population) [1] and it has complex word construction, so it is common to have spelling error while writing. It is not necessary that the reasons for spelling mistakes are only because of its complexity, it might also occur due to fast typing or lack of knowledge. As it is a common phenomenon to make spelling mistakes, automatic spelling correction processes are a must. There are lots of work on this issue but when considering Bangla there aren't many notable works. The main focus of this research work is checking the spelling of Bangla word following by the correction of it using fractional accuracy. Direct dictionary lookup and edit distance method will be used to check correctness and gain possible corrections of the word then the accuracy will be calculated. There are some works that used the edit distance method but, in this research, a much larger dictionary and new method to calculate accuracy are used which will give better results.

The rest of the paper is organized as follows: in section II a literature review is done where previously published works that are similar to our work has been discussed. Section III is about the methodology of our work where the process of our work described. Next, in the implementation section, results of our work is shown. Section V shows the comparison results with previous works. Finally section VI concludes the paper describing how this work is promising.

II. LITERATURE REVIEW

There are quite a number of works done on spelling mistake detection and correction. Spanish, English, Mandarin, Bangla etc. are some of the most spoken languages on earth, so it is normal to work on these languages but automated spell checker and correction is of such an importance that a number of works were also done on less spoken compared to the languages mentioned before or local languages such as Myanmar, Persian and Hindi [2], [3] and [4]. All over the world people are working on this topic not necessarily focusing on any specific language but also on language independent spell checker and corrections [5].

Languages such as English, Chinese and Arabic are largely spoken languages thus having more works focusing on them. Some notable works on English are done in [6], [7], [8], [9] and [10] using different methods such as edit distance, tree traversal etc. [11] and [12] are two works done on Chinese and Arabic.

As the main focus of this work is on Bangla, so if works done on automated spell checking in Bangla is considered then number of papers are small [13], [14], [15], [16], [17], [18], [19], [20] and [21] that focus on this common concern with this paper. All these papers have different techniques deployed but there are two common facts about all these papers and those are small corpus size and lack of balance. As a solution to the problem of Bangla spell checking using clustering and edit distance is performed by P. Mondal and B.M.M. Hossain claiming to have an accuracy of 99.8%, which was applied on 2450 words [13]. N.U. Zaman and M. Khan claimed accuracy of 91.67% with the data set of 1607 words using mapping based on double metaphor and edit distance as a solution [16]. To identify phonetic error, string matching is used by Bidyut Baran Chaudhuri [17]. He used 2 dictionaries to check and correct errors but his process mainly considered phonetic errors and it also needs double memory for the dictionary and the reverse dictionary he used for correction of misspelled words. In another one of M. Z. Islam, M. N. Uddin and M. Khan's paper they used stemming algorithm in which if the stem is not found, then it produces a list of possible suggestions using suggestion generation process [18]. They used to edit distance algorithm to find best match. Another work focusing on phonetic error is done by N. U. Zaman and M. Khan in which they modified phonetic encoding based on soundex algorithm for matching Bangla phonetic [15]. Using finite state automaton M. K. Munshi et. al. proposed an approach to suggest the probable list of correct words against the error word [19]. NH Khan, GC Saha, B Sarker, MH Rahman for checking error used character-based N-gram but it didn't

suggest any correction to the misspelled word [21]. To detect error words a direct dictionary, look up and binary search method and to suggest correct word against the incorrect word recursive simulation method is used by ABA Abdullah, A Rahman in their paper [20].

There are many English spell checking and correction works having the accuracy of 86.95%, 96.1% and 96.4% for both real word and non-word errors [8], [7]. Even though English spell checking and correction have been studied in [6], [7] corpus size or test data size were not mentioned. In [11] a work on Chinese language spell checking and correction, edit distance algorithm, soundex algorithm, and combining them with pinyin is used. In [12], a system for checking Arabic language spelling using context words and N-gram language models is used with a corpus size of 41,170,678 words and twenty-eight confusion sets for their experiment. Their average accuracy rate is 95.9% that handle both real-word errors and non-word errors.

A. Types of Error

Kukich [22] classified all spelling errors into two types. One is real-word error and the other one is non-word error. Contextually inappropriate though the word is valid is called real-word error. For example, in the sentence "I eat milk.", "eat" is not contextually appropriate but it is a valid word. Similarly, in Bangla, in the sentence "আমার আব্বু বাটি যাবে", বাটি is not contextually appropriate but it is a valid word. So, a real-world error is found here. Lexically invalid words are called Non-word error. For example, in the sentence "we aer going to play", 'aer' is not a valid word. In the same way in Bangla, 'নাটি' is a lexically invalid word in "আমার আব্বু নাটি যাবে". There is some more classification of non-word spelling errors offered by Kukich [22]. One is cognitive error and the other one is typographical error. Cognitive error occurs when user does not know the spelling of the erroneous word. Typographical error occurs due to typing mistake. For example, different errors of the Bangla word "চলমান", 'চলম্মান', 'চমান', 'চমমান' and 'চলমনা' are caused by insertion, deletion, substitution and transposition respectively.

III. METHODOLOGY

In this work we are handling non-word, especially typographical errors. We used direct dictionary lookup method to detect an error word. We have a dictionary from which we are searching if the typed word is available. If the typed word is not available then we can consider it as an error and then use edit distance algorithm to correct the erroneous word. To correct the error, we estimate similarity of structure between misspelled word and candidate correct words using minimum edit distance. It measures the minimum number of total operations required to transform one string into another. There can be various operations like insertion, deletion and/or substitution. To calculate edit distance, Levenshtein algorithm [23] is used as written in Algorithm 1. The working principle of Algorithm 1 is that, if we consider two strings X and Y (X= misspelled word, Y= candidate correct word) with the lengths m, n and create a matrix of order m×n edit distance between different sub-strings of X and Y. The corresponding values will express the minimum distance between X and Y. It is assumed that insertion and deletion each has a cost of 1 and substitution has a cost of 2.

Algorithm 1. Algorithm for calculating minimum edit distance

```
min_distance(misspelled_word,candidate_correction)
   m ← length(misspelled_word)
   n ← length(candidate_correction)
   create distance matrix dis [n+1, m+1]
   for each column i ← 0 to n
   for each row j ← 0 to m
                    dis[i,j]←min(dis[i-1,j]+ins-
                cost(candidate_correctionj),
                   distance[i-1,j-1]+subst-
                cost(misspelled_word,
                   candidate_correctioni),
                   distance[i,j-1]+ins-
                cost(misspelled_wordj))
   return min-distance
```

System chooses the word with the minimum edit distance for the misspelled word. We calculate the accuracy using fractional accuracy method [24]. For fractional accuracy the accuracy is measured by taking both the matching of suggested word with intended word as well as the order of matching into account. Therefore, accuracy rate is used in order to address this issue. Suppose w_m matches with v_{mi}, where w_m is intended word and v_{mi} is suggested word (i.e. w_m equals v_{mi}, where $1 \le i \le n + 1$), then the accuracy is

$$Acc = \frac{n+1-i}{n} \times 100\,\% \qquad (1)$$

Here n is length of suggestion list and i is the position of matching word. Failure occurs when i equals $n + 1$. $(n + 1)$-th match means no match has taken place, i.e. accuracy equals 0. The whole methodology is shown in Figure 1

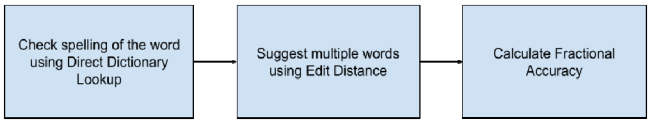


Fig. 1. Block Diagram of Proposed Method

IV. IMPLEMENTATION

As described earlier we are using a dictionary to check if the input word is correct or not. We have a large dictionary consisting of 4,039,603 unique entries. If a word of the input text is not available in the dictionary then it is marked as erroneous word. When an error word is detected, edit distance method is used to generate list of the correct words. Then we find which word from the generated list is similar to the intended word and we calculate fractional accuracy based on the method described above. For example, we have a list of 4 generated words, we found our intended word at position two then our fractional accuracy will be

$$\frac{4+1-2}{4} * 100\% = 75\%$$

The text used in our research are collected from social sites as well as auto generated. One third (50000 words) of the data is collected from social media and two third (100000 words) is auto generated. Percentage of accuracy based on edit distance range is shown in Table 1.

TABLE I. ACCURACY BASED ON EDIT DISTANCE RANGE

Edit distance range	Total number of error word	Total number of success word	Accuracy
1	135,000	132,000	97.78%
2-3	15000	13245	88.3%

We can see that if edit distance range is between 2 and 3 then we get higher accuracy. Another table (Table 2) shows the accuracy based on source of input text.

TABLE II. ACCURACY BASED ON SOURCE

Source	Total number of error word	Total number of success word	Accuracy
Social Media	50,000	47,145	94.29%
Auto generated	100,000	98,100	98.1%

V. COMPARATIVE RESULTS

After discussing different types of results of our work it is better to conclude with total accuracy result which shows that this research work is promising enough. As we see from Table 3 and Figure 2 that the higher the suggestion word list length is the higher accuracy rate we get. If the suggestion list length is only 1 then the accuracy is 65.17% whereas the accuracy is 96.83% if the length is 9.

TABLE III. ACCURACY TABLE BASED ON SUGGESTION LIST LENGTH

Suggestion list Length	Accuracy
1	65.17%
2	74.14%
3	82.62%
4	86.69%
5	88.76%
6	92.10%
7	93.79%
8	94.48%
9	96.83%

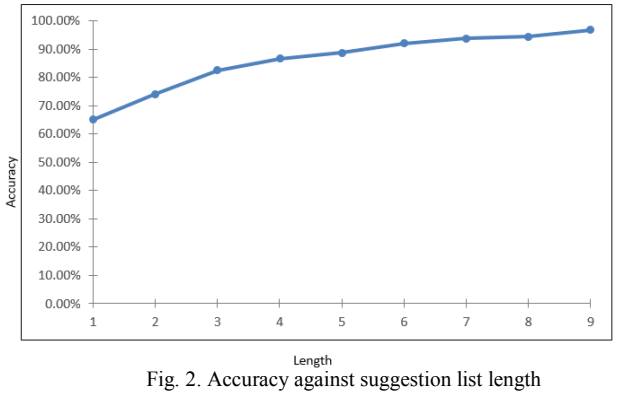


Fig. 2. Accuracy against suggestion list length

Also if we make a comparative table of all works done in Bangla spell checking, we see that our work is promising than others. Table 4 shows the comparative study.

TABLE IV. THE COMPARATIVE ANALYSIS OF ALL WORKS REPORTED

Work/Article	Type of Errors Handled	Algorithm	Test Data Size	Accuracy
This work	Non-word error	Edit distance	150,000 words	96.83%
[13]	Phonetic, typographical	Clustering	2450	99.8%
[14]	Phonetic, typographical, OCR generated	2-edit distance and phonetic encoding	1607	91.67%
[15]	Phonetic, Typographical	Phonetic encoding	*NM*	More than 80%
[16]	Orthographic rules in Bangla	1.Double metaphone 2. Mapping rule.	1607	Edit Distance 0: Correct:1473 Error:134 Accuracy:91.67% Error: 8.33% Error: 107 for edit distance1 Error: 27 for edit distance 2
[17]	Phonetic error	String matching algorithm	25,000 words	high accuracy with 5% false positive detection
[18]	Complex orthographic rules	Stemming algorithm and Edit distance	13,000 words	90.8% for correcting single error misspellings and 67% for correcting multiple error misspellings
[19]	Substitution errors, insertion errors	Finite state automaton	291 words	92% for correcting single character misspellings and 70% for correcting multiple character misspellings
[20]	Typographical errors and cognitive phonetic errors	Direct dictionary look up method and Recursive Simulation algorithm	*NM*	*NM*
[21]	Non-word error	*N*-gram Model (character based)	50,000 correct words and 50,000 incorrect words	96.17%

VI. CONCLUSION

So, the goal of this research was finding a better way which can handle non-word Bangla spelling mistakes. The authors proposed a model based on edit distance method but with different accuracy calculation method than the previous papers on Bangla spelling correction. It is clear that the model proposed here is promising as a large dictionary is used and the accuracy of the model is 96.83%. This research work dealt with non word errors. The authors are working on another research which will deal with real word errors.

REFERENCES

[1] List of languages by number of native speakers, Available at: https://en.wikipedia.org/wiki/List_of_number_of_native_speakers. (Last Accessed: March, 10, 2018).

[2] Aye Myat Mon, "Spell checker for Myanmar language", 2012 International Conference on Information Retrieval & Knowledge Management, 2012

[3] Mohammad Bagher Dastgheib, Seyed Mostafa Fakhr ahmad, Mansoor Zolghadri Jahromi, "Perspell: A new Persian semantic-based spelling correction system", Digital Scholarship in the Humanities, Vol. 32, Pages 543–553, 1 September 2017.

[4] Amita Jain, Minni Jain, "Detection and correction of non-word spelling errors in Hindi language", 2014 International Conference on Data Mining and Intelligent Computing (ICDMIC), 2014.

[5] Behrang Qasemi Zadeh, Ali Ilkhani, Amir Ganjeii, "Adaptive Language Independent Spell Checking using Intelligent Traverse on a Tree", IEEE Conference on Cybernetics and Intelligent Systems, 2006.

[6] Shashi Pal Singh, Ajai Kumar, Lenali Singh, Mahesh Bhargava, KritikaGoyal, Bhanu Sharma, "Frequency based Spell Checking and Rule based Grammar Checking",in International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) 2016.IEEE,2017,pp 4435 – 4439.

[7] Andrew Carlson, Ian Fette, "Memory-Based Context-Dependent Spelling Correction at Web Scale",in Sixth International Conference on Machine Learning and Applications (ICMLA 2007), IEEE 2007.

[8] Ya Zhou, Shenghao Jing, Guimin Huang, Shaozhong Liu, Yan Zhang, "A Correcting Model Based on Tribayes for Real-word Errors in English Essays." 2012 Fifth International Symposium on Computational Intelligence and Design.

[9] Pratip Samanta, Bidyut B. Chaudhuri, "A simple real-word error detection and correction using local word bigram and trigram", Twenty-Fifth Conference on Computational Linguistics and Speech Processing (ROCLING 2013).

[10] Yizheng Cai, Kevin Roland Powell, Ravi Chandru Shahani, Lei Wang, "LINGUISTIC ERROR DETECTION", United States, US 8,855,997 B2, Oct. 7, 2014.

[11] DongHui Li, DeWei Peng. "Spelling Correction for Chinese Language Based on Pinyin- Soundex Algorithm." Internet Technology and Applications (iTAP), 2011.

[12] Majed M. Al-Jefri, Sabri A. Mahmoud, "Context-Dependent Arabic Spell Checker using Context Words and N-gram Language Models."2013 Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences, 2013.

[13] Prianka Mandal, B M Mainul Hossain, "Clustering-based Bangla Spell Checker," in 2017 IEEE International Conference on Imaging, Vision & Pattern Recognition (icIVPR). IEEE 2017, pp 1-5.

[14] N. UzZaman and M. Khan, "A comprehensive bangla spelling checker," In the Proceeding of the International Conference onComputer Processing on Bengali (ICCPB), Dhaka, Bangladesh, 2006.

[15] N. UzZaman and M. Khan, "A bangla phonetic encoding for better spelling suggestions," in Proc. 7th International Conference on Computer and Information Technology, Dhaka, 2004.

[16] N. UzZaman and M. Khan, "A double metaphone encoding for bangla and its application in spelling checker," in 2005 International Conference on Natural Language Processing and Knowledge Engineering. IEEE, 2005, pp. 705–710.

[17] Chaudhuri, Bidyut Baran, "Reversed Word Dictionary and Phonetically Similar Word Grouping based Spell-checker to Bangla Text." Proc. LESAL Workshop, Mumbai. 2001.

[18] Md. Zahurul Islam, Md. Nizam Uddin and Mumit Khan, "A Light Weight Stemmer for Bengali and its Use in Spelling Checker," Proc. 1st Intl. Conf. on Digital Comm and Computer Applications (DCCA07), Irbid, Jordan, March 19-23, 2007.

[19] M. K. Munshi, M. Z. Islam, and M. Khan. "Error-tolerant Finite-state Recognizer and String Pattern Similarity Based Spelling-Checker for Bangla," Proceeding of 5th International Conference on Natural Language Processing (ICON). 2007.

[20] Abdullah, A. B. A., and Ashfaq Rahman, "A Generic Spell Checker Engine for South Asian Languages." Conference on Software Engineering and Applications (SEA 2003). 2003.

[21] Khan, Nur Hossain, et al. "Checking the Correctness of Bangla Words using N -Gram." International Journal of Computer Application 89.11 (2014).

[22] K. Kukich, "Techniques for automatically correcting words in text," ACM Computing Surveys (CSUR), vol. 24, no. 4, pp. 377– 439, 1992.

[23] Daniel Jurafsky and James H. Martin, (2000), "Speech and Language processing", USA: Prentice-Hall, Inc.

[24] Md. Tarek Habib, Abdullah Al-Mamun, Md. Sadekur Rahman, Shah Md. Tanvir Siddiquee, Farruk Ahmed, "An Exploratory Approach to Find a Novel Metric Based Optimum Language Model for Automatic Bangla Word Prediction", International Journal of Intelligent Systems and Applications, Vol. 2, pp. 47-54, 2018.

Filtering Bengali Political and Sports News of Social Media from Textual Information

Lutfun Nahar
Assistant Professor
Dept. of Computer Science & Engineering
International Islamic University Chittagong
Email:lutfacsecu@gmail.com

Zinnia Sultana
Assistant Professor
Dept. of Computer Science & Engineering
International Islamic University Chittagong
Email:zinniaiiuc@yahoo.com

Nilufar Jahan
B.Sc. in Computer Science & Engineering
International Islamic University Chittagong
Email:nilufarjahan@gmail.com

Ummay Jannat
B.Sc. in Computer Science & Engineering
International Islamic University Chittagong
Email:jannat.neeha@gmail.com

Abstract — **Web is being updated rapidly day by day. As a result of the modern media development, news also published very quickly. Generally, it takes less time to appear in the newspapers than an incident happens. But searching the relevant news through internet about one's choice is the foremost task. News classification is the promising area that fulfills what they want .As Bengali is the common language for a huge number of people; they feel comfortable to use it for publishing news in different news portal and social media. Different post and comments are also written in Bangla. So Bangla new in web are promptly increased, classification of Bangla news according to the subject is definitely needed. This paper focused on some machine learning methods such as Naïve Bayes Classifier that is based on probability theory, Support Vector Machine and Neural Networks to classify news as well as post or comments written in Bangla. 1000 long written post are collected from different news portal as well as social media. A Self- created News Corpus is used for experimental analysis. For training and testing a Feature Vector is constructed which is based on TF-IDF. The result shows that the Naive Bayes performs better than other three algorithms. In order to get better accuracy stop word are removed from documents. Bengali stemming has done to reduce the size of the dictionary. Moreover a comparison has been made to show the accuracy of different algorithm.**

Keywords— Naïve Bayes, Stemming, Dictionary, Support Vector Machine, Neural Network.

I. INTRODUCTION

Because of the availability of the Internet and its speed, electronic communication is now much more convenient. Satellite, submarine network everything have made possible to be connected with the global network. So people of every country are fully dependent on the electronic media, social network to collect the news within a short time. Especially youth generation is fully dependent on online publishing newspaper. The score of any games, schedule, commentary etc. are updated time to time in online news portal. Besides, now a day it is very common to advertise anything in social media due to the people's regular activity in social media. In political section, election campaign is very popular in social media for collecting public attention. So social media is also contributing a vital role for publishing important news with the news portal [2]. It's a good initiative to publish news for mass people with their native and mother language. That is why Bengali news portal is taking all kinds of privileges. So Bengali written news classification according to public sentiment is on demand. Previously, the task of news classification has done manually. At present due to success of machine learning process, this manual task running automatically uses some classification methods.When the information or data is much higher, the classification method is more effective. In addition classification of Bangla news is quite penitent because of high frequency of Bangla words and there is no available dictionary for Bangla words.

Recently Various measurable and machine learning procedures has been produced for news classification[1] in vaious dialect.In [5] , KNN is utilized to predict the sample text category in order to find the probability. Naive Bayes probabilistiic classifier is regularly used in content arrangement applications and tests due to its straight forwardness and adequacy which fundamental thought is to utilize the joint probabilities of classes given a record [7]. Naïve Bayes is highly used method for e-mail filtering such as anti-spam [13], [14], [15]. By using Naïve Bayes implement is easier than most other conventional methods [3].Support vector machine (SVM) classifier transforms the original data in a higher dimention using hyperplane. But it also acts much higher when there is poor recall [9].With the accrual of micro-blogging, researchers tried to acquire the sentiment of people and analyzed the opinion of people as negative or positive [10],[11],[12].During the election of Bangladesh Facebook and twitter[8] was used too much to criticize the political situation by the politician in 2014.Many researchers tried to classify post and comment of social media by using naïve Bayes method but without stemming of word. Because of same meaning of various words, the data dictionary can be huge [6].In the field of computational phonetics and probability, n-gram[16] is a contiguous subsequence of n-items in some random succession. An n-gram is a sub-sequence of n-items in any given sequence, where the sequence items or "grams "can be anything, from characters to words. In computational linguistics n-gram models are used most commonly in predicting words (in word level n-gram) or predicting characters (in character level n-gram) for the purpose of various applications [4]. In[17][18] they proposed four dimensionality reduction techniques and show that Principal Component Analysis (PCA) was found to be the most effective in reducing the dimensionality of the feature space. Therefore, the goal of this research is to filtering Bengali written political and sports news from social media using Naive Bayes classification Uni-gram technique in used to compute

probability occourences of each word .Moreover, Bangla stemming is being added to shrink the dictionary size. Thus the present section introduces the problem addressed in this research, while methodology is elaborated in the Section II. The results are presented in Section III and Section IV concludes the paper.

II. METHODOLOGY

The main purpose of news classification is to divide the documents into their respective categories. So we have followed the instructions given below:

1. To collect our data, we have selected some different online news portal based on popularity related to our content and some random post and comments from social media sites.

2. We use individual word file news document for the convenient of use.

3. After creating distinct document file we prepare the documents text by several techniques such as tokenization, punctuation removal, digit removal and so on.

4. In our training phase we fabricated a light Bengali lexicon such as Bengali data dictionary with proper keyword related to sports and political words and those are frequently uses in such articles.

5. We carry out some advanced text processing task like single letter words removal as well as stop word removal.

6. We induced a light Bengali stemmer to identify the stem of each word. It helps us mostly to reduce the size of our established dictionary.

7. In this way we derive each text document as a word-vector. Thus we found the trained document; the result are feed to new documents and we categorize the documents based on Naive-Bayes classifier.

The overview of our approach is depicted in Fig. 1.

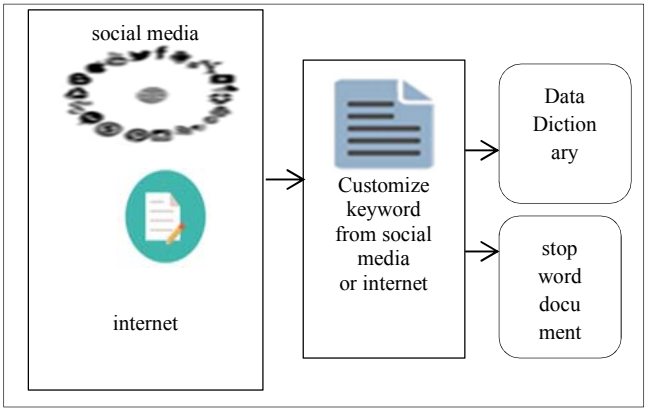


Fig.1.Overview of the Training Phase

A. Data collection

The source of our data is from different online news portal. Mainly we have collected long post and comments so that the sentiment of that post can be verified correctly.Sample collection of our data is given in Table I.

TABLE I. Different Types of Post

"ক্রিকেটারদের টাকা-পয়সা নিয়েও বেশ কিছু নিয়ম বেঁধে দিচ্ছে বিসিবি। যেসব ক্রিকেটারকে ক্লাব রেখে দিতে চায়, ড্রাফটের আগেই তাদের পাওনা টাকার ৫০ শতাংশ দিতে হবে, নাহলে তারা ড্রাফটের জন্য উন্মুক্ত হয়ে যাবেন। ড্রাফট থেকে নেওয়া ক্রিকেটারদের মোট টাকার ৫০ শতাংশ দিতে হবে ড্রাফটের পর বেঁধে দেওয়া সময়সীমার মাঝে, ২৫ শতাংশ দিতে হবে টুর্নামেন্ট চলার সময়ে, বাকি ২৫ শতাংশ টুর্নামেন্ট শেষ হওয়ার পর[19]. "
"হাথুরুসিংহে যোগ দেওয়ার পর পরেই তিন বছর আগে ফিল্ডিং কোচ হিসেবে যোগ দিয়েছিলেন হ্যালসলও। গত বছর হয়েছেন সহকারী কোচ। হাথুরুসিংহে চলে গেছেন, তবে হ্যালসল মনে করছেন এটা খেলোয়াড়দের জন্য শাপেবর হতে পারে[22][23]".
"৮ চার ও ১ ছক্কায় সাইফ তুলে নেন ফিফটি। ভিসাভাদিয়ার বলে ফিরলে বিপদে পড়ে দল। আফিফ কিছুটা সামাল দেওয়ার চেষ্টা করেছিলেন আমিনুল ইসলামকে নিয়ে, তবে সেটা বেশিক্ষণ স্থায়ী হয়নি। ১০ ওভার বাকি থাকতেই সব উইকেট খুইয়ে ফেলে বাংলাদেশ। ভিলজোয়েন একাই ৫ উইকেট নিয়ে ধসিয়ে দিয়েছেন বাংলাদেশের মিডল অর্ডার[24][25]। "

We used 1000 post and comments related to the sports, political and other topics. Table I is the sample of our collected sports news. In Table II, we have given some sample post and their source including their category. The sample dictionary keywords are given in Table III. The entire process is given in fig 2.

B. New Filter for Text Classification

The dictionary keywords related to the political and sports and the stopword set that we created in the training phase as we shown in figure 1.In order to predict the class level of a post as well as sentiment,first of all we have to choose a classification method.We used Naive bayes classification to find the probability of the class level. For testing we extract each of the word from a test text.After removing the stopwords we have done stemming process.It should be mentioned that the stoprwords are such kind of words that are not carrying any importance of meaning i.e conjunction. After the total stemming process words back to their own root word.then we count them and done their frequency checked.With the bayes theorem we found the probability of individual sports and political matching and nonmatching words with the dictonary keywords. Comparing two probabilities , we can predict the class level.

C. Tokenization

Tokenization is a process by which a text can be fractured and eliminnate white space , newlines , puncuations , digits and so on . It separates each word for the convenience of counting frequency as they have no importance but they only carry the weight of text. We also remove single words from our collected Bengali news document. Bengali verb, adjective and

noun are infected. So after tokenization we reform them to their root word through streaming.

TABLE II. SAMPLE COLLECTION OF DATA

	Data	Source	Data types
1	৫ ফেব্রুয়ারি থেকে শুরু হবে ঢাকা প্রিমিয়ার ডিভিশন ক্রিকেট লিগ। এ লিগের জন্য ‘প্লেয়ার বাই চয়েস’হবে ২০ জানুয়ারি। প্রত্যেক ক্লাব ৫ জন করে পুরোনো ক্রিকেটার ধরে রাখতে পারবে, তবে সেটা ‘এ প্লাস’ বা ‘এ’ ক্যাটাগরির ক্রিকেটারদের নয়। এরকম ১২ জন ক্রিকেটারের তালিকা প্রকাশ করবে বিসিবি, যারা ড্রাফটের জন্য উন্মুক্ত থাকবেন। সিসিডিএম চেয়ারম্যান কাজী ইনাম ক্লাব ও বিভাগভিত্তিক দলগুলোর প্রতিনিধিদের সঙ্গে সভার পর আজ জানিয়েছেন এসব তথ্য[20][21].	খেলার খবর	news
2	জনঅধিকার, সংবিধান ও গণতন্ত্র আজ প্রহসনে পরিণত হয়েছে। আইনের শাসন ও গণতন্ত্রের অভাবে জনগণই আজ সবচেয়ে বেশি শক্তিহীন হয়ে পড়েছে। অথচ এর বদলে সব ধরনের ক্ষমতায়ন হয়েছে দুর্বৃত্তদের। আর এর প্রায় পুরোটাই আজ রাজনৈতিক দুর্বৃত্তদের কব্জায়। যে দল ক্ষমতায় যায় রাষ্ট্রের সব শক্তিকে ব্যবহার করে তারাই হয়ে যায় একেকটা মহাদুর্বৃত্ত। সেখানে কদিন আগের ক্ষমতায় থাকা আরেক মহাদুর্বৃত্তও হয়ে পড়ে সবচেয়ে অসহায়। এ যেন আজব এক যুদ্ধ বিজয়ের মতো অবস্থা।	Facebook	Comments
3	প্রধানমন্ত্রী বলেন, সরকারে ধারাবহিকতা আছে বলেই উন্নয়নের ছোয়া লেগেছে। উন্নয়নের জন্যই আওয়ামী লীগে ভোট চাই। আপনাদের কাছে ওয়াদা চাই।	বাংলাদেশ আওয়ামী লীগ	news
4	বাসা থেকে প্রাইভেট পড়তে বেরিয়ে নিখোঁজ হওয়া শ্রীকান্ত দেব নাথকে পাঁচ দিন পর উদ্ধার করেছে নগর গোয়েন্দা পুলিশ। শুক্রবার দিনগত রাত দেড়টায় নগরীর কর্ণফুলী থানাধীন মইজ্জ্যারটেক এলাকায় ঢাকাগামী একটি বাস থেকে তাকে উদ্ধার করা হয়।	BD news 24	news

TABLE III: SAMPLE DICTIONARY KEYWORDS

Sports	ক্রিকেট	কোচ
	টুর্নামেন্ট	মাঠ
	খেলোয়াড়	খেলা
	রাউন্ড	ফিল্ডিং
Political	সরকার	ক্ষমতা
	জনগণ	রাজনৈতিক
	গণতন্ত্র	উন্নয়ন
	বিরোধী	নেতা

D. Stemming

To reduce the data size of the dictionary,stemming is the best procces. In this process a word is shorten as well is transform its root word without changing its morphological meaning.In our entire system, stemming is playing the most important part. Before stemming words were like this structure that is Shown in table IV. After stemming the result is shown in table V.

TABLE IV. BEFORE STEMMING

আনুষ্ঠানিক" "দল" "ক্যাম্প" "মূলত" "দল" "বোলার" "বল" "ব্যাটিং" "অনুশীলন" "অনুশীলন" "মুখোমুখি" "সংবাদ" "মাধ্যম"

TABLE V. AFTER STEMMING

"আনুষ্ঠানিকভাবে" "দলের" "ক্যাম্পে" "মূলত" "বোলারদের" "বোলাররা" "বল" "ব্যাটিং" "অনুশীলনও" "অনুশীলন" "মুখোমুখি" "সংবাদ" "মাধ্যমের"

E. Naive Bayes classifier

From Naïve Bayes theorem we measure the probability that helps us to do the classification. With the probability theory we measure matching words with the dictionary and count the frequency of words [1].Naive Bayes theorem is a probabilisticclassifier and mostly used classifier for text categorigation.[2]From Bayes theorem, we can compute the maximum a-posteriori probability (MAP) for the data. If Bayes rule is applied to a document *d* and a class *c*, then we can get

$$P_{MAP} = a\text{rgmax}_{c \in C} P(c|d) \qquad (1)$$

P (d |c): Probability of d given c.

P (c_j) can be estimated from the frequency of classes in the training examples and P(x_1, x_2.... x_n|c_j) can only be estimated if a very large number of training example is available. Here, we have applied Multinomial Naïve Bayes Classifiers (c_{NB}) to classify a text.

$$P_{NB} = a\text{rgmax}_{c \in C} P(c_j) \prod_i P(x_i|c_j) \qquad (2)$$

$$P_{NB} = \text{argmax}_{c \in C} P(c_j) P(x_1 = \text{“our”} | c_j) \ldots . P(x_n = \text{“text”} | c_j) \quad (3)$$

We can extract vocabulary from the training corpus and calculate the required P (cj) and P (xk| ck) for each cj in C as

$$P(x_k | c_j) = \frac{n_k + 1}{(n + |\text{vocabulary}|)} \quad (4)$$

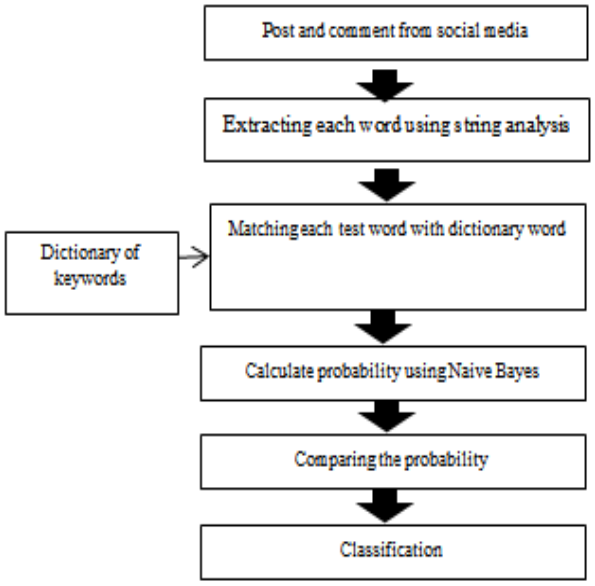


Fig.2. Entire process of our testing phase for Naïve Bayes

F. Support Vector Machine

Support Vector Machine [18] classifies data into a two class by using hyperplane. So it is a binary classifier. Data points which are closer to the hyperplane is called Support vector and it influence the location and direction of the hyperplane. Classifiers margin are maximize by using the support vector. The location of the hyperplane will change by removing the support vectors. These points benefited us to build our SVM. We extract the features and divide it into training data and testing data. 90% of the data is used for training and the rest10% is used for testing.

G. Neural Network classifier

To solve a specific problem neural network is used .It is classifier that is consist of a huge number of highly interconnected elements. In this research, Back propagation Multilayer Perceptron is used. The news classification learning methodology is given in Figure 3. It contains Preprocessing, Feature Extraction, Classification and Evaluation Phases. The experimental parameters learning rate alpha (ε) for the Neural network are used to determine the optimal results/outputs.

Framework for various machine earning technique is given in fig 3.

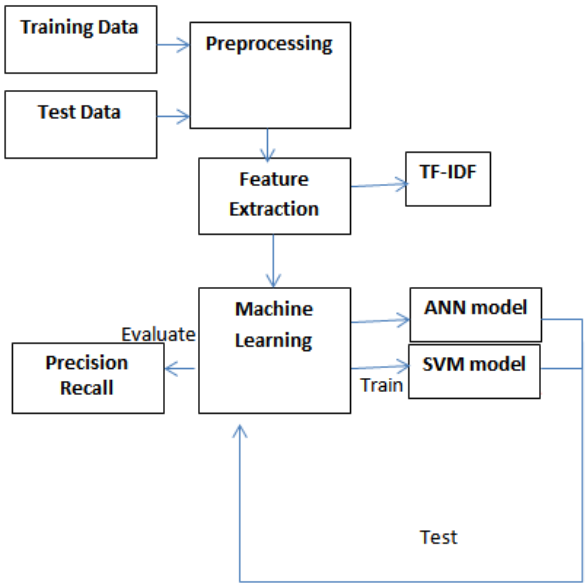


Fig.3. Framework for various machines learning technique

III. RESULT AND DISCUSSION

A. Evaluation Metrics:

To evaluate the accuracy of the system, we used recall and precision as a common Information Retrieval (IR) measure. Theoretically we can define them as follows:

Precision, P: Precision is the ratio of the number of items which define exactly by classifiers to the number of items which classifiers defined to this category, so the precision of is defined as following:

Precision;

$$P = \frac{items\ correctly\ identified}{items\ correctly\ identified + items\ incorrectly\ labeled\ as\ not\ items} \quad (5)$$

Recall, R: Recall rate is the ratio of the number of items which define exactly by classifiers to the number of items which are this category in fact, so the recall rate is defined as following:

Recall;

$$R = \frac{items\ correctly\ identified}{items\ correctly\ identified + individuals\ incorrectly\ labeled\ as\ items} \quad (6)$$

$$F1\text{-}Measure := 2 * \frac{R * P}{R + P} \quad (7)$$

B. Quantitative Analysis

We have tested 500 posts which were collected randomly from news portal and social media sites. Among those 200 posts are sports, 200 are political and 100 are random posts. The overall accuracy is 90%. Table VI shows the sample of class level prediction and table VII shows the testing dataset.

TABLE VI .SAMPLE OF CLASS LEVEL PREDICTION

S.No	Post	Matched probability	Non matched probability	Class

1	"ছুটি শেষে আনুষ্ঠানিকভাবে আজই বাংলাদেশ দলের ক্যাম্পে যোগ দিয়েছেন প্রথম দিন মূলত বোলারদের নিয়েই কাজ করেছেন বোলাররা তো বল করেছেনই ব্যাটিং অনুশীলনও করে নিয়েছেন হ্যালসল আজ অনুশীলন শেষেই মুখোমুখি হলেন সংবাদ মাধ্যমের[20][21] "	0.4985714	0.3333333	Sports
2	ক্ষমতা চিরস্থায়ী নয়।কারো না কারো গ্যারাকলে এটার জবাব তখনই দিতে হয়।বাংলার মানুষ নতুনত্বের পুজারী। বর্তমান যে সংকটে দাঁড়িয়ে আছে দেশ,সে সংকট হতে মুক্তির জন্য সাধারণ ভুক্তভোগীরা অবশ্যই তৃতীয় শক্তিকে স্বাগত জানাতে ভুল করবেনা।	0.5571321	0.2543451	Political
3	তমাল চন্দ্র দে (২৫) নামে এক শিক্ষার্থীকে এসিড নিক্ষেপের অভিযোগে দায়ের হওয়া মামলার আসামি মৌমিতা দত্ত অ্যানি ও তার স্বামী সুমিত দাশকে গ্রেফতার করেছে নগর গোয়েন্দা পুলিশ।	0.1212121	0.5613421	Others

TABLE VII. Testing Dataset

Sl.No	Domain	No. of Post	No of words in post
1	Sports	200	1500
2	political	200	2500
3	others	100	1000

As define above this research focused on classification of news for two domains namely sports and political news so we trained out the test set of document against each of them. Precision and recall were determined by looking at the number of documents that were correctly labeled for a particular section. We have tested 500 posts. The accuracy of the systems are given in table VIII, table IX and X in terms of Precision and recall. The experimental setup is done in two observations with variation in train and test splits. The experimental results are shown below.

B.1. Experiment results with Naive Bayes

Since From Naïve Bayes theorem we measure the probability that helps us to do the classification. With the probability theorem we measure matching words with the dictionary. The result obtained from experiment for Naïve Bayes is shown in Table VIII.

TABLE VIII. Experimental Result for Naïve Bayes

	Accuracy	Precision	Recall	F-Score
Sports	88	89	88	87
Political	89	80	89	88

B.2 Experiment results with SVM

The result obtained for SVM is given in TABLE IX

TABLE IX. Experimental Result for SVM

Exp.	Accuracy	Precision	Recall	F-Score
Sports	76	73	72	72
Political	72	74	73	72

B.3. Experiment results with Neural Network

The result obtained from experiments for Neural Network is shown in TABLE X.

TABLE X .Experimental Result for ANN

Exp.	Accuracy	Precision	Recall	F-Score
Sports	82	83	83	82
Political	83	84	84	83

To investigate the efficiency of results; a comparison has been made between Naïve bays, SVM and ANN. Fig.4 and fig 5.shows the results of all the three systems, including, Naïve Bayes, SVM and ANN. From the overall point of view, Naïve

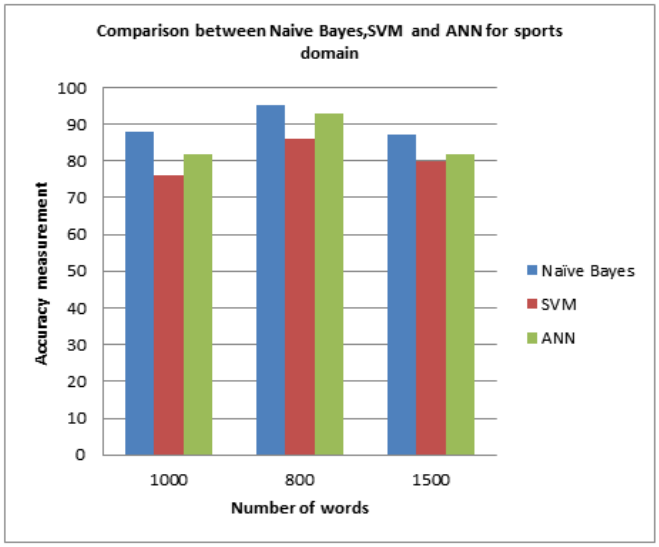


Fig.4. Comparison between Naive Bayes, SVM and ANN method for sports

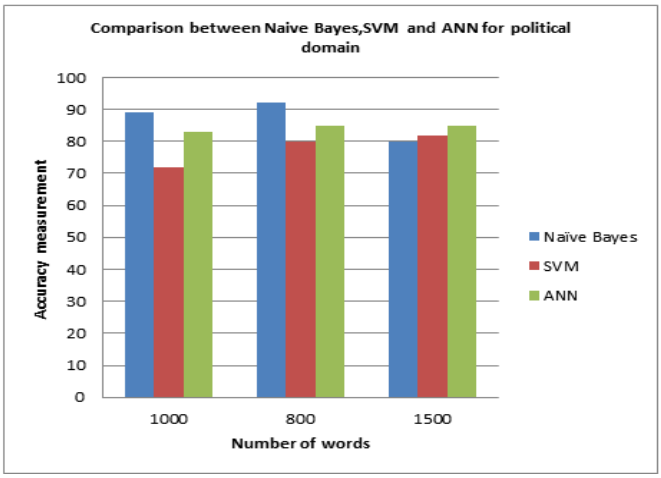


Fig.5. Comparison between Naive Bayes, SVM and ANN method for political news

IV. CONCLUSION AND FUTURE WORK

Due to rapid growth of social media and online news portal, it's an ultimate necessary of filtering to find the relevant news in a shortest way. For this a proper classification need to be used to solve the heterogeneity of categorization. In our paper we used Naïve Bayes classification that based on probabilistic theory to find a satisfactory result. The appropriateness of our classification system mainly reclines on the data set of dictionary and stemming process as well as training phase. Because of stemming we reduced our dictionary size. The accuracy can be different if we use different classification method further. This classification only can be possible if the post or comments are written in Bengali. So correct spelling should be needed for post as well as comments. As we only show the text based classification; adding image that contains information can mislead to find proper result and correct result can be varies to identify exact probability. So text conversion of images also an ultimate necessary to find an exact classification. Therefore, for better filtering as well as rapid classification, other n-gram techniques are also being a useful feature. The future recommendations on the development of this work are extract textual data from image and apply CNN based approach to solve the Bangla new classification problem.

REFERENCES

[1] A.Chy, M.Seddiqui and S.Das. "Bangla news classification using naive Bayes classifier". 16th Int'l Conf. Computer and Information Technology.(2014)

[2] T.Islam, A.Bappy , T.Rahman and L.Uddin (2016). "Filtering political sentiment in social media from textual information".5th International Conference on Informatics, Electronics and Vision (ICIEV)2016.

[3] M.Seddiqui ,AAM.Maruf and A. Chy,"Recursive Suffix Stripping to Augment Bangla Stemmer." 1st Int'l Conf. on Advanced Information and Communication Technology(ICAICT)2016.

[4] M.Mansur, N.Uzzaman N and M. Khan. "Analysis of -Gram Based Text Categorization for Bangla in a newspaper Corpus ".Center for Research on Bangla Language Processing, BRAC University, Dhaka, Bangladesh.

[5] Y.Zhou,Y. Li and S.Xia ."An Improved KNN Text Classification Algorithm Based on Clustering". Journal of Computers, 4(3). (2006)

[6] G.Kaur, K.BAJAJ , "News Classification and Its Technique : A Review" , IOSR JOURNAL OF COMPUTER ENGINEERING (IOSR - JCE) E-ISSN: 2278-0661,P-ISSN: 2278-8727, VOLUME 18, ISSUE 1, VER. III (JAN – FEB. 2016), PP 22-26.

[7] D. Lewis, "Naive (Bayes) at forty: The independence assumption in information retrieval," Machine Learrning: ECML-98, PP. 4–15, 1998

[8] CH.Chan, A.Sun and E.Penglim Automated online news classification with Personalization", Proceddings of the 4th International Conference of asian Digital Library (ICADL)2001, ,Bangalore, India

[9] A.Pak ,P. Paroubek ."Twitter as a Corpus for Sentiment Analysis and Opinion Mining", Language Resources and Evaluation Conference (LREc), 2010 ,vol. 10, pp.1320-1326.

[10] E.Kouloumpis, T.Wilson, J.Moore "Twitter Sentiment Analysis:The Good the Bad and the OMG!", .Proceedings of the fifth International AAAI Conference on Weblogs and Social Media .2011, vol. 11, pp. 538-541.

[11] A.Agarwal, B.Xie, I.Vovsha and R.Passonneau "Sentiment Analysis of Twitter Data". Proceedings of the workshop on languages in social media, (LSM) Association for Computational Linguistics, pp. 33-38, 2011.

[12] V.Metsis, I. Androutsopoulos, G.Paliouras "Spam Filtering with Naive Bayes – Which Naive Bayes" (CEAS), Mountain View, California USA,2006.

[13] P.Domingos, M.Pazzani " On the Optimality of the Simple Bayesian Classifier under Zero-One Loss".

[14] Machine learning ,Kluwer Academic Publishers. Manufactured in the Netherlands.1997.

[15] DM. Rennie, L.Shih, J.Teevan and D. Karger."Tackling the Poor Assumptions of Naive Bayes Text Classiers".

[16] International Conference on Machine Learning (ICML), Washington DC.

[17] F.Urnkranz "A study using n-gram features for text categorization", Austrian Research Institute for Artificial Intelligence.1998

[18] W.Wang, and B.Yu."Text categorization based on combination of modified back propagation neural network and latent semantic analysis". Neural Computing and Applications, 18(8), pp.875-881.

[19] DailyAmaderSomoy.[Online]Available:http://dainikamadershomoy.com /todays-paper [Accessed: 11-Jan2018].

[20] Prothom Alo. [Online] Available: http://prothi.com/politics/article/ [Accessed:12-Jan2018].

[21] Daily Inqilab. [Online] Available: http://samakal.com/politics/article/ [Accessed:12-Jan2018].

[22] Samakal. [Online] Available: http://samakal.com/politics/article/ [Accessed:12-Jan2018].

[23] BanglaTribune.[Online]Available:http://www.banglatribune.com/sport/news/[Accessed:20-Jan2018].

[24] BDCricketTIme.[Online]Available https://bn.bdcrictime.com

[25] Accessed:20-Jan2018].

[26] Newspapers71.com..[Online]Available:http://www.newspapers71.com [Accessed:20-Jan2018].

Sentiment Analysis & Emotion Extraction: A Review of Research Paradigm

Lutfun Nahar
Assistant Professor
Dept. of Computer Science & Engineering
International Islamic University Chittagong
Email:lutfacsecu@gmail.com

Zinnia Sultana
Assistant Professor
Dept. of Computer Science & Engineering
International Islamic University Chittagong
Email:zinniaiiuc@yahoo.com

Farzana Tasnim
B.Sc. in Computer Science & Engineering
International Islamic University Chittagong
Email:farzanatasnim34@gmail.com

Afifa Chowdhury
B.Sc. in Computer Science & Engineering
International Islamic University Chittagong
Email:afifamim06cseiiuc@gmail.com

***Abstract*—** Micro-blogging is one of the most popular social network services that enable users to share their thoughts, ideas and discuss current issues every day. Some popular micro-blogging sites are Tumblr, Twitter, Pinterest, Flattr, Meetme. This research will work to represent the study of sentiment analysis to extract emotions from textual contents by using several methods of tracking mode. Several classifiers are trained to extract basic moods such as Joy, Trust, Fear, Surprise, Sadness, Disgust, Anger, and Anticipate from text. The aspect of this research is to give a nearly concept about sentiment analysis techniques and the related areas.

Keyword—micro-blog; sentiment analysis; emotion extraction; emotion.

I. INTRODUCTION

Among all users, micro-blogging has become one of the most well-liked social networking applications. It plays a pertinent role in providing information. Lots of messages appear daily in popular sites where users can share their feelings, views etc. Various systems have developed to detect user's opinions or sentiments by processing natural languages on some micro-blogging platform. Based on the 20 latest articles where different approach is used , the contribution of this survey is- it provides futuristic categorization, which will help researchers of Sentiment Analysis (SA) to use appropriate techniques in certain field. As various SA approaches are categorized here, new comers can have an extensive view on entire SA field.

This section focuses on sentiment analysis and it's classification as well as detection method. Sentiment or opinion Analysis is a computational study that refers to process natural languages and analyzes texts to extract and determine people's attitudes and emotions. Sentimental analysis can be classified into two types- Machine learning approach and Lexicon based approach. Figure 1 shows the Sentiment Classification methods.

A. Machine learning approach

Machine learning approach uses linguistic features and appliesvarious famous algorithms on these for analyzing their sentiments. J. Bollen *et. al.* and Wenbo Wang *et. al.* have used this approach for their research. Machine learning approach can be divided into 2 classes.

A.1. Supervised learning

A supervised learning problem can classify sentiment can with two labels (i.e. positive and negative). Input dataset is considered to be a labeled one when machine learning approaches are used to classify sentiment. Dmitry Davidov *et. al.* has used this approach. Various types of supervised learning classifiers are used for classifying sentiment like Naive Bayes, Maximum Entropy, Stochastic Gradient Descent, Support Vector Machine, Decision Tree, K-nearest neighbour, Artificial Neural Network and n-gram method etc. *Ze-Jing Chuang et. al.* and Agarwal *et. al.* have used SVM classifier. Alexander Pak *et. al.* has used Naive Bayes classifier. John *et. al.* has used Maximum Entropy classifier.

A.2.Unsupervised learning

An unsupervised learning approach classifies sentiment based on some fixed words and phrases. Datasets are consisting of input data without labeled resources.

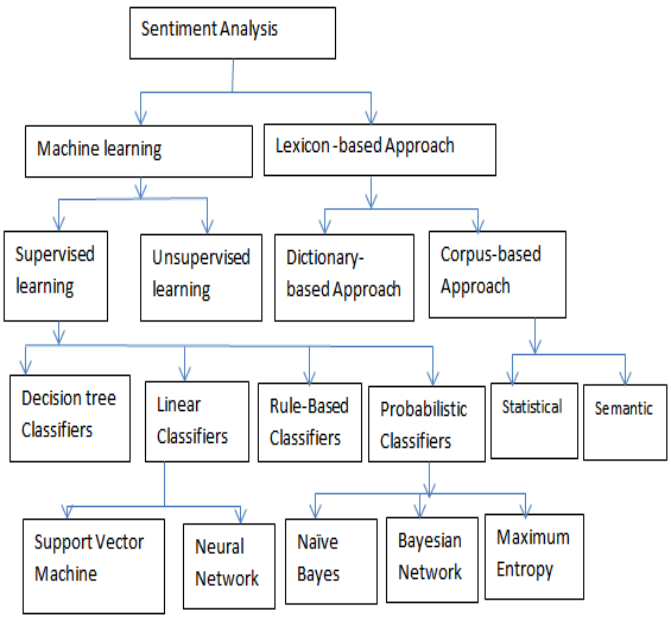


Fig.1. *Sentiment classification methods [1]*

B. Lexicon based approach

Lexicon based approach depends on reliable and organized sentiment labels. Opinion words like positive words or negative words are used to classify sentiment. Lexicon based approach can be divide into 2 classes.

B.1. Dictionary based approach

This simple approach is based on bootstrapping and uses small set of simple words or online dictionary such as WordNet, thesaurus to classify sentiment.

B.2 .Corpus based approach

This approach uses a list of words and words pattern to find other words in corpus with the help of context specification orientation [1]. Carlo Strapparava *et. al.* has used this approach for her research.

There are another method known as Hybrid approach which is the combination of both machine learning approach and lexicon based approach [2].Again SA can be carried out in three basic levels. These are sentence level, document level and aspect level SA. Document level SA is used to categorize an opinion or sentiment as positive or negative. Sentence level SA is used to discover whether the sentence is biased or aimed and then find out whether the verdict expresses positive or negative sentiment. Aspect level SA is used to recognize the entities and then determine their aspects [3].In the last few years, the field of Sentiment Analysis has enhanced with various algorithms that were proposed and developed. These fields comprise Emotion Detection (ED), Building Resources (BR) and Transfer Learning (TL). Emotion detection is used to detect and analyze explicit or implicit emotions of users in the sentences to identify user's emotional state such as anger, confusion or joy across both voice and non-voice channels. Building resources focuses to create data set, corpora in which emotions of users are annotated based on their divergence. Transfer learning classification is used to analyze data from one particular field and then use the consequences in marked field.

C. Emotion Detection Techniques

There are basically four techniques for emotion detection including Keyword spotting technique, Lexical affinity method, Learning based method and Hybrid method.

C.1 Keyword spotting

This technique can be defined on the basis of some definite predefined keywords which have been classified into some categories such as surprised, angry, sad, joy and disgusted. This technique analyzes the intensity of words to identify emotion. The process is shown in the fig. 2

C.2 Lexical Affinity

Lexical Affinity approach works with probabilistic 'affinity' for a specific opinion to random words spaced out from picking up sentimental keywords.

C.3 Learning and Hybrid method

Learning based method is used to categorize the input string into different emotions. Learning based methodology and keyword based spotting methodology is combined to create Hybrid methods .

Thus the present section introduces the problem in this work while methodology review is elaborated in the Section II. And Section III concludes the paper.

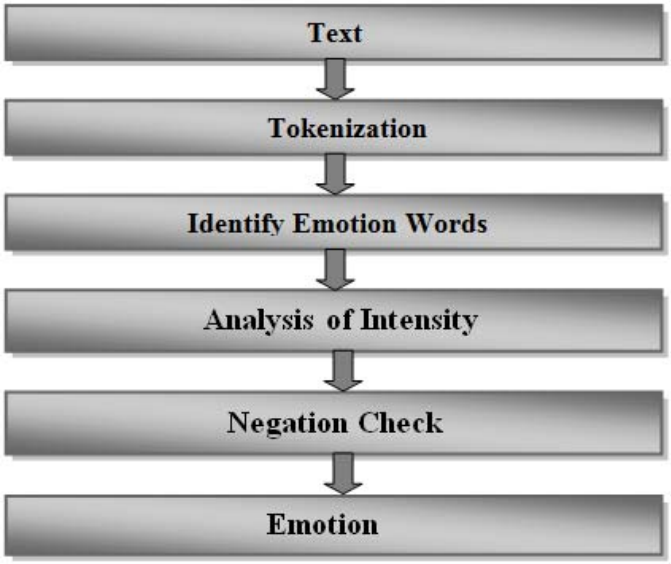


Fig 2. *Keyword Spotting Techniques [2]*

II. METHODOLOGY REVIEW

In 2002, X. Zhe *et. al.* [4] worked on online textual information and build a machine that can extract emotion. The engine is able to analyze a given string from online conversation and extract the emotion. The output is delivered to the communicating user display. There are three functions of the engine: Input analysis function, tagging system and the parser. The operation of the engine is showed in Fig. 3.

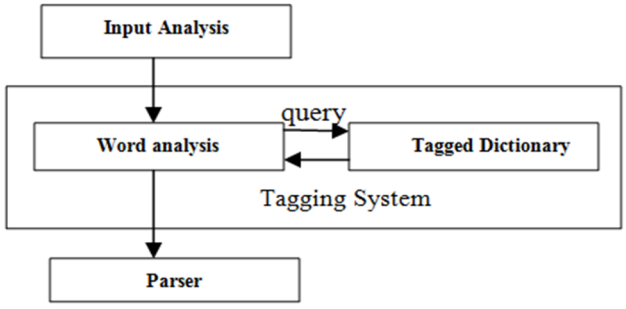


Fig. 3. *The working flow of engine [4]*

Semantic analysis is used to recognize the emotions of text. Two types of tests are used to test the engine performance. They are static and dynamic. Static test construct emotional sentences artificially. Dynamic test enables users to input any sentences freely. Here, the emotion extraction engine requires extremely low bandwidth for real time communications.

Therefore, image transmission is avoided here because of large bandwidth.

In 2003, Tetsuya Nasukawa *et. al.* [5] have demonstrated a sentiment analysis on some organizations and their products to extract positive or negative sentiment for particular subjects from a textual context rather than categorizing the entire content into positive or negative towards the subject. The key issue of this task is to recognize the semantic relationships between subjects and sentiment-related terms because the polarization of the emotion may be wholly different depending on the relationships. They define the framework of the sentiment lexicon. They had 3,513 records in the sentiment analysis lexicon, as shown in Table 1.

TABLE I. DISTRIBUTION OF SENTIMENT TERMS

POS	Total	Positive	Negative	Neutral
Adjective	2465	969	1495	1
Adverb	6	1	4	1
Noun	576	179	388	9
Sentiment verb	357	103	252	2
Transfer verb	109			

POS tagging is used to disambiguate some polysemous expressions and syntactic parsing is used to find out the relationships between sentiment expressions and the subject term. The framework is not suitable to deal with the long and complex sentences because of negating the local sentiment for the whole.

In 2004, *Ze-Jing Chuang et. al.* [6] have recognized emotion from voice signals and textual content. They extract thirty-three visual features in the analysis of voice signals from the voice signal input. . Principle component analysis (PCA) is used to select most significant and Support Vector Machines (SVM) are applied to classify the sentimental state of speech input and fed into the textual emotion recognition module. They define all sentiment modification words and emotional keywords physically. The final result is regulated from the sentiment outputs from the visual and textual approaches. The final result shows that the multi-modal approach gives an optimistic result for emotion extraction than uni-modal strategy.

In 2005, Xiao *et. al.* [7] have extracted and analyzed a feature set by "Fisher's Discrimination Ratio" and selected few best features that are computed from F0-fundamental frequency with range between 60-450 Hz. Including the feature list and selection result, they have introduced a new definition of emotions as 3-states emotion with active, passive and neutral state. They have also built a database to verify states of emotional speech with speech samples. In this paper fifty feature are used to analyze and they also examined voice samples separately for male and female and finally sadness, anger and disgust these tree emotion are categorized. Those confused emotions should be found out in our future work such as neutral and fear, boredom .Their future work will include the description about sub-states and difference between sub-states. For verifying the states definition of emotional speech a new database is being built.

In 2008, Alastair J. Gill *et. al.* [8] study the capability of naive raters of sentiment to identify eight emotional categories (joy, anger, fear, sadness, surprise, disgust acceptance, anticipation) from a real data set and use 200 samples. Principles of CMC and potential applications in HCI is used to range from recognizing emotion from blog posts to allow dynamically adaptive interfaces.. Contributors used the activation-evaluation wheel exposed in Fig. 4. There do not apply any machine learning classification, which sets a challenges in future work for possible judging naive human including emotion monitoring of blog posts, or dynamic interfaces.

They unravel the text-based communication of emotion which is built upon Hancock *et al.* [9] basically in three ways: (1) they develop their classification of emotion into the eight main emotion categories from positive and negative sensation. (2) they emphasis on exactly classification of emotion based on asynchronous short blog post excerpts of 50 and 200 words, resulting from (3) real sentimental blogs.

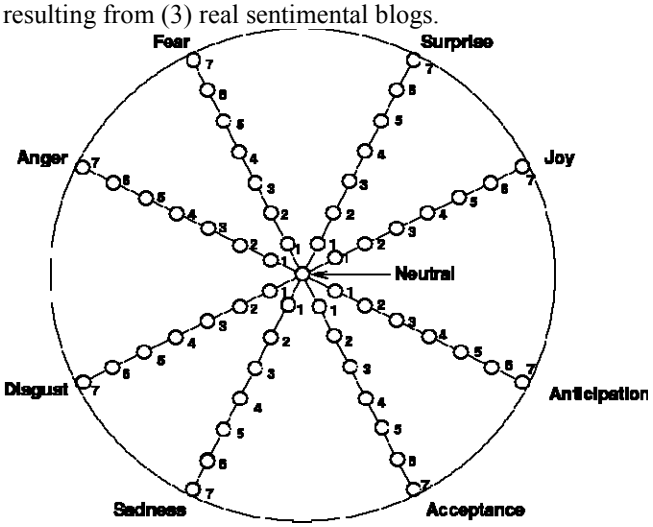


Fig. 4. *Emotion wheel used for text rating [8].*

In 2008, Carlo Strapparava *et. al.* [10] have extracted emotion of news headlines from newscast sites. Headlines are normally made of a few words and are often written by inventive persons to seek the reader attention. That's why a short sentence may represent the emotional features which help to create this type of text predominantly appropriate for use in an automatic sentiment identification system. The lexicon contains news headlines taken from popular newspapers. Two types of lexicon were available one is progress lexicon containing 250 annotated headlines and another is a test lexicon containing 1,000 annotated headlines. A Web-based annotation interface was established to display one headline to perform the annotations at a time, together with six slide bars for sentiments and one slide bar for valence. Inter-tagger agreement studies are conducted for each of the six emotions. Using the Pearson correlation measure, the agreement calculations were carried out which is shown in Table 2.

TABLE II. PEARSON CORRELATION FOR INTER-ANNOTATOR AGREEMENT

Emotion	Detection in %
Anger	49.55
Disgust	44.51
Fear	63.81
Joy	59.91
Sadness	68.19
Surprise	36.07

They carried out on a big lexicon of 1,000 deadlines through relational estimations of a number of knowledge-based and corpus-based techniques. Future work is to implement deeper semantic processing of the text and analyze the emotions lexical structure in the knowledge-based and corpus based classification.

In 2009, John *et. al.* [11] have built a system that will classify the emotions of "the word of interest" of a tweet with "Maximum Entropy Classifier" and if the emotions related to "the word of interest" are different from the emotions of tweets, this system will try to predict the emotion of "the word of interest". Though their classifier was able to classify emotions versus neutral texts of twitter over 70% of time but they did not get enough data to distinguish between negative and positive emotions. In Twitter their classification model is capable to identify emotion vs neutral text by selecting 3 categories of emotion such as positive, negative and neutral. They also believed that for a big corpus of data they will get an excellent performance and also include 50% data on 500 tweets that are labeled is perfect for optimism.

TABLE III. DATA SETS USED IN CLASSIFICATION

Data	Description
3-way	Compares neutral, positive and negative in a three way classifier.
+,#	Compares emotion present versus no emotion (2-way).
+,-	Compares positive emotion to negative emotion (2-way).

In 2009, Go *et. al.* [12] have shown that using emoticons as noisy labels with tweets, is an effective way for performing distant supervised learning. They have developed a novel approach using "Distant Supervision" to classify the sentiment of twitter messages automatically as either negative or positive with respect to query and achieved high accuracy when using this method. Their method could use tweets with emoticons to identify opinions of individuals for distant supervised learning. As features, they have explored the uses of parts of speech, unigrams, bigrams, unigrams and bigrams. The accuracy of their classifier is shown below:

TABLE IV. CLASSIFIER ACCURACY

Features	Keyword	Naïve Bayes	MaxEnt	SVM
Unigram	65.2	81.3	80.5	82.2
Bigram	N/A	81.6	79.1	78.8
Unigram +Bigram	N/A	82.7	83.0	81.6
Unigram + POS	N/A	79.9	79.9	81.9

There works with English sentences, in future use this approach to classify sentiment in other languages and attention will be paid to neutral sentiment.

In 2010, Dmitry Davidov *et. al.* [13] proposed a supervised sentiment classification framework is in this paper which is based on twitter data.50 Twitter tags and 15 smiley's are used here. They consider dependencies and overlap between different smileys and Twitter hash tags sentiment. They assessed diverse feature types for sentiment exaction including punctuation, words, patterns and n-grams and assuring that each feature type come up with the sentiment classification framework. They recommended two different methods for automatic reorganization of sentiment type overlap. The first is tag co-occurrence and the second type is feature overlap. In future these methods will be applied for automated sentiment clustering and sentiment dependency rules and multi-class sentiment identification and classification in other domains of Twitter data .They construct an engine which can identify and classify different types of sentiment automatically for the text which are not so long. They used supervised classification with hash tags and smileys as training labels. They considered different types of feature for sentiment extraction such as words, n-grams, punctuation and patterns to confirm that each type of feature is used for the classification of sentiment. By using hashtag and different algorithm the accuracy are given below

TABLE V. ACCURACY OF ALGORITHM

Algorithm	Data Set	Accuracy	gc	%above guessing
MaxEnt I	3-way	47.60%	5.1%-	14%
MaxEnt II	3-way	41.09%	5.9%	7.76%
MaxEnt II	=.#	72.79%	5.1%	22.79%
MaxEnt II	+.-	51.08%	4.5%	1.08%
MaxL N	+.-	55.91%	7.2%	4.91%

In 2010, Alexander Pak *et. al.* [14] carried out semantic analysis of the composed corpus and describe discovered phenomena in their research. Multinomial Naive Bayes classifier is used to build a sentiment classifier that uses POS-tags and also N-gram features, using the corpus which is capable to define positive, negative and neutral opinions of corpus. To gather an input with positive and negative opinion they present a method in which no human exertion is desired for categorizing the documents. They also perform statistical semantic analysis of the collected input and bearing experimental evaluations on real micro blogging posts. The size of the collected corpora is not restricted. POS-tagging is used to build tree Tagger and discovered the variance in flows among negative, positive and neutral sets. Syntactic structures are used to describe emotions or state facts. Their work is to gather a multilingual corpus of Twitter information and compare the attributes of the corpus crosswise over various languages and also implement a multilingual negative and neutral sets their classifier is capable to fix three sentiments like positive .negative and neutral of documents.

In 2011, Agarwal *et. al.* [15] have worked to analyze the sentiment of twitter data and for this purpose they have introduced a feature based model i.e POS-specific prior polarity features and tree kernel and demonstrate their performance on unigram baseline by SVM-Support Vector Machine. They also have reported a gain for two classification tasks- i) positive vs. negative vs. neutral i.e a 3-way task and ii) positive vs. negative i.e a 2-way task. For their experiment they have took those features that merge the preceding schism of words and parts of speech tags in twitter data.

In 2011, J. Bollen *et. al.* [16] have used psychometric instrument to extract six different types of emotions from a sentiment analysis of tweets between august 1 to December 20,2008. For each day in the timeline they evaluate a six dimensional mood vector. Then they compare their results with some media events. Profile of Mood state (POMS) is used to measure the sentiment of each tweets. Their method used an analytic tool rooted in decades of empirical psychometric research which showed a valid replacement of machine learning to determine sentiment and associate its oscillations with a timeline of media events. In this article ,an extended version of the Profile of Mood States (POMS) is used for sentiment analysis of messages published on. We know that political and social event are significantly correlated and economic and cultural also correlated, it fluctuate depend on the level of public mood in a range of different mood of dimensions. Moreover they proposed a substitute to machine learning for the detection of public sentiment.

In 2012, Shiv Naresh Shivhare *et. al.* [17] have proposed a framework which is separated into two main mechanisms: Emotion Ontology and Emotion Detector.

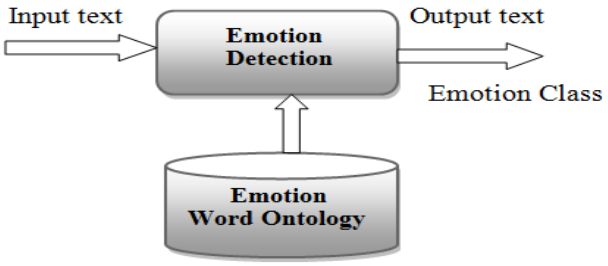


Fig. 5. *Working flow of framework [17]*

Ontologies allow an understanding of particular domain to be connected between publics, application systems and institutions. Sentiment word hierarchy is transformed into ontology. The ontology, which was proposed, contains the relationship layout of sentiment class's and subclasses. High weight age is appointed to the upper level sentiment classes and low weight age is appointed to the lower level sentiment classes. Emotion detection algorithm is used to recognize the emotion of the textual data. The weight for certain emotion is computed by the algorithm by storing weights assigned at each level of hierarchy and same calculation is done for its counter emotion, then both scores are compared and larger one is occupied as the identified emotion. Different parameters are considered to determine weight age to be allocated for several sentiment words so that they can be fixed based on it.

In 2012, Wenbo Wang *et. al.* [18] automatically formed a large emotion-labeled lexicon by hitching sentiment-related hash tags existing in the tweets. Two different types of machine learning algorithms are applied for opinion extraction to learn the efficiency of several features and the effectiveness of the scope of the collected data. They show the grouping of unigrams, bigrams and sentiment bearing words. Using data of 2 million tweets for 7 sentiment categories the uppermost accuracy was attained. They collected 5 million tweets and to verify their quality they arbitrarily selected a set of tweets as sample. They developed a set of filtering heuristics based on the set of collected tweets. They explore various features which are enabling to provide better perceptions for emotion identification. In this paper emotion are identified automatically. They collect more than 2.5 million tweets to cover these seven categories of emotion. This is one of the biggest datasets. The results illustrate that the combination of feature bigrams, unigrams, current sentiment and also emotion lexicons, and POS achieves the higher accuracy which is 65.57%.To identify the fine-grained emotion lexicon based and POS are less effective .And also there was no dataset for not so popular emotions such as surprise and fear etc.

In 2012, Zhang *et. al.* [19] have constructed an emotional vector with hierarchical composition based on taxonomy to act upon opinion analysis. Through this emotional vector algorithm, their system could soak up new emotional words and could perform multi-level analysis for emotional perspectives. Their experimental assessments have showed a well-built connection between bursty actions and public emotions. Their proposed "Emotional Vector Algorithm" analyze sentiment effectively rather than traditional orientation

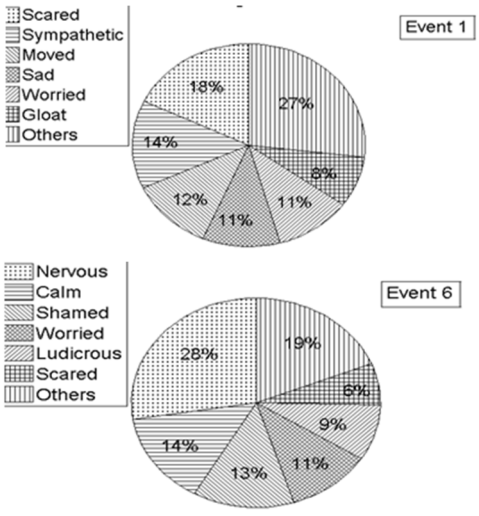


Fig. 6. *Distribution of emotional vector of events [19].*

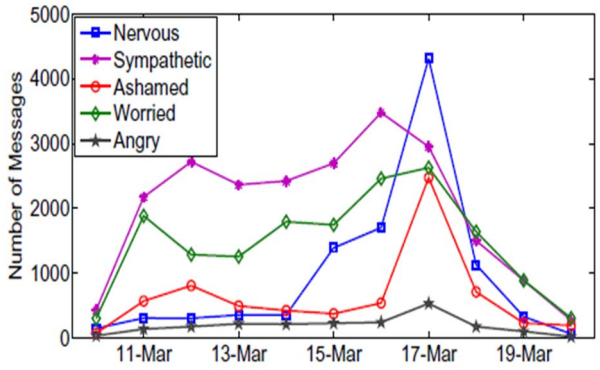


Fig. 7. *Five typical dimensions of emotional vector overtime [19]*

In this paper, they build an emotional vector based on hierarchical structure to achieve sentiment analysis and multi-level analysis is done to show how emotional changes time to time. Their experiment exhibited that their methods could effectively evaluate public moods.

In 2012, Mohammad *et. al.* [20] have created emotion word hashtag type input from twitter texts. He has shown consistent self-labeled hashtag annotations. He also proposed a technique to progress the correctness in the classification of emotion by using tweeter corpus in different domain of micro-blogs. Using TEC classifier, he has extracted emotion based on lexicon from twitter and exhibited that their system performs better than conventional crafted WordNet Affect lexicon in the task of emotion classification and combines labeled data to improve classification accuracy.

In 2012, Robert *et. al.* [21] have introduced a corpus annotated with seven emotions: Love, Sadness, Fear, Joy, Surprise, Anger and Disgust. They have analyzed how emotions are distributed in text, applied a simple supervised method for detecting emotions, compared it with other emotion-annotated corpora and developed a baseline approach for this comparison. For expressing emotions of corpus, they have also characterized the linguistic style as below:

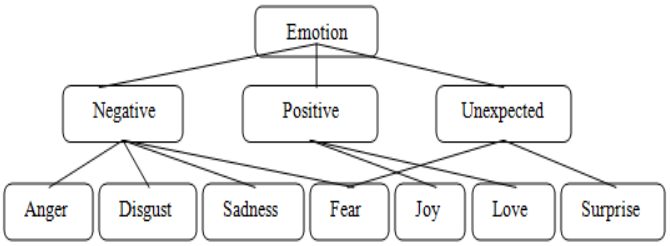


Fig.8. *Emotion ontology for the six Ekman emotions (plus LOVE). Solid lines indicate inheritance [21]*

System used for automatically identifying emotions in the corpus is given below:

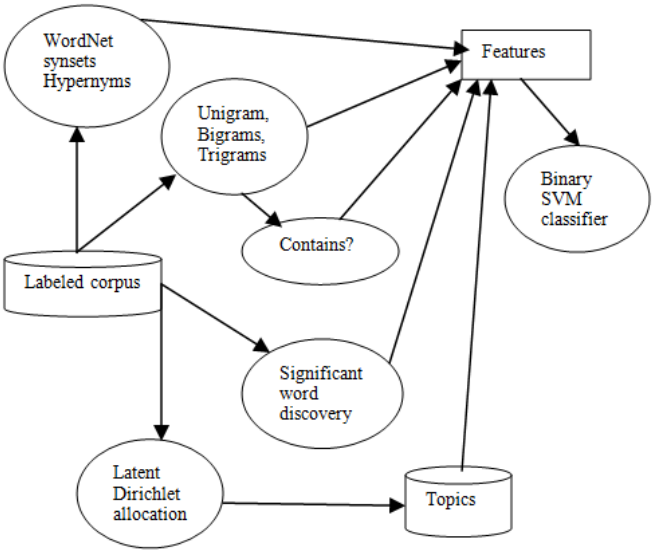


Fig. 9. *System used for automatically identifying emotions in the corpus [21]*

In their paper they made an emotion corpus created from the micro-blogging service Twitter. It categories seven different emotions. They also established a standard approach to use for Bench marking and used then compare their corpus with several present emotion corpora. We have also supervised method is also used for detecting emotions can be trained on this corpus. Moreover this analysis should lead to the design of novel supervised and unsupervised emotion detection technique

In 2013, Neviarouskaya *et. al.* [22] have focuses on phrases that are related to causes of emotion automatically. For this purpose, they have developed a set of phrases with 22 specific emotions from the OCC model. They have introduced a method for detecting linguistic correlations between an emotion and its cause and extracting those, which are related to emotion causes. Their method has employed the enslavement and syntactic information, provided by parser and a system for analyzing 8 types of emotion-cause linguistic relations to extract emotion causes automatically.

In their result section error analysis on the sentence is performed and also illustrates how other method failed to extract correct phrases. Rules are added and modified in the emotion extraction method; they got the following evaluation results: precision in 0.821, recall in 0.852, and F-score in 0.810. In that way, their method performed with about 15% gain in accuracy.
The main objective of their research are the creation of a corpus of emotion causes specific for 22 emotions from the OCC model and the advance of a novel method for extraction of emotion causes from sentences based on the analysis of syntactic and dependency information provided by the parser. causes as positive, negative, or neutral.

In 2014, Li *et. al.* [23] have developed a rule-based system to detect and extract event of emotional post automatically. They have built an emotion corpus which has been labeling by human annotators with Chinese micro-blog-"Weibo" texts.

They have selected features by using Emotion Cause Extraction and trained their classifier to classify emotions using SVR. To detect and extract emotions from original texts, they have designed a rule-based sub-system which generates a set of features for training the classifier. Their experiment has improved traditional emotion classification models with emotion cause extraction technique. Due to multi-class problem, their experiments has ‘‘one-vs-rest’’ manner. it is important for them to compare their system with a baseline system which based on convention feature selection method. They use 75% posts in their corpus as the training dataset, and 25%as the test dataset

In 2014, Hasan *et. al.* [24] have established a novel approach called “EMOTEX” to categorize twitter messages of individuals to assume their emotions. They have used Circumflex model that characterizes experiences with valence and arousal to mock-up emotional states. Their classifier could identify multiple classes of emotion on large data set with no manual endeavor. In their experimental assessments, they have investigated the utility of some features like-unigrams, punctuations, negations and emoticons. Again they have used Lexicon of emotions to solve the sparse problem and high dimensional feature space problem in twitter data set. Their technique has achieved an accuracy of 90% over other machine learning algorithms like- KNN, SVM, Naïve Bayes and Decision Tree for classify twitter messages.
They use a lot of number of hash-tags corresponding for each class of emotion depending to each class of emotions. Four classes of emotions are used here namely, Happy-Inactive, Happy-Active, Unhappy-Active, and Unhappy-Inactive The initial list of keywords that are used listed below: • Class Happy-Active: Happy, Excited, Delighted, Astonished, Aroused • Class Happy-Inactive: Serene, Contented, Satisfied, Relaxed, Calm • Class Unhappy-Active: Tense, Angry, Afraid, Annoyed, Distressed • Class Unhappy-Inactive: miserable, Depressed, Sad, Gloomy. They also compared the accuracy of several machine learning algorithms such as SVM, Naive Bayes, KNN, and Decision Tree for classifying the moods of Twitter messages. This method are not used for other pieces of information such as, sleep data, exercise and physical activities, and food information because of limitation of keyword used

.

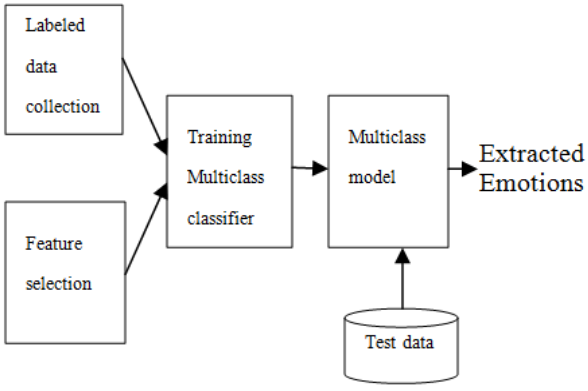


Fig.10. *Model of Emotex [24]*

TABLE VI. CLASS IDENTIFICATION

Class	Number of tweets before pre-processing	Number of tweets after pre-processing
Happy-active	39600	34000
Happy-Inactive	41000	29200
Unhappy-Active	44000	37000
Unhappy-inactive	40700	33900
Total	165300	134100

III. CONCLUSION

This research presents a synopsis on opinion analysis and emotion extraction related fields that use a variety of sentiment analysis techniques such as Machine learning approach and Lexicon based approach for various real-world applications. Twenty of the recently introduced related workings are summarized here. Comparing different Machine learning approach, it is clear that Support Vector Machines Classifiers (SVM) is the most frequently used technique for solving Supervised Learning problem. On the other hand, Semantic approach is the best technique to solve problems of Corpus based in Lexicon based approach. WordNet is the most common lexicon source used in analysis.Natural language process (NLP) is one of the most facilitated tools used in the SA process. It helps to produce more accurate results in SA related problems by processing natural language. In future we will build an emotion extraction method that can automatically detect emotion from tweets from an experience of emotion specified by emotion token and the classification of causes as positive, negative, or neutral. Also collect a multilingual corpus of Twitter data and compare the characteristics of the corpus across different languages.

REFERENCES

[1] B. Liu and L. Zhang, "A survey of opinion mining and sentiment analysis," in Mining text data, ed: Springer, 2012, pp. 415-463.

[2] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," Ain Shams Engineering Journal, vol. 5, pp. 1093-1113, 2014.

[3] A. Tripathy, A. Agrawal, and S. K. Rath, "Classification of sentiment reviews using n-gram machine learning approach," Expert Systems with Applications, vol. 57, pp. 117-126, 2016.

[4] X. Zhe and A. Boucouvalas, "Text-to-emotion engine for real time internet communication," in Proceedings of International Symposium on Communication Systems, Networks and DSPs, 2002, pp. 164-168

[5] T. Nasukawa and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing," in Proceedings of the 2nd international conference on Knowledge capture, 2003, pp. 70-77. [3] A. Tripathy, A. Agrawal, and S. K. Rath, "Classification of sentiment reviews using n-gram machine learning approach," Expert Systems with Applications, vol. 57, pp. 117-126, 2016.

[6] Z.-J. Chuang and C.-H. Wu, "Emotion recognition using acoustic features and textual content," in Multimedia and Expo, 2004. ICME'04. 2004 IEEE International Conference on, 2004, pp. 53-56 [5] T. Nasukawa and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing," in Proceedings of the 2nd international conference on Knowledge capture, 2003, pp. 70-77.

[7] Z. Xiao, E. Dellandrea, W. Dou, and L. Chen, "Features extraction and selection for emotional speech classification," in Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on, 2005, pp. 411-416.

[8] A. J. Gill, D. Gergle, R. M. French, and J. Oberlander, "Emotion rating from short blog texts," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2008, pp. 1121-1124.

[9] J. T. Hancock, C. Landrigan, and C. Silver, "Expressing emotion in text-based communication," in Proceedings of the SIGCHI conference on Human factors in computing systems, 2007, pp. 929-932.

[10] C. Strapparava and R. Mihalcea, "Learning to identify emotions in text," in Proceedings of the 2008 ACM symposium on Applied computing, 2008, pp. 1556-1560

[11] J. S. Lewin and A. Pribula, "Extracting emotion from twitter," Technical Report2009.

[12] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," CS224N Project Report, Stanford, vol. 1, p. 12, 2009.

[13] D. Davidov, O. Tsur, and A. Rappoport, "Enhanced sentiment learning using twitter hashtags and smileys," in Proceedings of the 23rd international conference on computational linguistics: posters, 2010, pp. 241-249.

[14] A. Pak and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," in LREc, 2010, pp. 1320-1326.

[15] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment analysis of twitter data," in Proceedings of the workshop on languages in social media, 2011, pp. 30-38..

[16] J. Bollen, H. Mao, and A. Pepe, "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena," ICWSM, vol. 11, pp. 450-453, 2011 [16] J. Bollen, H. Mao, and A. Pepe, "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena," ICWSM, vol. 11, pp. 450-453, 2011.

[17] S. N. Shivhare and S. Khethawat, "Emotion detection from text," arXiv preprint arXiv:1205.4944, 2012.

[18] W. Wang, L. Chen, K. Thirunarayan, and A. P. Sheth, "Harnessing twitter" big data" for automatic emotion identification," in Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom), 2012, pp. 587-592.

[19] L. Zhang, Y. Jia, B. Zhou, and Y. Han, "Microblogging sentiment analysis using emotional vector," in Cloud and Green Computing (CGC), 2012 Second International Conference on, 2012, pp. 430-433.

[20] S. M. Mohammad, "# Emotional tweets," in Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation, 2012, pp. 246-255.

[21] K. Roberts, M. A. Roach, J. Johnson, J. Guthrie, and S. M. Harabagiu, "EmpaTweet: Annotating and Detecting Emotions on Twitter," in LREC, 2012, pp. 3806-3813.

[22] A. Neviarouskaya and M. Aono, "Extracting Causes of Emotions from Text," in IJCNLP, 2013, pp. 932-936.

[23] W. Li and H. Xu, "Text-based emotion classification using emotion cause extraction," Expert Systems with Applications, vol. 41, pp. 1742-1749, 2014.

[24] M. Hasan, E. Rundensteiner, and E. Agu, "Emotex: Detecting emotions in twitter messages," 2014.

Topic Classification from Text Using Decision Tree, K-NN and Multinomial Naïve Bayes

Md. Ataur Rahman
Dept. of Language Science and Technology
University of Saarland
Saarbrücken, Germany
arahman@coli.uni-saarland.de

Yeasmin Ara Akter
Dept. of Computer Science and Engineering
East Delta University
Chittagong, Bangladesh
yeasmin.a@eastdelta.edu.bd

***Abstract*—One of the central motivations behind Natural Language Processing is detecting patterns. Given a text document, the task of identifying the context is known to be as topic classification. This paper explores the performance of three different classifiers namely Decision Tree, K-Nearest Neighbors and Multinomial Naive Bayes on a topic classification task (with six topic class). The evaluation is done on the basis of accuracy, precision, recall, and f1-score based results. Among those three aforementioned classifiers, we have selected the Multinomial Naïve Bayes as our best model using which we achieved 91.8% accuracy.**

***Keywords*—Text Classification, Natural Language Processing, Machine Learning**

I. INTRODUCTION

Topic classification (TC) is one of the classic subdomains of natural language processing which deals with allocating predefined classes to text documents. Due to the abundance of textual information in digitalized format, the need for categorization of documents has become a thriving interest for the past few decades. The first idea probably originated back in 1961 from the seminal work of Maron on probabilistic text classification [1]. One dominant approach to tackle this problem is based on machine learning techniques, which recently took the research communities by storm.

The prime inquisitiveness in most of the text classification task spans from finding out the best possible combinations of features with best Machine Learning (ML) classifiers. Our motivation in this article also extends towards the investigation of how different supervised machine learning classifier performs on a multi-label text classification task. We tried to experiment with the parameters of three different classifiers (Decision Tree, K-NN and Multinomial Naive Bayes) implemented using the scikit-learn library. We managed to elevate the performance of each of these baseline classifiers by employing some form of pre-processing (stemming, stop-word removing, lemmatization etc.). Apart from that, both *tf-idf* and *count* features were explored in order to compare the results for the aforementioned classifiers under different settings.

II. RELATED WORKS

Perhaps one of the earlier prominent work in this domain is the usage of Support Vector Machine (SVM) for learning to classify text [2]. The paper delineates a comparison between SVM and four traditional machine learning approaches. Two corpora were used for evaluation and an SVM with an RBF-kernel gained 86.4% and 66.0% accuracy respectively.

This paper [3] describes the correction scheme of the Reuters Corpus Volume I (RCV1-v2) from the original RCV1-v1. Apart from this, they have used the above amended corpus to evaluate three popular ML algorithms (SVM, k-NN, Rocchio). A micro F1 score of 0.816 and macro F1 of 0.619 was obtained using an SVM.

A paper [4] by Sebastiani et al. provided an in-depth exploration of nearly every type of classification techniques including the neural network and could be recalled as a starting point in our research. The paper also compared state-of-art systems at that time among which the best-reported method [5] using SVM obtained an accuracy of 92% and Decision Tree (DT) gave 88.4% accurate results using the Reuters corpus.

Another article [6] provides a heuristic on character-level Convolutional Neural Networks (CNN) for the task of textual document classification. They have compared several traditional as well as deep learning models and claimed that character-level CNN is a better choice in terms of data scalability.

III. DATASET

For our experiment, we have considered a fraction of Amazon's product review corpus consisting of 6000 text documents from which 75% (4500) documents are used for *training* purpose. The rest 25% (1500) data are used for *testing* the performance of the system. Each review is annotated as one of *six* topic type categories (*books, camera, dvd, health, music, software*) which corresponds to multi-class variants. The distribution of classes in both training and testing data is shown in TABLE I.

TABLE I. DISTRIBUTION OF VARIOUS LABELS IN THE DATASET.

Labels	Training Set	Testing Set
Books	760	233
Camera	730	258
DVD	770	242
Health	743	243
Music	767	260
Software	730	264
Total	4500	1500

IV. Methodology Of The System

In this section, a brief idea about the settings and approach taken for all the three classifiers is given. But before going to the details about the classifiers we will have a look at the *pre-processors* used in our system:

A. Preprocessing

We have experimented with three types of preprocessing techniques that are popular in text analysis domain. First of all, we have removed the *stop words* (such as *a*, *of*, *the* etc.) which occurs regularly in English text and do not generally contribute that much in classification. The second pre-processing techniques that we followed is the *stemming* of multiple forms of a single word into the same root. We have used *Porter's Stemmer* from NLTK for stemming. Both of these improved the results substantially for example in the case of K-NN classifier. The accuracy improves from 75.7% to 82.6% (7% for tf-idf feature) and from 41.87% to 59.1% (17.23% for count feature). The other classifiers performance also improves between 1-2%.

Another tool that we experimented with our system is the WordNet *lemmatizer* from NLTK which didn't improve the performance, this may be due to the reason that, for topic classification task it is important to preserve the word as it is instead of converting it to lemma.

B. Decision Tree

Decision Tree (DT) constructs classification models by breaking down a dataset into smaller subsets while incrementally building an associated tree. The final tree contains *decision nodes* and *leaf nodes* (Fig.1). A decision node consists of two or more branches (e.g., color, size and shape) while a leaf node (e.g., edible) represents a decision or classification. The best predictor, also referred to as the **root node**, is marked by the topmost decision node. DT can operate on either numerical or categorical data. Perhaps the intuition for the decision tree can be best described by an example (Table II):

TABLE II. An Example Dataset.

Color	Size	Shape	Edible
Yellow	Small	Round	Yes
Yellow	Small	Round	No
Green	Small	Irregular	Yes
Green	Large	Irregular	No
Yellow	Large	Round	Yes
Yellow	Small	Round	Yes
Yellow	Small	Round	Yes
Yellow	Small	Round	Yes
Green	Small	Round	No
Yellow	Large	Round	No
Yellow	Large	Round	Yes
Yellow	Large	Round	No
Yellow	Large	Round	No
Yellow	Large	Round	No
Yellow	Small	Irregular	Yes
Yellow	Large	Irregular	Yes

A decision tree utilizes *entropy* to compute the homogeneity from a sample. *Information gain is* another terminology which indicates the mitigation of entropy when a dataset is partitioned based on an attribute. Constructing a node in the DT thus involves finding attribute that produces the maximum *information gain*. Thus, the first Branch of the tree in Fig.1 consist of the attribute *size* because it has the maximum Information Gain (0*.*106) among all the three features [$G(color) = 0.036$; $G(shape) = 0.036$].

In the scikit-learn implementation of the Decision Tree classifier, we have experimented with several parameters such as *maximum depth* of the tree, *minimum samples* to split an internal node, *minimum samples* per leaf node, *maximum number of features* and etc. We have tried different

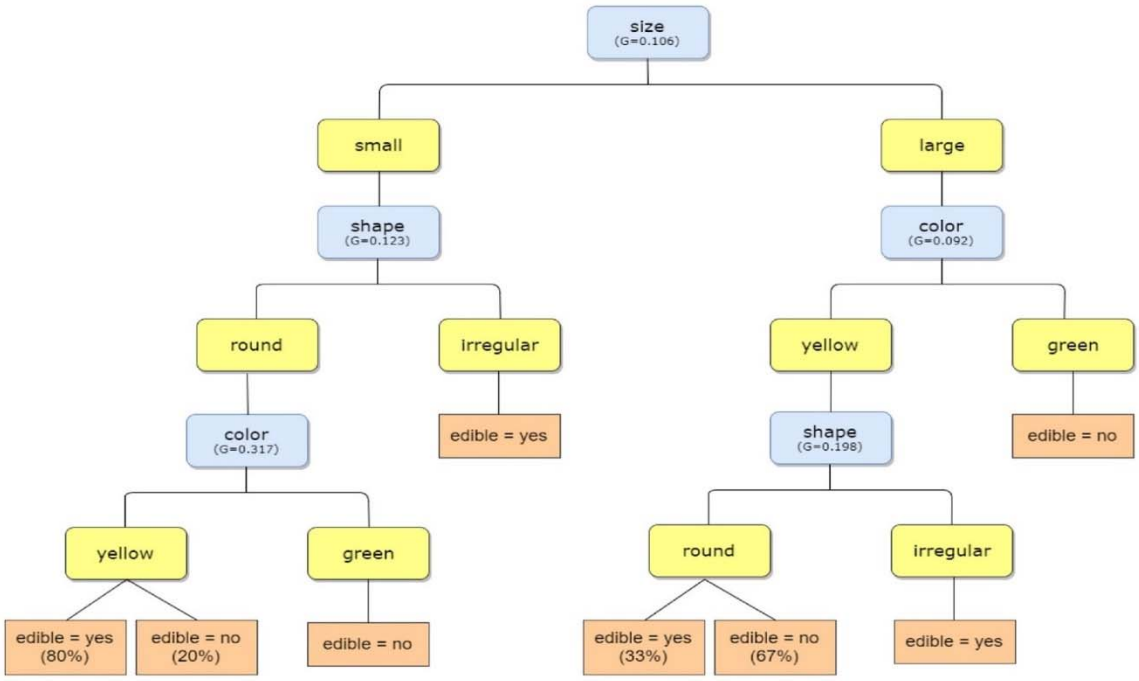


Fig.1. Decision Tree for the dataset of TABLE II.

combinations by changing these parameters. One such setting for example, is given below:

```
DecisionTreeClassifier
(min_samples_split=3,
min_samples_leaf=2, max_depth=10,
max_features=1000)
```

Changing the value of parameter reduces the performance of the classifier from an accuracy of 79.4% (tf-idf feature) and 79.93% (count feature) to 48.47% and 47.4%.

C. K-Nearest Neighbors

The second classifier that we compared for performing a six-way classification was the K-NN classifier. By experimenting with different values of *K*, we have observed the following phenomenon after plotting the *accuracy* and *f1-score* (Fig.2.). The observations are as follows:

a) Accuracy gets better with a higher K in the observed experiment.

b) The overall performance (both accuracy and f1-score) increases drastically from *K=1* to *K=10*. After that, the curves seem to have more like a plateaued progression having upward trends with some fluctuations.

c) The *bias* for K-NN classifier increases monotonously with the value of *K*, while the variance drops off as *K* is increased. In fact, under reasonable assumptions, the bias of the 1-nearest neighbor (*K=1*) classifier disappears completely with the size of the training data approaching infinity.

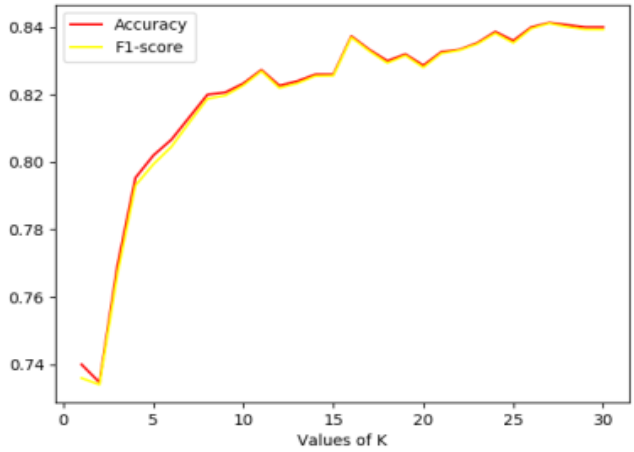


Fig.2. Performance of K-NN classifier for different values of K.

A comparison based on *time* (in seconds) taken for training and testing the other two classifiers with K-NN is given below in Table III.

TABLE III. COMPARISON OF TRAINING AND TESTING TIME.

	NB	**DT**	**KNN**
TRAINING(S)	0.906	3.914	0.789
TESTING(S)	0.215	0.208	1.015

From Table III, we can see that the different classifier has different training and testing time. The lowest training time is achieved by K-NN classifier (0.789) while the highest training time is for Decision Tree (3.914). The testing time, on the other hand, is highest for K-NN (1.015). These training and testing time heavily depend on the architecture of the classifier itself. As building tree takes linear time $O(n)$ with a searching cost of $O(\log n)$, the overall complexity gets high for DT. For Naïve Bayes it's fairly simple, it takes only the time to construct the frequency table and the multiplication time to calculate the conditional probability of the individual words. On the contrary, K-NN has properties that are quite different from most other classification algorithms. Training a K-NN classifier simply involves the task of determining *K* and preprocessing documents. In fact, if we pre-select a value for *K* and do not pre-process, then K-NN requires no training at all. For *n* samples in *D* dimensions, the test time for k-NN is $O(\text{D n log } n)$. It is linear in the size of the training set as we need to compute the distance of each training document from the test document. Test time is independent of the number of classes. K-NN therefore, has a potential advantage for problems with large class size.

D. Multinomial Naïve Bayes

The third classifier that we have used is the "Multinomial Naive Bayes" classifier which is a slightly modified version of the Naive Bayes classifier. The only difference is that while the Naive Bayes works with conditionally independent features the modified Multinomial version could be used to estimate the likelihood of conditionally dependent features by considering them as independent features. As it uses a bag of word feature in the text, it assumes that there is no relation between each word through the text, which is okay as long as we don't consider the semantic relationships of words in a sentence or among the sentence.

V. RESULTS AND EVALUATION

In this section, we will compare the results of three different classifiers mentioned in the previous section (Section III). For our evaluation purpose, we considered *accuracy*, *precision*, *recall* and *f1-score* based measures. We have also used a *10-fold cross validation* over the entire dataset for further evaluation regarding each type of classifiers performance.

For the purpose of simplification, we have only illustrated the comparison of the overall score (micro average) for the three classifiers (Table IV and V). Based on the comparison, we have chosen *Multinomial Naive Bayes* classifier with *Tf-Idf feature* as our *Best Model*. A detailed result for each of the class label for the best model is depicted in Table VI.

TABLE IV. COMPARISONS OF RESULTS FOR DIFFERENT CLASSIFIER (TF-IDF FEATURE).

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
DT	0.788	0.794	0.788	0.790
KNN	**0.826**	0.831	0.826	0.826
NB	**0.918**	0.920	0.918	0.918

TABLE V. COMPARISONS OF RESULTS FOR DIFFERENT CLASSIFIER (COUNT FEATURE).

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
DT	**0.796**	0.802	0.796	0.798
KNN	0.591	0.613	0.591	0.590
NB	**0.905**	0.908	0.905	0.905

TABLE VI. DETAILED RESULTS OF MULTINOMIAL NAÏVE BAYES CLASSIFIER.

	precision	**recall**	**F1-score**	**support**
Books	0.935	0.923	0.929	233
Camera	0.862	0.942	0.900	258
DVD	0.920	0.901	0.910	242
Health	0.971	0.840	0.901	243
Music	0.947	0.962	0.954	260
Software	0.892	0.936	0.913	264
Avg/Total	0.920	0.918	0.918	1500

To have a clearer look at the accuracy results, we have also evaluated our classifiers with the 10-fold cross-validation scheme. These are depicted in Table VII.

TABLE VII. 10-FOLD CROSS VALIDATION ACCURACY RESULTS FOR THREE DIFFERENT CLASSIFIERS(BASED ON BEST RESULTS).

Folds	**NB**	**DT**	**KNN**
Fold-1	0.926	0.797	0.856
Fold-2	0.863	0.778	0.810
Fold-3	0.882	0.750	0.796
Fold-4	0.893	0.800	0.793
Fold-5	0.886	0.731	0.825
Fold-6	0.812	0.731	0.731
Fold-7	0.792	0.772	0.778
Fold-8	0.926	0.832	0.852
Fold-9	0.872	0.757	0.750
Fold-10	0.858	0.723	0.777

VI. CONCLUSION

From the result section, we can see that the overall accuracy for Multinomial Naive Bayes classifier for the task of topic classification in our experiment was 91.8% which is better than Decision Tree (79.6%) and K-NN (82.6%) classifier at their best settings. This is the prime motivation behind selecting *Multinomial Naive Bayes* as our best model. Again, the time consumed by MNB (Table III) is a good trade-off in terms of both training and testing perspective. The approach that MNB takes is also straight forward to understand as it directly considers the prior and posterior probabilities for classification purpose and there is no further complexity under-the-hood. We have used three different pre-processing techniques in our experiment. Although, it would have been nice to incorporate other normalization and feature selection techniques to see if the result improves or not, nevertheless what we have achieved is pretty much similar to the state-of-the-art systems in this domain.

REFERENCES

[1] M. E. Maron, "Automatic indexing: an experimental inquiry," *Journal of the ACM (JACM),* vol. 8, pp. 404-417, 1961.

[2] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features," in *European conference on machine learning*, 1998.

[3] D. D. Lewis, Y. Yang, T. G. Rose and F. Li, "Rcv1: A new benchmark collection for text categorization research," *Journal of machine learning research,* vol. 5, pp. 361-397, 2004.

[4] F. Sebastiani, "Machine learning in automated text categorization," *ACM computing surveys (CSUR),* vol. 34, pp. 1-47, 2002.

[5] S. Dumais, J. Platt, D. Heckerman and M. Sahami, "Inductive learning algorithms and representations for text categorization," in *Proceedings of the seventh international conference on Information and knowledge management*, 1998.

[6] X. Zhang, J. Zhao and Y. LeCun, "Character-level convolutional networks for text classification," in *Advances in neural information processing systems*, 2015.

Survey on Text-Based Sentiment Analysis of Bengali Language

Nayan Banik*, Md. Hasan Hafizur Rahman*, Shima Chakraborty†, Hanif Seddiqui†, Muhammad Anwarul Azim†
*Department of Computer Science & Engineering
Comilla University, Comilla - 3506, Bangladesh
Email: cse.nayan@gmail.com, hhr@cou.ac.bd
†Department of Computer Science & Engineering
University of Chittagong, Chittagong - 4331, Bangladesh
Email: shimactg@gmail.com, hanif@cu.ac.bd, azim@cu.ac.bd

***Abstract*—Digital contents in Bengali text are increasing with the advancements of World Wide Web (WWW) to express the sentiment of content originators from mixed perspectives. These high voluminous unstructured web contents can be utilized to build smarter tools to assist people through Natural Language Processing (NLP). Though Bengali NLP tools are still inadequate due to its inherent complexities, research on Sentiment Analysis in Bengali is flourishing as a challenging area and is getting researcher's attention at a rapid pace. To do research in this area, researchers spend a lot of their time to investigate previously published works to understand the advancements for further enhancement; one of the most stressful and challenging parts of research. In this regard, we are going to demonstrate domain specific available pre-processing steps along with existing research methodologies, respective datasets and evaluation metrics of notable works related to sentiment analysis of Bengali language as our research finding. Moreover, we also incorporate a series of directions for future researchers to augment their research in this exciting domain.**

Keywords: *Sentiment Analysis*, *Opinion Mining*, *Machine Learning*, *Information Extraction*

I. Introduction

Textual documents in Bengali are continuously increasing together with web contents of other languages on the World Wide Web (WWW) consisting of our regular activities like enhancing Banglapedia[1], producing social media posts, publishing documents of financial and economic markets, spreading political views, sharing reviews of online products along with published online news and so on. These available web information express originators sentiments and emotions toward various topics in daily life and they are important to make human decisions as well as for making automatic decisions through machines. Though, human being can extract sentiment in a limited extent by utilizing their intuitive learning, machines are incapable of doing so. In this regard, researchers performed their research in the last few decades all over the world to develop smarter methods by experimenting with classical machine learning and statistical tools to detect major properties of the text including subjectivity detection and emotion or polarity detection. The subjectivity detection determines the text is either subjective or objective, while the polarity detection asserts that the text is either positive, negative or neutral[1].

Furthermore, textual documents of a language are analyzed from document level, sentence level, phrase level or lexical level perspectives. The analysis in these perspectives are performed using supervised, unsupervised, semi-supervised or case based reasoning techniques[2]. In the supervised approach, researchers applied classical machine learning models, such as Naïve Bayes (NB), Maximum Entropy (ME), Artificial Neural Network (ANN), Support Vector Machines (SVM), whereas Latent Dirichlet Allocation (LDA) and theme based clustering are used as unsupervised techniques[3][4]. Moreover, graph based [5] and deep learning[6] approaches are also applied as modern trend for Sentiment Analysis (SA) research now-a-days.

Additionally, domain-specific data plays an important role in SA research; however, the lack of annotated and standard benchmarking datasets make the research much harder in this area[7]. Research in these multi-domain areas is considered as a part of Natural Language Processing (NLP) tasks which attract research community to perform their research[8]. As a result, researchers have developed a series of language-specific NLP tools to extract text properties of many languages like *Arabic, Hindi, Chinese, Czech, French, German, Italian, Japanese, Russian, Thai* to enrich their linguistic area for knowledge extraction[9]. In spite of having some research works in SA on the Bengali language, this area is still underdeveloped and therefore, research in this area would provide extensive opportunities to develop linguistic NLP tools for analyzing Bengali texts in the next few decades. In this connection, it is important for researchers to understand the chronological improvement in this area because imminent research constantly relies on the advancement of outcomes of previous works.

The rest of the paper is organized as follows. **Section II** introduces general terminologies together with related processing activities. Overview of a system related to SA is articulated in **section III**, while **section IV** focuses on the review of SA research along with our judgment in Bengali language text. Concluded remarks in addition to some directions for future research are described in **Section V**

[1]http://bn.banglapedia.org/

II. General Terminologies and Related Processing Activities

In this section, we introduce some definitions of general terminologies along with various processing activities of SA research those are found in our investigation. This will help researchers to understand the essence of our review on textual documents of Bengali language.

A. Sentiment Analysis and Emoticons

Sentiment Analysis (SA), also known as Opinion Mining (OM) refers to extract the subjective impressions or attitudes of content originators from published texts in terms of topics or contextual polarity through the use of NLP, statistics or classical machine learning techniques. On the other hand, emoticons are graphical facial expressions that represent the meaning of verbal communications without observing the change of body language. These pictorial presentations express a range of tone or feeling which plays an important role in social media as well as messaging in cyber-world[10][11]. The Fig.1 represents a snippet of Bengali text document that includes emoticon, hash tags, punctuation, and URLs.

'আয়নাবাজি' দেখে এলাম; যেমন ভেবেছিলাম, তার চেয়েও হাজার গুন ভালো লেগেছে :) ছবির এই দৃশ্য অসাধারণ ছিল http://images.myreviewsite.com/aynabaji.jpg। সময় করে দেখে নিন। #আয়নাবাজি #Aynabaji

Fig. 1. A snippet of document written in Bengali language

B. Stop Words, Punctuation, Hash Tags and URL Removal

Single letter words or less influential extremely high frequent words known as stop-words should be filtered out before processing the text. Since stop words and punctuation characters have no association with sentiment polarity, they are also removed by using their Unicode representations. Uniform Resource Locators (URL) are used to identify resources in the WWW but have no significance in SA. Moreover, hashtags in social networks are also marked with a number sign (#) in front of a word or unspaced phrase in a document[12][13]. The Fig. 2 shows the output of Fig. 1 after removing stop-words, punctuations, hash tags and URLs to understand the significance of this process in SA.

আয়নাবাজি দেখে এলাম ভেবেছিলাম হাজার গুন ভালো লেগেছে ছবির দৃশ্য অসাধারণ সময় দেখে নিন

Fig. 2. Stop-Words, Punctuation, Hash Tags and URL remove from Fig. 1

C. Tokenization and Normalization

Tokenization or text segmentation is a process of dividing the written text into meaningful units or tokens, such as words, sentences, or topics delimited by whitespace, tab, new line and so on. In bag-of-words approach, each input sentence is tokenized into words which act as the primary features of SA. Normalization is the process of converting each token into unified scale for further processing. The Fig. 3 demonstrates the process of tokenization of texts after removing stop-words, punctuations, hashtags, URLs of Fig. 1 to close look the process of recognized individual words in a system.

'আয়নাবাজি', 'দেখে', 'এলাম', 'ভেবেছিলাম', 'হাজার', 'গুন', 'ভালো', 'লেগেছে', 'ছবির', 'দৃশ্য', 'অসাধারণ', 'সময়', 'দেখে', 'নিন'

Fig. 3. Tokenization of words from Fig. 2.

D. Vectorization

Vectorization is a process to reconstruct a collection of texts in a document into numerical feature vectors. A document is considered as a collection of words with or without considering the information of their relative position. High frequent words having very little meaning must be excluded to prioritize rarer yet more interesting terms in the corpus. In order to reweight these count features into floating point values, Term-Frequency and Inverse Document-Frequency ($TF-IDF$) transformation or word embeddings are used. Fig. 4 illustrates the vectors of words of Fig. 1 after removing less important words and performing tokenization operation.

(আয়নাবাজি, 9), (দেখে, 34), (এলাম, 69), (ভেবেছিলাম , 54), (হাজার, 21), (গুন, 60), (ভালো, 45), (লেগেছে, 11), (ছবির, 76), (দৃশ্য, 51), (অসাধারণ, 75), (সময়, 23), (দেখে, 34), (নিন, 97)

Fig. 4. Vectorization of text from Fig. 3

E. Stemming and Lemmatization

Stemming is used to reduce inflectional forms of a language by splitting a word into its constituent root part and affix without doing the complete morphological analysis. As Bengali is a highly inflected language with pragmatically free word order i.e., Bengali (verb, noun, adjective) words are inflected forms of roots stemming is performed on the tokenized words[14] to reduce the dictionary size. In Fig. 5, we portrayed stemmed words of sentences that are given in Fig. 1 to comprehend the essence of a stemmer in SA. On

'আয়না', 'দেখ', 'এল', 'ভাব', 'হাজার', 'গুন', 'ভালো', 'লাগ', 'ছবি', 'দৃশ্য', 'অসাধারণ', 'সময়', 'দেখ', 'নে'

Fig. 5. Stemmed words from sentences of Fig. 1

the other hand, $lemmatization$ is the process of grouping different inflected words so that these words analyzed as a single item. Though the stemmer operates on a single word without contextual knowledge, lemmatization process understands the context of a word in a sentence. To familiar with the lemmatized words as output Fig. 6 is given to demonstrates the importance of lemmatization process.

'আয়নাবাজি', 'দেখে', 'এলাম', 'ভেবেছিলাম', 'হাজার', 'গুন', 'ভালো', 'লেগেছে', 'ছবির', 'দৃশ্য', 'অসাধারণ', 'সময়', 'দেখে', 'নিন'

⇓

'আয়না', 'দেখ', 'এল', 'ভাব', 'হাজার', 'গুন', 'ভালো', 'লাগ', 'ছবি', 'দৃশ্য', 'অসাধারণ', 'সময়', 'দেখ', 'নে'

Fig. 6. Lemmatization after tokenization of contents of Fig. 1

F. Parts-of-Speech (POS) Tagger

Automated POS tagger is used as a NLP tool to detect each word with a corresponding POS in a sentence. It helps

to verify and validate textual data, annotate named entities in the large corpus, translate one language to another language, correct the grammatical errors, and so on. As a result, Bengali POS tagger is an essential tool to enrich Bengali language, especially in the area of NLP. To comprehend the influential effect of a POS tagger, the Fig. 7 recognize respective POS of texts those are given in the Fig. 1.

'আয়না_NN', 'দেখ_VRB', 'এল_VRB', 'ভাব_VRB', 'হাজার_NN', 'গুন_NN', 'ভালো_ADJ', 'লাগ_VRB', 'ছবি_NN', 'দৃশ্য_NN', 'অসাধারণ_ADJ', 'সময়_NN', 'দেখ_VRB', 'নে_VRB'

Fig. 7. Parts-of-Speech Tag of words of contents in Fig. 1

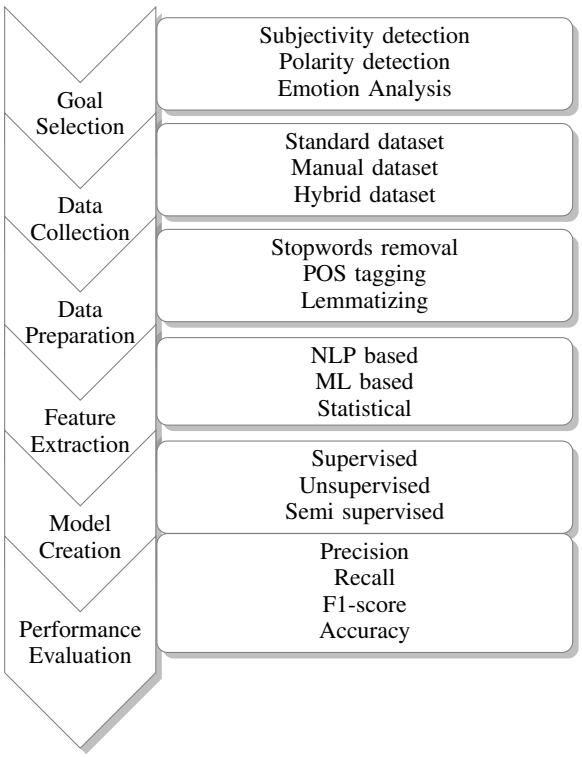


Fig. 8. Steps in a Sentiment Analysis System

G. WordNet and SentiWordNet

$WordNet$ is a lexical database of English language having nouns, adjectives, verbs and adverbs grouped into sets of cognitive synonyms called $synsets$. This database is used for word-sense disambiguation, information retrieval, automatic text classification, text summarization, machine translation and so on. On the other hand, $SentiWordNet$ is a lexical resource for SA which assigns three numerical sentiment scores (positivity, negativity or objectivity) to each synset of WordNet for the quantitative analysis of associated synsets and the use of vectorial term representations for semi-supervised synset classification[15].

III. Overview of a Sentiment Analysis System

*S*entiment *A*nalysis (SA) is a complex system that follows step-by-step logical phases to execute specific operations to achieve particular goals of the research. These phases of SA system are demonstrated in Fig. 8 to figure out the significance of SA research.

a) Goal Selection: Considering the application areas and associated text properties, the system goal is defined in the first phase which determines whether the given text is subjective or not. If the text is identified as subjective then the polarity is extracted and emotional level is investigated by quantifying sentiment scores.

b) Data Collection: Domain-specific datasets play an important role to achieve the desired goal where they can be either standardized or manually developed for benchmarking. Though standard benchmark datasets are available for some rich languages, most of the low resource languages use manual dataset either created by the expert system or crawling diverse online sources by custom crawlers.

c) Data Preparation: Raw data consisting extraneous noise like emoticons, stop words, punctuations, hashtags, $URLs$ must be removed. In this phase, stemming or lemmatization is used for highly inflected languages to convert words to its base form while POS tagger is used to detect each word with a corresponding POS in a sentence.

d) Feature Extraction: Influential features having a positive impact on the performance of the system can be examined from various perspectives such as NLP, Machine Learning (ML) or Statistical. For NLP based feature extraction, the system implements linguistic knowledge where ML approaches use training on corpus data and the statistical techniques consider the weighted importance of features in the system.

e) Model Creation: Selected features from the feature extraction phase are fed to a model that can be either lexicon based, supervised, unsupervised or semi-supervised. As a supervised model, NB, ME, SVM, ANN can be considered where clustering approaches and bootstrapping are considered as unsupervised and semi-supervised models respectively.

f) Performance Evaluation: The performance of SA system is evaluated by several metrics including expert analysis, precision, recall, F-score, accuracy to justify the acceptance of the system under considerations. Therefore, remarks on limitations and future directions are given by discussing the performance of a system.

IV. Sentiment Analysis in Bengali Language

Numerous amount of textual electronic documents in the Bengali language from diverse domains such as online news portals, social media text, online product reviews and so on are published on the WWW and it is continuously increasing where subjective information plays a great role from both the user and organizational perspectives. Due to this twofold importance and the exponential growth of Bengali text on the web, researchers are attracted to perform their research in this area. Although Bengali is an highly inflected language suffering from major challenges such as lack of labeled dataset or absence of rich NLP tools[16], researchers performed their individual research in the document level, sentence level, phrase level or lexicon level.

Document level SA represented in [17] use Mutual Information (MI) as a feature extractor for NB classifier. The researchers used Amazon Watch Reviews and their word by word translation in Bengali as datasets. After performing standard preprocessing, the experiment results show that the system works well for Bengali considering negation. But the drawback is word by word translation of the English dataset. Another similar works in [18] utilizes translated Bengali dataset from Amazon watches for polarity detection. But these texts are cross-checked using dictionary word mapping to avoid meaningless sentences. The system recognize positive and negative polarity where each polarity is subdivided into weak, steady and strong. Though the work presents 85% accuracy for Bengali, it requires double translation for negation handling. In [19], the authors use NB classifier to extract sentiment from Facebook user comments (3000 statuses from 70 users) where negations are handled by antonym substitution. Standard preprocessing steps are performed before stemming operation extract bigram as features. The system demonstrates 77% precision but the small size to dataset is a limiting factor in this work. Relating to that, researchers in [20] claims NB model is incapable to detect complex multi-line sentences of a document for SA. The system results 83.20% accuracy for NB with bigram, stemmer, and normalizer. It also shows that WE performs better when the the dataset size is increased. Moreover in [21], the authors claim that NB classifier is ineffective for social media texts as these texts are not well written by maintaining a standard format. Using Facebook user comments on food reviews as dataset, the work represents lexicon-based approach for SA. Though they claim that dictionary-based approach exhibits better accuracy; however, they do not perform any sort of experiments to evaluate performance metrics regarding their statement.

Sentence level polarity detection on twitter posts is presented in [22] where the authors us SVM and ME text classifiers. This work demonstrates semi-supervised bootstrapping method where a small labeled dataset is used to label other unlabeled texts and generate sentiment lexicon. Investigators consider several features in this scholary work and show that SVM (accuracy 93%) outperformed ME (accuracy 85%). Relating to that, polarity detector for Bengali online news texts is presented in [23] where the authors utilize several features in SVM classifier. Separating subjective texts from the acquired dataset, authors use sentiment lexicons from SentiWordNet(Bengali)[24] in order to score the sentences. The system considers many features like POS tags, chunk level information, functional words, stemming cluster, negative words, and dependency tree features. The experimental result by considering all of these features demonstrate precision of 70.04% where the baselines are 63.02%. Moreover in [25], the authors investigate linear and non-linear SVM using N-gram based modeling on diverse social media documents. The system preprocess manually collected data and shows that N-gram vectorization approach results in highest accuracy 91.581% and 89.271% for linear and non-linear SVM respectively. Another traditional SVM based SA system having accuracy of 64% on Bengali cricket review is done in [26].

Furthermore, word level experiment is performed in [38] by demonstrating data standardization techniques where the automatic translation process is accomplished to organize positive and negative words of SentiWordNet. But this elementary research without dataset is less efficient for complex systems and it does not consider Bengali inflected word terms, spelling errors, colloquial terms for sophisticated analysis. Therefore, different researchers perform experiment in [39] to exhibit shared SA task of Bengali tweets along with Hindi and Tamil. Since the datset is very small, this work in not compatible with modern deep learning techniques. On the other hand, one more research in [27] presents a theme cluster-based approach in a corpus developed from news and blog domains to recognize the subjectivity of texts. As a subjectivity classifier researchers deal with various features in their experiment and they examine the results by using Samsad English-Bangla Dictionary, Multi-Perspective Question Answering ($MPQA$) corpus and International Movie Database ($IMDB$). In addition, document-level SA in [28] propose Contextual Valence Analysis (CVA) for polarity detection in Bengali language texts using WordNet and SentiWordNet. To evaluate the proposed system, authors use expert analysis to justify the approach to a satisfactory level. The deficiency of this work is the earlier translation process from Bengali to English to perform their analysis. One more research on horoscope data is observed in [30] where the data are annotated by two native speakers and when disagreement appears between them, a third annotator finally annotate these sentences. After the dataset preparation, five classifiers are trained (NB, SVM, $K-Nearest\ Neighbours$ (KNN), Decision Tree (DT) and Random Forest (RF)) with their different parameters. Among these five classifiers, SVM with unigram features produce an accuracy rate of 98.7% as the best sentiment classification model. Relating to that in [34], the authors demonstrate SVM outperforms NB when applied to a manually developed Bengali movie review dataset. The system also provides a comparative analysis on different features having precision of 0.86 while considering stemmed unigram in SVM classifier. In the same fashion, sentence level research introduces deep learning methods for both plain Bengali and Romanized Bengali texts to extract the sentiment in the context of either positive or negative polarity [29]. In this research, they use binary cross entropy and categorical cross-entropy as loss functions with Recurrent Neural Network (RNN) specially Long Short Term Memory ($LSTM$) model. The major challenge of this research is lack of standard and large dataset since deep learning models require a huge dataset compared to other machine learning models. The successive research in [32] shows the Word Embedding (WE) with Hellinger PCA can be useful for detecting sentiment from Bengali text. In this stage, WE extracts the syntactic context of the words rather than the sentiment words. The authors implement Sentiment Specific Word Embeddings ($SSWE$) to recognize their respective sentiment in this research. The results depict that the performance can be increased with large corpus. In

TABLE I
SUMMARY OF INVESTIGATION OF SENTIMENT ANALYSIS ON BENGALI LANGUAGE

Paper	Analysis	Context	Dataset	Preprocessing	Methods	Result
[27]	Subjectivity Detection	Reviews, news, blog sites	Manual, Standard	POS tagging, clustering	CRF	Precision - 72.16% and 74.6%
[28]	Document Level		Manual	Word translation, sense values from SentiWordNet	CVA	
[17]	Document Level	Product reviews	Manual	Review rating selection, stopwords removal, negation handling	NB	Accuracy - 85%
[29]	Sentence Level	Social media, news, product reviews	Manual	Emoticons, proper noun removal	RNN (LSTM)	Accuracy - 70%
[30]	Sentence Level	Horoscope	Manual	Punctuations, stopwords removal, IG index	NB, SVM, KNN, DT, RF	Accuracy - 98.7% (SVM)
[18]	Sentence Level	Product reviews	Standard	Translation, negation handling	NB	Accuracy - 85%
[19]	Sentence Level	Social media	Manual	Hashtag, url removal, stemming	NB	Precision - 77%, Recall - 68%, F-score - 72%
[22]	Sentence Level	Social media	Manual	Tokenization, normalization, POS tagging	SVM,ME	Accuracy - 93%
[31]	Sentence Level	Social media	Manual	Removal of punctuation marks, extra spaces, unrecognized characters	word2vec, WE	Accuracy - 75.5%
[32]	Sentence Level	Microblogging sites	Manual	Removal of punctuation marks, extra spaces, unrecognized characters	WE	Accuracy - 70.0%
[23]	Phrase Level	News	Manual	POS tagging, clustering	SVM	Precision - 70.04%, Recall - 63.02%
[33]	Sentence Level	Social media	Standard	Special characters removal, irrelevant features removal, word level sentiment polarity tagging	CNN	Accuracy - 46.80%
[25]	Sentence Level	Social media	Manual	POS tagging, emoticons and punctuation removal, stemming, negativity separation	SVM	Accuracy - 91.684% and 89.271%
[34]	Sentence Level	Social media	Manual	URL, emoticon, punctuation, stop words removal	NB, SVM	Precision - 0.86%
[26]	Sentence Level	Reviews, Social Media	Manual, Standard	Tokenization, stop words removal, punctuation and number removal	SVM	Accuracy - 64%
[35]	Sentence Level	Social Media	Manual, Standard	Emojis, symbols, numbers, stickers, English letters removal	RNN	Accuracy - 85.67%
[36]	Sentence Level	Reviews, Social Media	Manual	Non Bengali character and symbol removal	Word2Vec	Accuracy - 83.79%
[37]	Sentence Level	Social Media	Manual	Tokenization, URL, punctuation, user tags, mentions and stop words removal	LSTM, CNN, NB, SVM	Accuracy - 65.97% and 54.24%

addition, the research in [31] depicts words2vec model for Bengali SA in the sentence level; where researchers claim that WE of sentences determine the characteristics of words and their respective context. In this research, investigators remove all noise from the data source through some general preprocessing steps. Then positive and negative scores are calculated using both syntactic score and mutual score among these data. Although the performance is low, the authors use WE to capture the sentiment properties of words in a sentence. Researchers also neutralize the effects of valence shifting words by using threshold value for highly positive and highly negative in their corpus. With the recent improvements in deep learning techniques, authors in [33] develop a SA system based on deep Convolutional Neural Network (CNN) for twitter dataset[39]. The proposed method includes preprocessing for removing special characters, irrelevant features and lexical level sentiment polarity from Sentiwordnet[24]. In order to compare the performance, the authors have implemented a traditional Deep Belief Network (DBN). More related deep learning based systems in [35] and [36] shows that the scarcity of benchmark dataset in Bengali poses a challenge for deep learning techniques in SA task. Table I provides a summary of discussed notable research works in Bengali SA task.

V. CONCLUSION

This review article demonstrates existing procedures, achievements, and limitations of available works of Sentiment Analysis (SA) on Bengali texts. In our findings, we represent

preprocessing steps regarding the domains, context of research, comparisons of different experimental methods as well as performance evaluation of a system which will help the new researchers to do their research in this field. In our judgment, a small number of evaluation metrics are satisfactory regarding problem domains on Bengali texts; however, there are a large number of domains available that are not explored till now and hence it can be an exciting field to contribute. Moreover, Bangla Natural Language Processing ($BNLP$) tools, standardized datasets for benchmarking are not up to date for Bengali language to develop SA system effectively which must be kept in mind during the future experiments.

REFERENCES

[1] B. Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, vol. 5, no. 1, pp. 1–167, 2012.

[2] K. M. Rahman and A. Khamparia, "Techniques, applications and challenges of opinion mining."

[3] S. Sun, C. Luo, and J. Chen, "A review of natural language processing techniques for opinion mining systems," *Information Fusion*, vol. 36, pp. 10–25, 2017.

[4] K. D. Varathan, A. Giachanou, and F. Crestani, "Comparative opinion mining: a review," *Journal of the Association for Information Science and Technology*, vol. 68, no. 4, pp. 811–829, 2017.

[5] R. Feldman, "Techniques and applications for sentiment analysis," *Communications of the ACM*, vol. 56, no. 4, pp. 82–89, 2013.

[6] X. Glorot, A. Bordes, and Y. Bengio, "Domain adaptation for large-scale sentiment classification: A deep learning approach," in *Proceedings of the 28th international conference on machine learning (ICML-11)*, 2011, pp. 513–520.

[7] T. Al-Moslmi, N. Omar, S. Abdullah, and M. Albared, "Approaches to cross-domain sentiment analysis: A systematic literature review," *IEEE Access*, 2017.

[8] R. Piryani, D. Madhavi, and V. Singh, "Analytical mapping of opinion mining and sentiment analysis research during 2000–2015," *Information Processing & Management*, vol. 53, no. 1, pp. 122–150, 2017.

[9] D. Vilares, M. A. Alonso, and C. Gómez-Rodr íguez, "Supervised sentiment analysis in multilingual environments," *Information Processing & Management*, vol. 53, no. 3, pp. 595–607, 2017.

[10] B. Pang, L. Lee *et al.*, "Opinion mining and sentiment analysis," *Foundations and Trends® in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.

[11] S.-M. Tζιωρτζής, "Sentiment analysis by emoticons and unsupervised comment summarization in greek e-government data," 2013.

[12] I. Hemalatha, G. S. Varma, and A. Govardhan, "Sentiment analysis tool using machine learning algorithms," *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, vol. 2, no. 2, pp. 105–109, 2013.

[13] S. Rosenthal, A. Ritter, P. Nakov, and V. Stoyanov, "Semeval-2014 task 9: Sentiment analysis in twitter." in *SemEval@ COLING*, 2014, pp. 73–80.

[14] M. H. Seddiqui, M. N. Hoque, and M. H. H. Rahman, "Semantic annotation of bangla news stream to record history," in *Computer and Information Technology (ICCIT), 2015 18th International Conference on*. IEEE, 2015, pp. 566–572.

[15] S. Baccianella, A. Esuli, and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining." in *LREC*, vol. 10, 2010, pp. 2200–2204.

[16] M. Karim, *Technical challenges and design issues in bangla language processing*. IGI Global, 2013.

[17] A. K. Paul and P. C. Shill, "Sentiment mining from bangla data using mutual information," in *Electrical, Computer & Telecommunication Engineering (ICECTE), International Conference on*. IEEE, 2016, pp. 1–4.

[18] K. A. Hasan, M. S. Sabuj, and Z. Afrin, "Opinion mining using naïve bayes," in *Electrical and Computer Engineering (WIECON-ECE), 2015 IEEE International WIE Conference on*. IEEE, 2015, pp. 511–514.

[19] M. S. Islam, M. A. Islam, M. A. Hossain, and J. J. Dey, "Supervised approach of sentimentality extraction from bengali facebook status," in *Computer and Information Technology (ICCIT), 2016 19th International Conference on*. IEEE, 2016, pp. 383–387.

[20] M. Al-Amin, M. S. Islam, and S. D. Uzzal, "A comprehensive study on sentiment of bengali text," in *Electrical, Computer and Communication Engineering (ECCE), International Conference on*. IEEE, 2017, pp. 267–272.

[21] S. Akter and M. T. Aziz, "Sentiment analysis on facebook group using lexicon based approach," in *Electrical Engineering and Information Communication Technology (ICEEICT), 2016 3rd International Conference on*. IEEE, 2016, pp. 1–4.

[22] S. Chowdhury and W. Chowdhury, "Performing sentiment analysis in bangla microblog posts," in *Informatics, Electronics & Vision (ICIEV), 2014 International Conference on*. IEEE, 2014, pp. 1–6.

[23] A. Das and S. Bandyopadhyay, "Opinion-polarity identification in bengali," in *International Conference on Computer Processing of Oriental Languages*, 2010, pp. 169–182.

[24] ——, "Sentiwordnet for bangla," *Knowledge Sharing Event-4: Task*, vol. 2, 2010.

[25] S. A. Taher, K. A. Akhter, and K. A. Hasan, "N-gram based sentiment mining for bangla text using support vector machine," in *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)*. IEEE, 2018, pp. 1–5.

[26] S. A. Mahtab, N. Islam, and M. M. Rahaman, "Sentiment analysis on bangladesh cricket with support vector machine," in *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)*. IEEE, 2018, pp. 1–4.

[27] A. Das and S. Bandyopadhyay, "Subjectivity detection in english and bengali: A crf-based approach," 2009.

[28] K. A. Hasan, M. Rahman *et al.*, "Sentiment detection from bangla text using contextual valency analysis," in *Computer and Information Technology (ICCIT), 2014 17th International Conference on*. IEEE, 2014, pp. 292–295.

[29] A. Hassan, M. R. Amin, A. K. Al Azad, and N. Mohammed, "Sentiment analysis on bangla and romanized bangla text using deep recurrent models," in *Computational Intelligence (IWCI), International Workshop on*. IEEE, 2016, pp. 51–56.

[30] T. Ghosal, S. K. Das, and S. Bhattacharjee, "Sentiment analysis on (bengali horoscope) corpus," in *India Conference (INDICON), 2015 Annual IEEE*. IEEE, 2015, pp. 1–6.

[31] M. Al-Amin, M. S. Islam, and S. D. Uzzal, "Sentiment analysis of bengali comments with word2vec and sentiment information of words," in *Electrical, Computer and Communication Engineering (ECCE), International Conference on*. IEEE, 2017, pp. 186–190.

[32] M. S. Islam, M. Al-Amin, and S. D. Uzzal, "Word embedding with hellinger pca to detect the sentiment of bengali text," in *Computer and Information Technology (ICCIT), 2016 19th International Conference on*. IEEE, 2016, pp. 363–366.

[33] K. Sarkar, "Sentiment polarity detection in bengali tweets using deep convolutional neural networks," *Journal of Intelligent Systems*.

[34] N. Banik and M. H. H. Rahman, "Evaluation of naïve bayes and support vector machines on bangla textual movie reviews," in *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)*. IEEE, 2018, pp. 1–6.

[35] A. A. Sharfuddin, M. N. Tihami, and M. S. Islam, "A deep recurrent neural network with bilstm model for sentiment classification," in *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)*. IEEE, 2018, pp. 1–4.

[36] S. H. Sumit, M. Z. Hossan, T. Al Muntasir, and T. Sourov, "Exploring word embedding for bangla sentiment analysis," in *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)*. IEEE, 2018, pp. 1–5.

[37] N. I. Trinto and M. E. Ali, "Detecting multilabel sentiment and emotions from bangla youtube comments," in *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)*. IEEE, 2018, pp. 1–6.

[38] D. Das and S. Bandyopadhyay, "Developing bengali wordnet affect for analyzing emotion," in *International Conference on the Computer Processing of Oriental Languages*, 2010, pp. 35–40.

[39] B. G. Patra, D. Das, A. Das, and R. Prasath, "Shared task on sentiment analysis in indian languages (sail) tweets-an overview," in *International Conference on Mining Intelligence and Knowledge Exploration*. Springer, 2015, pp. 650–655.

Doly: Bengali Chatbot for Bengali Education

1st Md. Kowsher
Dept. of Applied Mathematics
Noakhali Science and Technology University, Noakhali-3814,Bangladesh.
ga.kowsher@gmail.com

2nd Farhana Sharmin Tithi
Dept. of CSE
Daffodil International University, Dhaka, Bangladesh
farhanasharmin3tithi@gmail.com

3rd M Ashraful Alam
Dept. of Applied Mathematics
Noakhali Science and Technology University, Noakhali-3814,Bangladesh
mdashrafulsikdar1580@gmail.com

4th Mohammad Nurul Huda
MSCSE
, United International University
United City, Madani Avenue, Badda, Dhaka 121, Bangladesh
mnh@cse.uiu.ac.bd

5th Mir Md. Moheuddin
Dept. of Applied Mathematics
Noakhali Science and Technology University, Noakhali-3814,Bangladesh
mirmdmoheuddinkhan@gmail.com

4th Md. Golam Rosul
Dept. of Applied Mathematics
Noakhali Science and Technology University, Noakhali-3814,Bangladesh
golamrosul1506040@gmail.com

***Abstract*— This Scientific Research paper is a synopsis of an automated system "Doly: Bengali Chatbot" which gives a reply to a user query on behalf of a human for the education system in the Bengali language. This is an AI-based Chatbot system, mainly based on machine learning algorithms and Bengali Natural Language Processing (BNLP). The machine gets embedded with this knowledge to identify the desired sentences and making a decision within itself, as a response to answer questions. There are many English Chatbot's which used in education, web query, banking sector & various sectors. In this case, we've propounded a complete data-driven retrieval based closed domain Chatbot which is easily colloquy in the Bengali language with the users. We've created the train function adapter to train the Doly by encoding (encoding="utf8") our corpus from bot data. An input adapter has been created to take input and for output, an output adapter has been created to generate automated responses to a user's input. We've also used a machine learning algorithm like search algorithm for finding an appropriate list of matching results from the corpus and use Naïve Bayesian algorithm to select right answer form list of matching result. The main aim of this Chatbot based system is to bridge the gap between the knowledge sources by providing instant replies to the questions and queries that have to ask in the Bengali language.**

***Index terms*--- Chatbot, Dialogue & Corpus, Naïve Bayesian algorithm, AI, Bengali Natural Language Processing (BNLP).**

I. INTRODUCTION

By Technology-based education, learners can easily get access to unlimited online resources. It creates huge facilities for technical learning, distance learning, and self-learning. Nowadays chatbot is using as an only one learning platform for the education system. Personal education chatbot will help students to interest by self-learning & also it can also remove the dependency from tutor & other technology. The chatbot is a conversational intelligent agent that can easily communicate with human users by using Natural Language Processing. It is likewise known as a chatterbot, Bot, IM bot, talkbot, Interactive agent, or Artificial Conversational Entity. The user can interact with the bot in written, oral, or mixed format of conversation. The applications of chatbots can be applied in various fields such as Health, Tourism Business,Customer Confirmation and so on. The architecture of Chabot integrates language models with computational algorithms. Normally, a chatbot works by a user asking a question or developing a comment and replying the question or initiating a new topic. We can deliver instance examples of Cleverbot [1] ALICE [1] etc.The Bengali language is the main language of Bangladesh and the fourth most popular language in the world. Bangla is spoken by about 245 million people of Bangladesh and two states of India [3]. The literacy rate of Bangladesh is 75.4% and the numbers of total students are 23,907,151. The rate of literacy is increasing day by day in Bangladesh. Nowadays the educations system is being dependent on Intelligent Instructing Systems of computer circumstance for education.This is why the paper proposes the Bengali language chatbot named 'Doly' based on the Bengali language with the help of Natural Language Processing (NLP) as well as Machine Learning (ML) techniques for developing an education tool of Bangladesh. The trial spectacle is that our chatbot is capable to impart responses to the users in the real time. To constructing Bengali chatbot 'Doly', we installed Python 3.7.0 & Anaconda in our Machine. For Natural Language Processing (NLP) we installed Natural Language Toolkit (NLTK). We also need to install Chatterbot that comes with built-in adapter classes that allow it to connect to various types of databases. We can't run directly the Bengali chatbot in chatterbot as it did not support Unicode.As a Unicode Decode Error for the implementation of Bengali chatbot 'Doly', we have to encode Bangla NLP with a chatterbot. We created the train function adapter to train the Doly by encoding our corpus from bot data. It this case we use encoding="utf8" in our .yml corpus to read Bengali data. To take input we created an input adapter and for output, we created an output adapter to generate automated responses to a user's input. This makes it easy to create chatbots and automate conversations with users. The task separated by four parts, training data, input data, process input, and return response or output. We also use machine learning algorithm like search algorithm for finding the appropriate list of matching results from the corpus and also use a Bayesian algorithm to select right answer form list of matching result.This paper described a created chatbot that returns a relevant answer to questions asked by the user in Bangla language.

The contributions are summarized as follows:

- We have proposed a Bengali chatbot as 'Doly' by imparting Bengali Natural Language Processing.
- For the pre-processing Bengali language, we have created our own BLTK tool in python.
- As algorithms, we have used Naive Bayes classifier and popular algorithm Edit distance of dynamic programming
- Various adapters were developed for our proposal task.

The paper is composed as follows, Section 2 gives the details of the related works on the chatbot technology. Section 3 contains the Bengali Natural Language Processing (BNLP). Section 4 described the design and methods used for Creating Doly: Bengali Chatbot. In section 5 offers the results and analysis. Section 6 included Comparison between others Chatbot. Section 7 concludes research work with future directions. Finally, Section 8 highlights the conclusion.

II. Related Work

Weizenbaum, J. et al.[1] discussed the first idea of chatbots originated in the Massachusetts Institute of Technology [2] where the Eliza chatbot (also known as "Doctor"). The users of this chatbot imagined Computer programs really could hear their problems and understand, and help them in a constructive way. There are other chatbots called ALICE that used heuristic pattern matching to follow the rules from the input [3].Breazeal, C. et al.[4] research has a robotic chat-bot which understand and relate to users in human or social terms, adapt and learn throughout their lifetime. It also communicates with users. Vinyals, O. et al.[5] discuss the converse model in natural language interaction that predicting the very next command given the previous command in conversations. Bickmore et al.[6] researched agent based chatbot. This chatbot is relational computer agents that they mapping social-emotional relationships for users. An artificial intelligence chatbot like Jabberwacky [7], that studies language and context-based human interactions. To find appropriate response to some stored data such as conversations and comments.Bickmore et al.[8] also research a relational agent which applied different relational behaviors to establish social bonds with museum visitors and evaluate the ability of an experimental relational agent to comfort users in stressful situations [9]. Cassell, J. et al.[10] The most popular agent called Rea that shows non-verbal behaviors such as hand gestures, facial displays, an eye gaze, body posture etc. This behaviors improve the effectiveness of communication and can create a more immersive experience for users. Augello et al.[11] proposed a social chatbot model for a communicative skill learning game based on "Social practice" that can choose the most appropriate dialogue. Hill J. et al.[12] shows a different chatbot such as human-chatbot conversations lack in content, quality and vocabulary and human-human conversations over IM.Pan, X. et al.[13] discussed a virtual agents chatbot. They suggested a projection based system virtual party like an environment. In this system, a conversational male user approached by a female character called Christine. Christine also can formulate personal questions and statements. This female character is interested in a male user that involves smiles, headnotes, eye contact, leaning towards. Crutzen, R. et al.[14]studies an information chatbot that helps users to know about sex, drugs, and alcohol use among adolescents. This chatbot also called question related chatbot or search engines. Shawar et al.[15] described existing work on chatbot knowledge acquisition is mainly based on human annotated datasets. Their approaches are helpful to develop sense knowledge for Bangla chatbots, but there are not available of extracting knowledge for the specific task and technical challenges and design issues in Bangla language processing[16]. Shrestha et al. [17] proposed a way is to detect pairs in an email conversation for the work of email summarization. Zhou and Hovy [18] developed a summarization system for technical chats and emails about the Linux kernel.

Unlike these works, we have created a search based chatbot for Bengali language using BNLP. For Bengali Natural Language processing, we have created BLTK and BRE tools.

III. Bengali Natural Language Processing (BNLP)

The Bengali language is the fourth most spoken language in the world. It is the mother tongue of the people of Bangladesh; there are also few parts of India like as Calcutta who talk with Bangla. But it is really a matter of great mournful that the Bengali language deliberated as the low-density language because of the insufficient digitized text element is available in the Bengali language. But nowadays Bengali Natural Language Processing (BNLP) is the most interesting research field and AI (NLP) filed in Bangladesh. Developing an ideal Chatbot is mandatory because it needs a huge database of Bangla Language and must give reasonable answers to all interactions. Due to lack of availability of large Bengali conversation corpus, for the prototype, the system enhanced with a sample hand-crafted dataset for basic interactions. So the main goal of our task is to develop an ideal chatbot in Bengali which will able to interact in Bangla Language as well as building a corpus in Bangla. We name it "Doly".



Fig. 2. Bengali Alphabat.

IV. Proposed Work

We propound a simple encoder and decoder based conversational method of an intelligent agent that will provide for chatbot users an existence with Bengali knowledge base by interactively asking for its characteristic. In the section of BNLP helps to encode or decode Bengali corpus for training, input, and output return. BNLP is also imparted for the pre-processing questions and answer.

In the beginning, 'Doly' is nothing but an unskillful chatbot which introduction herself with no sense of how to connect with conversations. In a certain moment a user inputs affirmation, the store protects the affirmation that they are

given and the affirmation that was in reply. As 'Doly' accepts more input the number of affirmation protected in the library and that it can reply and the accuracy of each answer increase the relation with the input affirmation.

The system chooses the best closest matching reply by investigation the closest matching acquainted statement that matches the input; it then chooses an answer from the chosen of acquainted reply to that affirmation.

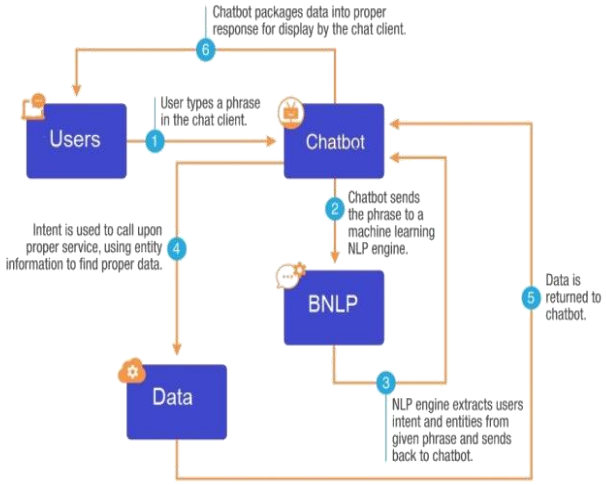


Fig. 1. Propose Model

A. Bengali Corpus

For dialog corpus of Bengali chatbot Doly we made some simple Bengali Language corpus which is saved by .yml or yaml file. YAML is an acronym that keeps for "YAML Ain't Markup Language" which is very immensely common, intelligent, human-friendly and ideal for data serialization for all programming languages. To complete our novel task, we collected 2393 questions and its related answer as our inserted data of Bengali eductions. We separated the whole data as their related group.

```
categories:
- scinece
conversations:
- - ভর কাকে বলে?
  - কোন বস্তুতে পদার্থের পরিমাণকে ভর বলে
- - পদার্থ কি?
  - যার ভর আছে ,স্থান দখল করে,বল প্রয়োগ করলে বাধাঁ সৃষ্টি করে এবং যা পাঁচটি ইন্দ্রিয় দ্বারা অনুভব করা যায় , তাকে পদার্থ বলে।
- - পদার্থ কই প্রকার?
  - তিন প্রকার। যথাঃ (১) তরল পদার্থ (২) কঠিন পদার্থ ও (৩) বায়বীয় বা গ্যাসীয় পদার্থ।
- - মোল কাকে বলে ?
  - দার্থের পারমানবিক ভর বা আণবিক ভরকে গ্রামে প্রকাশ করলে যে পরিমান পাওয়া যায় তাকে সংশ্লিষ্ট পদার্থের এক মোল বলে
- - হাইড্রোজেনের এক মোল সমান কত গ্রাম?
  - ১.০০৮ গ্রাম
- - রেখা কাকে বলে? যার দৈর্ঘ্য আছে কিন্তু ,প্রস্থ ও বেধ বা উচ্চতা নেই তাকে রেখা (line)বলে।
  - যার দৈর্ঘ্য আছে কিন্তু ,প্রস্থ ও বেধ বা উচ্চতা নেই তাকে রেখা বলে।
- - বক্ররেখা কাকে বলে?
  - যে রেখা এক বিন্দু থেকে অন্য বিন্দুতে যেতে দিক পরিবর্তন করে অর্থাৎ সোজাসুজি চলে না তাকে | বলে।
```

Fig. 3. Bengali Corpus use in Doly Chatbot.

B. Environmental setup

Natural Language Processing method and Natural Language Toolkit for Python are provided developing an apparatus in order to explore speech, intelligent explanation and human-like responses. While constructing of Bengali chatbot 'Doly', we introduced Python 3.7.0 to our Machine. Python is a high-level object-oriented language (OOP) which is suitable for scientific examination and tools development. We took the Anaconda as the apportionment of Python. Anaconda creates the best stage for open source data science which is powered by Python. For Natural Language Processing (NLP) we installed Natural Language Toolkit (NLTK). We also need to install ChatterBot that comes with built-in adapter classes that allow it to connect to various types of databases.

Python 3.7.0 is used for the implementation of Bengali chatbot 'Doly' because of providing "Unicode Decode Error". Unicode Decode Error is a runtime error which is caused by non-English language like Bengali with a large number of letters in the alphabet. The Unicode range of Bengali is 0980–09FF which has 11 vowels and 40 consonants. We use encoding="utf8" to encode the Bengali Language Unlike English Language, it has consonant conjuncts, modifier, and other graphemes. So Bengali cannot be set up with ASCII decoding system. In many pre-processing steps, NLTK has no value for the Bengali Language such as to remove a stop word, lemmatization etc. In this case, we've created BLTK for the preprocessing all inserted questions and answer.



Fig. 4. Official Unicode Consortium Code Chart of Bengali.

C. Procedure

We've created the train function adapter to train the Doly by encoding our corpus from bot data. It this case, we've applied the encoding ="utf8" in our .yml corpus to read Bengali data. To take input we've created an input adapter and for output, we've also created an output adapter to generate automated responses to a user's input. This makes it easy to create chatbots and automate conversations with users. The task divided into four parts, training data, input data, process input and return response or output.

1) Pre-Processing: To enable data for the training need to be pre-processed widely. We've applied various types of pre-processing techniques for the best deal with an answer such as anaphora, removing stop words and lemmatization. 'Anaphoric' which describe to a word that is led in a sentence and replacing pronoun from a noun. It needs a successful identification and resolution of Natural language processing (NLP). In the proposed chatbot, we have limned a review of the task done in the field of anaphora resolution in the Bengali language which is the influence of cases on pronouns, mainly personal pronoun. Using the Hobbs' algorithm of Anaphora resolution. Another preprocessing technique is to remove stop words. A stop word removing

is an action which commonly applied to delete a word that does not impress on documents or sentences for an instance of Bengali stop words. The other preprocessing process is lemmatization. Lemmatization is an efficient process to turn a word to its root word. It's one kind of Natural Language Processing and active for imparting in several purposes such as text mining, questions replying or Chatbot and others that boost to deal with root words of any language. In our proposed Chatbot, we've used two novel techniques for lemmatization based on data structure and the hash table of computer programming. The novel techniques are 'BDSEA' and 'Trie'. 'DBSRA' refers to 'Dictionary Based Search by Removing Affix' which is a really easy concept and more suitable for the lemmatization of Bengali words with the lowest time and space complexity.

2) Training Via Corpus Data: This training action authorizes to 'Doly' to be trained through a catalog of strings where the catalog illustrates a conversation. In this case, the order of each response is based on its allocation from a conversation or the list of strings. To instruct the data or corpus we've imparted the train function. In Doly, a lot of corpus data and effectiveness module that trained our chatbot in quickly for communication. We have to simply individualize the corpus data modules that we want to use. This training function approves the 'Doly' to be instructed using data from the dialog corpus that we made. While working with string type data in Python, it is practical to encounter errors. In this case, we use 'utf8' for encoding our corpus. When a chatbot trainer has provided with a data set, it builds the essential way in the chat bot's knowledge graph so that the questions or statement inputs and responses are correctly enacted. It sorted all the states from the corpus like graph sorted database in Fig 5.

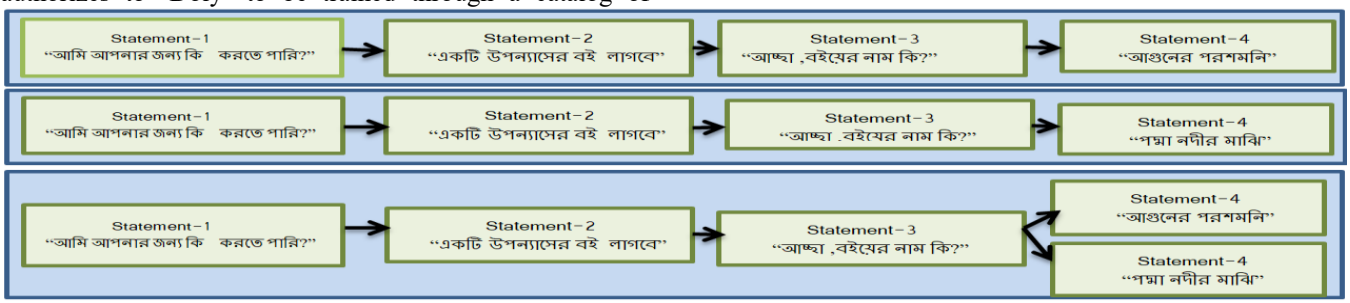


Fig. 5. Training data.

3) Input Adapters: Doly's input adapters are developed to authorize the chatbot to have a versatile method of adoption or restoring input from a given source. Input adapter is applied for getting input from some source, and then to convert it into a format that 'Doly' can access. So we created a format for the affirmation of our inputted data.

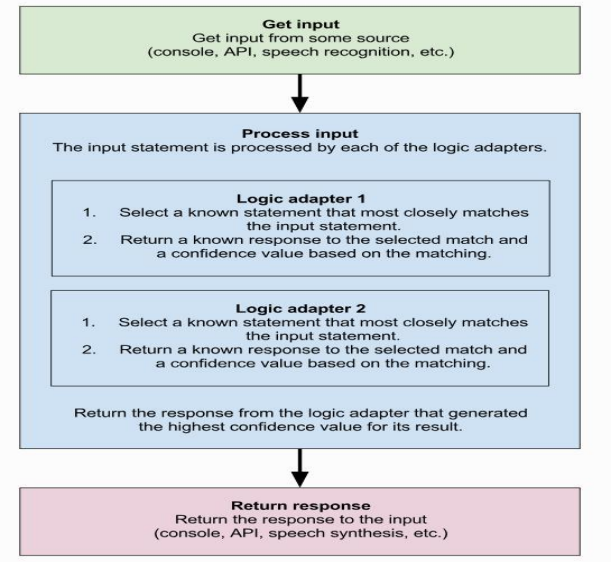


Fig. 6. Process flow diagram of Doly.

4) Process Unit: We've created the procedure of the adopted input. The input affirmations have been applied by an algorithm which would find the best likelihood valued reply for the query. The algorithm has been selected all the acquainted affirmations that most closely resemble with the input affirmation. It will return the acquainted repercussion to the selected match and a confidence estimate value based on matching after computation of each of the response. Here the confidence score is the likelihood value of the response. The algorithm will return the explanation that generated the highest likelihood value for itself.

5) Algorithams: To create 'Doly' we apply different types of machine learning methods to yield its responses. We also use machine learning algorithm like search algorithm for finding the appropriate list of matching results from the corpus and also use a Bayesian algorithm to choose the best possible answer form list of matching result. In probability theory, Bayes' theorem is also known as Bayes' law or Bayes' rule which recounts the probability of a happening, based on previous knowledge of conditions that may be related to the happening. The Bayesian hypothesis is the most common application of Bayes' theorem that is a statistical hypothesis. Testing of the drug, productions, materials as well as computing the entire output of a company based on machines are the one kinds of example.

This theorem states mathematically in the following equation:

$$M(A \mid B) = (M(B \mid A)\ M(A))/(M(B))$$

where A and B are events and $M(B) \neq 0$ and $M(A \mid B)$ is a conditional probability of event A happening given that B is true. Similarly, $M(B \mid A)$ is also a conditional probability happening given that A is true. Yet M(A) and

M(B) are the probabilities of watching A and B independently of each other which is called the marginal probability. In a general way some of the various machine learning and artificial intelligence way that are implemented throughout Doly's methodology.

Search Algorithm: A searching system is a basic principle to find out data form given corpus or database in artificial intelligence. Search is the most momentous part of how Doly quickly and efficiently recuperate the possible statements that it can respond to the user. A few ways of characteristics that help the Doly select a response include

I.The similarity of an input statement to known statements

II.The frequency in which similar known responses occur.

III.The likeliness of an input statement to fit into a category that known statements.

Classification Algorithm: In various types of logical adapters 'Doly' make sense to conduct Naive Bayesian classification algorithms to calculate if an input question is gathered to a particular set of criteria that take a response to be generated from that logic adapter. So simply we can mention that Doly is a Search-based artificial intelligence technique. For the sake of Search-based artificial intelligence, there are also some following optional algorithms implemented in Doly that can be used as the 'statement_comparison_function' for a logical adapter.

Levenshtein distance: Levenshtein distance or editing distance is dynamic programming algorithms where we can measure the minimum number of single characters are dissimilar between two words or strings. In Edit distance there are different sets of string operations such as removal, insertion, or substitution. The Edit distance follow some properties such as (a)A non-negative cost has been edited by every operation (b)For each operation, there is a reverse operation with equal costs.

There are several applications in Levenshtein Distance such as Natural Language Processing and Computational biology.

$$\mathrm{lev}_{a,b}(i,j) = \begin{cases} \max(i,j) & \text{if } \min(i,j) = 0 \\ \min \begin{cases} \mathrm{lev}_{a,b}(i-1,j)+1 \\ \mathrm{lev}_{a,b}(i,j-1)+1 \\ \mathrm{lev}_{a,b}(i-1,j-1)+1_{(a_i \neq b_j)} \end{cases} & \text{or} \end{cases}$$

Equation 1: Levenshtein Distance between two string

Here $a_i and\ b_j$ is an exponent. If $(a_i = b_j)$means the contained character of ai and bj are not unequal there are unchanged and edit=0. Otherwise if $(a_{i!} = b_j)$ then edit=1. If we want to strings lemmatization than we have to check the entire corpus of root words with the inserted sentences through the help of minimal edit distance algorithm. The minimal edit can be greater than zero. If a corpus's word and an inserted string are equal then the edit will be zero or it is changeable.

Synset Distance: For calculating the synonymous similarity between statements, NLTK provides Synset distance. In Bengali words, we developed a corpus of synonymous similarity which covers to implement for data.

Jaccard Similarity: The Jaccard index also called the Jaccard similarity coefficient is a statistic for comparing the similarity and diversity of sample sets which are distant. The Jaccard similarity measured the similarity between finite sample sets and is defined as the ratio between the size of the intersection and union of sample sets. If P and Q be two sets the Jaccard index formula is given

$$\mathrm{J_{index}}(P,Q) = \frac{|P \cap Q|}{|P \cup Q|} * 100 = \frac{|P \cap Q|}{|P|+|Q|-|P \cap Q|} * 100$$

$$\text{here, } 0 \leq \mathrm{J_{index}}(P,Q) \leq 1$$

The Jaccard distance measures dissimilarity between sample sets that are complementary to the Jaccard coefficient.

$$\mathrm{J_{distance}}(P,Q) = 1 - \mathrm{J_{index}}(P,Q) * 100$$

Now, let an example using set notation and Venn-diagram

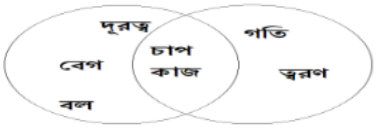


Fig-2: Jaccard similarity.

P= { দূরত্ব, বেগ, বল, চাপ, কাজ },Q= { কাজ, চাপ,গতি,ত্বরণ}

$$\mathrm{J_{index}}(P,Q) = \frac{|P \cap Q|}{|P \cup Q|} * 100$$

$$= \frac{|\{\text{কাজ, চাপ}\}|}{|\{\text{দূরত্ব, বেগ, বল, চাপ, কাজ, গতি, ত্বরণ}\}|} * 100$$

$$= \frac{2}{7} * 100 = 28\%$$

$$\text{And } \mathrm{J_{distance}} = 1 - \mathrm{J_{index}}\% = 1 - 28\% = 72\%$$

So,we conclude four postulates such as (a) If $J_{index} = 50$ similar which means the sets share all members.(b)If $J_{index} \ll 50$ similar which means the sets share few members. (c) If $J_{index} \cong 50$ similar which means the sets share half of the members. (d) If $J_{index} = 0$ similar which means the sets share no members.

D. Logic Adapters

The Logical adapters are identified as the logic for making responses of a conferred input statement or question by a Chatbot. The logical adapter which Doly used can be determined by applying the logic adapters parameter to the recruitment path of the logic adapter that needs to use. It is pretty possible to input any number of logic adapters for our chatbot in order to make jurisprudence. If multiple adapters are exercised, then the chatbot will retire the response with the best-enumerated confidence value.

E. Response Selection Method

Response selection methods determine which response should be used in the event that multiple responses are generated within a logic adapter. Doly uses Statement objects to hold information about things that can be said. An important part of how a chat bot selects a response is based on its ability to compare two statements to each other. This module contains various text comparison algorithms designed to compare one statement to another.

We use the get first response method for the selection of a response. This method takes the input statement and selects the statement in the knowledge base which closely matches the input to the chatbot from a list of statement options to choose a response from.

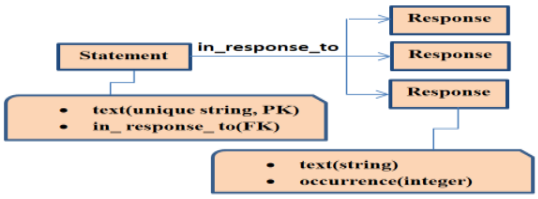


Fig. 7. Response System.

V. Results & Analysis

The difficulty of appraisal is intrinsic as each conversation having an effect on each other and the equivalent conversation will not happen to exceed once; one slightly different answer will lead to a completely different conversation [12]. So, for the excellent improvement of the chatbot, we compared our chatbot system with previous existing chatbots. When we checked the 2776 questions releted to inserted corpus on our chatbot system ,it returned 2448 correct reply.That means 88 % accurate.

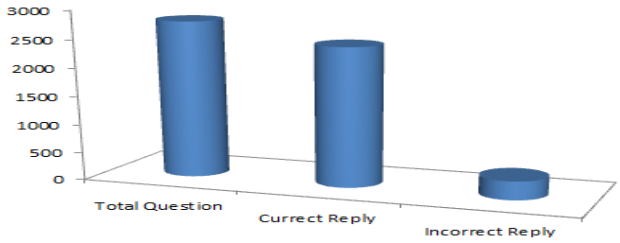


Fig. 8. Results Analysis.

VI. Comparison Between Others Chatbook

Our purpose is creating Bengali chatbot in the sake of the educations system of Bangladesh. So, we compared 'Doly' with two chatbots which are Neural Conversational Machine (NCM) and Cleverbot. Neural Conversation Machine is a chatbot which can be developed with the end-to-end and this model chats by predicting the next sentence given the previous sentences or sentence with a few handcrafted rules [12]. So NCM is a generative based conversational agent whereas our task is a retrieval-based system. We give some random questions to the NCM and our system for the comparative study between the two. It responds to the input by finding how a human responded to that input when it was asked, in part or in full, by Cleverbot. In this case, 'Doly' acted pretty good and responded with the best way.

VII. Future Work

In future works, we will try to enable 'Doly' among all Bengali native speaker students as in the sake of effective educations tool with the Bengali language. We will create a chatbot with the help of neural network technique as well as a sequence to sequence chatbot via Bengali Language Corpus. Also, we will try to generate own large Bengali Language corpus and try to develop a great system for Bangla Natural Language Processing (BNLP).

VIII. Conclusion

This research is a pioneering work in the field of dialogue system in Bengali. The main challenge of this work is to create a chatbot based on the accurate knowledge base. Due to the lack of large dataset, we implemented a retrieval based closed domain chatbot which will converse with the user based on the pattern matching algorithm and will improve its performance measure by learning from the interaction. Our work will provide a Bengali conversation corpus which will assist in the construction of a system for Bengali Language Processing research.

IX. References

[1]. Weizenbaum, J.: Contextual understanding by computers. Commun. ACM 10(8), 474–480 (1967)

[2]. Weizenbaum, J.: ELIZA—a computer program for the study of natural language communication between man and machine. Commun. ACM 9, 36 –45 (1966)

[3]. ALICE:Artificial Intelligence Foundation. http://www.alicebot.org/bios/rechard wallace.html

[4]. Breazeal, C.: Designing Sociable Robots. MIT Press, Cambridge (2002)

[5]. Vinyals, O., Le, Q.: A neural conversational model. In: Proceedings of ICML Deep Learning Workshop (2015)

[6]. Bickmore, T.W., Picard, R.W.: Establishing and maintaining long-term human-computer relationships. ACM Trans. Comput. Hum. Interact. (TOCHI) 12(2), 293–327 (2005)

[7]. Jabberwacky. http://www.jabberwacky.com

[8]. Bickmore, T., Schulman, D., Vardoulakis, L.: Tinker - a relational agent museum guide. J. Auton. Agents Multi Agent Syst. 27(2), 254–276 (2013)

[9]. Bickmore, T., Schulman, D.: Practical approaches to comforting users with relational agents. In: ACM SIGCHI Conference on Human Factors in Computing Systems (2007)

[10]. Cassell, J.: Embodied conversational agent: representation and intelligence in user interfaces. Al Mag. 22(4), 67–83 (2001)

[11]. Augello, A., Gentile, M., Weideveld, L., Dignum, F.: A model of a social chatbot. In: De Pietro, G., Gallo, L., Howlett, R.J., Jain, L.C. (eds.) Intelligent Interactive Multimedia Systems and Services 2016. SIST, vol. 55, pp. 637–647. Springer, Cham (2016). doi:10. 1007/978-3-319-39345-2_57

[12]. Hill, J., Ford, W.R., Farreras, I.G.: Real conversations with artificial intelligence: a comparison between human–human online conversations and human–chatbot conversations. Comput. Hum. Behav. 49, 245–250 (2015)

[13]. Pan, X., Slater, M.: A preliminary study of shy males interacting with a virtual female. In: Presence 2007: The 10th Annual International Workshop on Presence (2007)

[14]. Crutzen, R., Peters, G.-J.Y., Dias Portugal, S., Fisser, E.M., Grolleman, J.J.: An artificially intelligent chat agent that answers adolescents' questions related to sex, drugs, and alcohol: an exploratory study. J. Adolesc. Health 48, 514–519 (2011)

[15]. Shawar, Bayan Abu, and Eric Atwell. *A comparison between Alice and Elizabeth chatbot systems*. University of Leeds, School of Computing research report 2002.19, 2002.

[16]. Abdul-Kader, Sameera A., and J. C. Woods. "Survey on chatbot design techniques in speech conversation systems." *International Journal of Advanced Computer Science and Applications* 6.7 (2015).

[17]. Shrestha, Lokesh, and Kathleen McKeown. "Detection of question-answer pairs in email conversations." *Proceedings of the 20th international conference on Computational Linguistics*. Association for Computational Linguistics, 2004.

[18]. Zhou, Liang, and Eduard Hovy. "Digesting virtual geek culture: The summarization of technical internet relay chats." *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 2005.

Threat and Abusive Language Detection on Social Media in Bengali Language

Puja Chakraborty, Md. Hanif Seddiqui
Department of Computer Science and Engineering
University of Chittagong
Chittagong-4331, Bangladesh
puja.csecu@gmail.com, hanif@cu.ac.bd

***Abstract*—Threat and abusive languages spread quickly through social media which can be controlled if we can detect and remove them. Since there exist many social media like Facebook, Twitter, Instagram etc and a huge number of social media users, we need a robust and effective automatic system to identify threat and abusive languages. In our proposed system Machine Learning and Natural Language Processing techniques have been implemented to build an automatic system. Previous research on Bengali abusive language detection used Multinomial Naïve Bayes (MNB), Support Vector Machine(SVM) algorithms and considered Bengali Unicode characters to build their system. We considered both Unicode emoticons and Unicode Bengali characters as valid input in our proposed system. Besides MNB and SVM algorithm, we implemented Convolutional Neural Network (CNN) with Long Short Term Memory(LSTM). Among three algorithms, SVM with linear kernel performed best with 78% accuracy.**

***Index Terms*—Natural Language Processing (NLP), Multinomial Naïve Bayes (MNB), Support Vector Machine(SVM), Convolutional Neural Network (CNN), Long Short Term Memory (LSTM), Bengali Text Classification, Bengali Hate speech Identification**

I. INTRODUCTION

Social media is an interactive platform to share important information, opportunities, news and good thoughts. But one of the major problems of social media is threat and abusive language which can also be referred as cyberbullying. A text can be considered as threat or abusive if it contains sexist or racist slurs, attacks or criticizes any community or religious view, provokes criminal activities etc. Offensive language has a fatal detrimental effect on teenage people which even can lead them to suicide. Bully victims are more likely to (between 2 and 9 times) commit suicide than non-victims [1].

Removing offensive and abusive data can reduce the negative effects from society. It was reported that 4.021 billion people accessed internet and 3.196 billion people used social media in 2018 and the rate of internet and Facebook users increases by 7 percent and 13 percent each year respectively[2]. According to Bangladesh Telecommunication Regulatory Commission (BTRC), the total number of internet subscribers in Bangladesh has reached 91.348 million at the end of December, 2018 [3]. Since so many people use social media, checking each and every post and/or comment manually is not possible. So, we need a strong and effective automatic system which can detect whether a post and/or comment is abusive or not.

Many machine learning and Natural Language Processing(NLP) techniques have been used to detect hate speech or abusive languages. Support Vector Machine (SVM) classifier with unigram, bi-gram, tri-gram features weighted by its TF-IDF is used to classify hate speech and offensive language by Thomas Davidson, Dana Warmsley, Michael Macy, Ingmar Weber [4]. Zeerak Waseem and Dirk Hovy proposed a model to detect hate speech from Twitter in which they detected 10 most repeated words for each class and used as classification feature. Besides, gender and location of the users were also collected which were also used in their proposed system [5]. Maral Dadvar and Franciska de Jong also used gender specific language features and implemented the classification approach using SVM algorithm in their research [6]. Template based strategy to generate features was used by William Warner and Julia Hirschberg in which they implemented 10-fold cross validation SVM and obtained 94% accuracy [7]. In another approach, prposed by Z. Rui Zhao and K. Mao, Latent Semantic Analysis (LSA) was used to extend pre-defined insulting words which was further used for classification [8]. In a different research, Multinomial Naïve Bayes (MNB) and SVM were used to detect abusive Bengali text with count vector and TF-IDF by Shahnoor Chowdhury Eshan and Mohammad Shahidul Hasan [9].

For our proposed system, we have collected data from different Facebook pages. Prothom Alo [10] and BBC News Bangla [11] are most popular Facebook pages of Bangladesh [12]. A great number of data have been collected from these two pages. We considered both Unicode Bengali words and emoticons as valid input in our proposed system. Besides, consecutive exclamation and question marks are taken in special consideration. After preprocessing, we used three different techniques, MNB, SVM and CNN-LSTM, to classify the data as threat and abusive or not.

This paper is divided into different sections. A brief description of related work in this field is mentioned in section II, Dataset preparation is explained in section III, classification approach of our proposed system is illustrated in section IV. Experiment and evaluation of the proposed systems is described in section V and finally the paper concludes in section VI.

II. Related Work

For the last couple of years, many researches have been conducted to detect hate speech automatically and most of the related works are based on English text.

Thomas Davidson, Dana Warmsley, Michael Macy, Ingmar Weber (2017) conducted a supervised learning on Twitter dataset [4]. They classified their training data into three classes, namely *Hate*, *Offensive* and *Neither*. After converting each tweet into lowercase character, they used stemmer, logistic regression with L1 regularization, linear SVM etc classification tools which resulted in 90% F1 score. This model was trained on 24802 labeled data and the data were biased. Only 5% of the tweets belonged to *Hate* class and as a result 40% of *Hate* data were misclassified. To train any type of model, the data should be balanced. Our proposed system is a binary classification problem where we used balanced dataset.

Jin Wang, Liang-Chih Yu, K. Robert Lai and Xuejie Zhang (2016) used regional CNN-LSTM model for sentiment analysis where they used Stanford Sentiment Treebank (SST) (Socher et al., 2013) and Chinese ValenceArousal Texts (CVAT) (Yu et al., 2016) as their dataset [13]. In their research they concluded that among lexicon-based method, regression based method, CNN, RNN, LSTM and regional CNN-LSTM model, for both English and Chinese dataset, regional CNN-LSTM model resulted in minimum error because CNN tracked down local information and LSTM found long distance dependency.

Another system, proposed by Shahnoor Chowdhury Eshan and Mohammad Shahidul Hasan(2017) is based on Bengali Unicode text implemented both count vector and TF-IDF for unigram, bi-gram and tri-gram. Then MNB and SVM (linear and RBF) classifiers were used[9]. This system only dealt with Unicode Bengali words while we considered both words and emoticons. Besides, consecutive exclamation marks and questions marks are considered as single word in our system.

III. Dataset

Total 5644 (*Table I*) posts and comments have been collected from different pages of Facebook which include Facebook page of newspapers, celebrities, entertainment pages etc. Among the collected data approximately 50% are threat and abusive type (which are labeled as "YES" in the dataset) and the rest of the data are not abusive (which are labeled as "NO" in the dataset).

TABLE I
Dataset

NO	YES	Total
2905	2739	5644

IV. Classification Approach

Overview of our classification approach is exposed in **Fig 1**. Dataset is preprocessed and after that the training data are used to train the model. Finally, the trained model predicts the class of test data and based on the predictions the performance of the models are calculated.

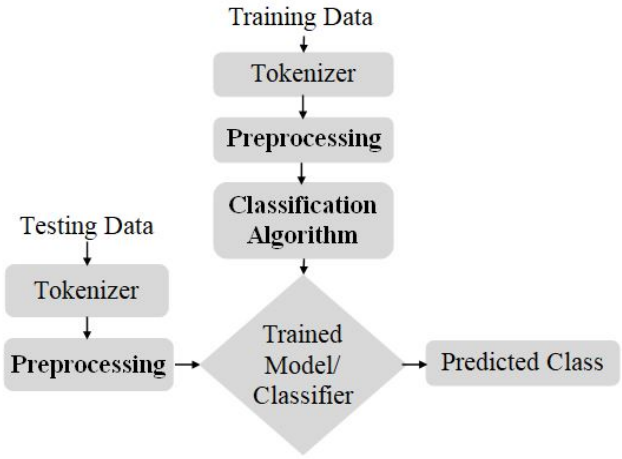


Fig. 1. Classification Framework

A. Preprocessing

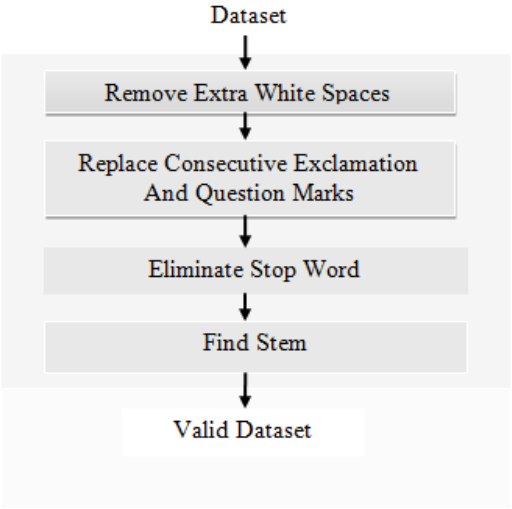


Fig. 2. Data Preprocessing

Input data are processed before they are used to train and test the models. Basically, preprocessing(**Fig 2**) step is a combination of few individual steps.

In the first step of preprocessing extra spaces from each document are removed. Consecutive exclamation marks (e.g '!!!', '!!!!!') and question marks (e.g.'??','????') are considered as a single word because consecutive exclamation and question marks do not represent exclamatory or interrogative emotions always, rather they are used to express anger, disgust or negative emotions. In the code, sequential exclamation marks are replaced with 'exn' and question marks are replaced with 'qsn' so that consecutive marks can be identified uniquely. Unicode

emoticons, Bengali Unicode, exclamation(!) and question() marks are considered as valid input. All special characters i.e. '$', '#', '@' etc and punctuation marks are removed. Stop words are removed as well. Finally a Bengali stemmer[14] is used to find the stem of the Bengali words.

B. Classifiers

Multinomial Naïve Bayes (MNB), Support Vector Machine (SVM) and CNN-LSTM are used as classification algorithm in our proposed system.

1) Multinomial Naïve Bayes (MNB): For **Fig 3**, n-grams ranged up to three(unigram, bigram, trigram) are considered to calculate TF-IDF. Thereafter prior probability of each class is determined based on the training documents. Then the conditional probability of each n-gram (which are referred as token in Fig 3) is computed. The probability that a test document should be of "YES" type and "NO" type both are calculated based on the prior probability of each class and the conditional probability of the tokens contained in the test document. Between them, the class with maximum probability is selected as the label of the test document.

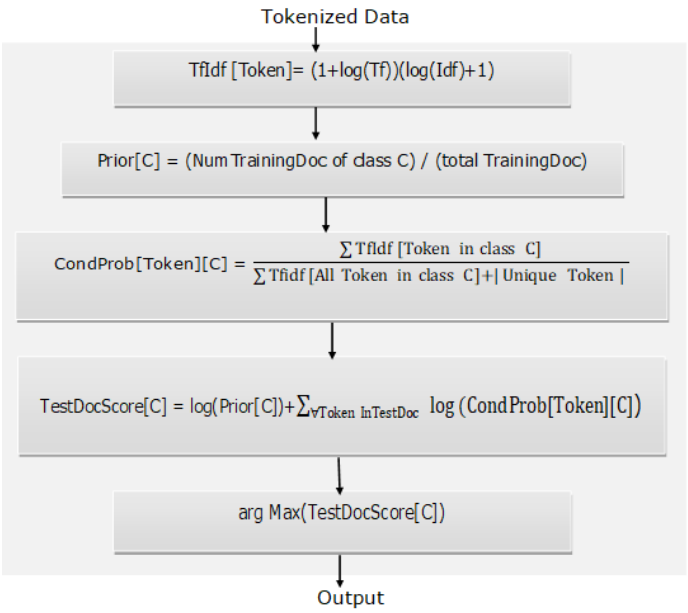


Fig. 3. MNB Classifier

2) Support Vector Machine (SVM) : According to **Fig 4**, same feature extraction techniques (TF-IDF of n-gram ranged up to three) as MNB classifier is used. The notations used in Fig 4 are stated below:

$W : Word\ Vector$
$X : Input\ Vector$
$b : Bias\ Value$
$y : Document\ Labels$
$e : Vector\ of\ all\ ones$
$\alpha : Langliers\ Multiplier$

At first each document is mapped into a higher dimensional space. Then we need to find out a hyperplane which will divide the documents on the high dimensional space in two region such that each region contains only one type of document. In our system, dual form of primal optimization problem is solved. W is determined after α is calculated. Finally after finding the support vectors, $b, y_i\alpha_i \forall i$ these values are stored which are used in the prediction of test document.

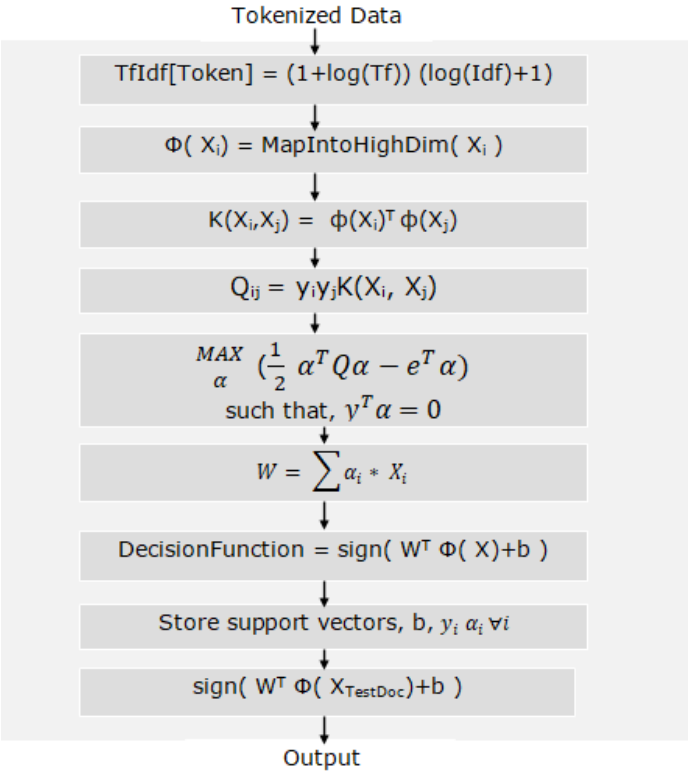


Fig. 4. SVM Classifier

3) CNN-LSTM: Each word, emoticons or consecutive exclamation or question marks are considered as token for CNN-LSTM network. For **Fig 5**, each unique token of data-set is converted to an integer value before feeding to the network. After that the model is designed with four layers namely embedding layer, convolution layer, LSTM layer and a fully connected layer. Each consecutive layer is denoted by different shades in **Fig 5**.

The document which has maximum length is set as the input length for embedding layer. Each token is converted to a word vector of length 32 and therefore each document is represented as a list of vectors.

In convolution layer, 32 different filters each with length 3 is added. Every filter overlays over the word vectors and performs an element-wise multiplication for all its 3x32 elements, and then aggregate up the multiplication result and obtain one numeric value which is stored as the element of the output sequence.

In LSTM layer 100 hidden nodes are used where each node consists of a forget gate, an input gate, an output gate and a

combined gate to calculate current cell memory. Output from each node of LSTM layer are used as the input in the next layer where sigmoid function is used as the activation function to give final output.

In total we used 50 epochs and the batch size was 128. The dataset is divided into training dataset, validation dataset and testing dataset.

From training dataset, at a time 128 data are used to train the model which are fed to the network. After that the whole process described in **Fig-5** is conducted on those data and based on the loss obtained from this output the parameters in different layers are updated. Then, the next 128 data are fed to the network and the same process repeats until each and every documents in the training dataset are passed through the network. All these steps form a single iteration.

The symbols used in the LSTM layer of the process in Fig-5 are stated below:

$X : Input\ Vector$
$H_{t-1} : Previous\ Cell\ Output$
$C_{t-1} : Previous\ Cell\ Memory$
$H_t : Current\ Cell\ Output$
$C_t : Current\ Cell\ Memory$
$W : Weight\ Vector$
$B : Bias\ Value$

The target of neural network is to minimize the loss of the system. After training the network, documents of the test dataset are passed through the network and it predicts the labels of the documents. Based on the prediction, the accuracy of the system is determined which is stated in **Table VII**.

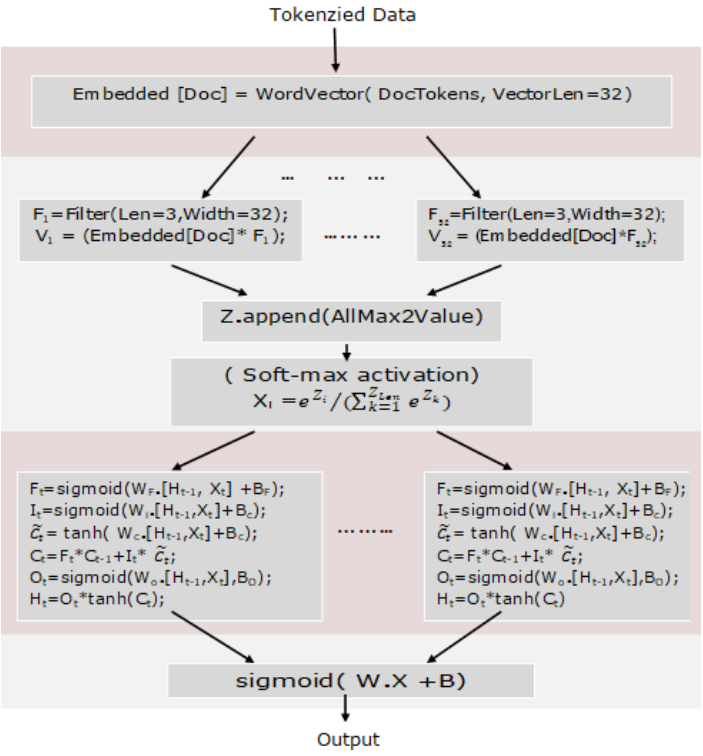


Fig. 5. LSTM Classifier

V. EXPERIMENT AND EVALUATION

A. MNB and SVM Performance

According to **Table II**, with 1426 documents MNB classifier is able to detect 99% of *NO* type documents accurately where it detects only 6% of *YES* type documents correctly. If we observe the precision rate, we see that only 40% of the test documents which are predicted as *NO* by MNB classifier are actually *NO* type documents and among the predicted *YES* type documents, 97% are correctly predicted. Linear SVC classifier correctly predicts 87% and 42% documents of *NO* class and *YES* class respectively. According to the precision rate of Linear SVC classifier, among the predicted documents 45% of *NO* type documents and 85% of *YES* type documents are predicted right. RBF SVC classifier correctly detects 63% of *NO* type data and 70% *YES* type data. It should be mentioned here that among the test data, 35% are *NO* type documents and the rest 65% are *YES* type documents. Besides training data are biased toward *NO* type class.

From **Table III**, we can see that with the increased data, the performance of Linear SVC has improved. Where the performance of other two classifiers have decreased. Linear SVC identified 46% *NO* type and 47% *YES* type documents accurately. Both Linear SVC and RBF SVC have same F1 score and accuracy in this case. Though MNB classifier predicts almost 99% *NO* type documents correctly, it predicts only 17% of total *YES* type documents right. Among the documents which are predicted as *NO* by MNB classifier, 60% of them are actually belong to *YES* type class.

According to **Table IV**, accuracy of MNB classifier was highest than other two classifier. Here MNB classifier detects 97% of *NO* class documents accurately. Though the documents that are labelled as *NO*, among them 61% documents are actually of *NO* type. Linear SVC and RBF SVC identify 77% and 70% *NO* class data properly. On the contrary, MNB classifier could recognize only 38% *YES* type data but linear SVC and RBF SVC could identify 55% and 62% *YES* type documents perfectly. One thing that should be mentioned, any stemmer was not used up to these three runs. Besides the data were biased. Among 3604 data, 3404 data were used as training data where 63% of the data were labeled as *NO* and 27% data were of *YES* class.

Using a Bengali stemmer and running the module on the same 3604 data we noticed that Linear SVC performed better than MNB and RBF SVC classifier. From **Table V**, we observe that linear SVC performed better to identify both of the classes where MNB and RBF SVC could identify data of one type of class much better than other type of class. If we compare with Table IV, we notice that for linear SVC the percentage of correctly identifying *NO* and *YES* type documents has increased from 77% to 79% and from 55% to 67% respectively.

After replacing consecutive exclamation and question marks with a specific word(which is less prone to appear in dataset)

we trained the models again and the obtained result is displayed in **Table VI**. We can see that MNB classifier is able to detect 95% of *NO* type documents and only 5% of *YES* type documents accurately. Here, SVM could identify both type data in a balanced way.

TABLE II
MNB, SVC, RBF COMPARISON WITH 1426 DATA

(a) Confusion Matrix

	MNB Classifier		Linear SVC		RBF SVC	
	NO	**YES**	**NO**	**YES**	**NO**	**YES**
Actual Class	0.99	0.01	0.87	0.13	0.63	0.37
	0.81	0.19	0.58	0.42	0.30	0.70
	Predicted Class					

(b) Efficiency Parameter

	MNB Classifier		Linear SVC		RBF SVC	
Class/Label	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Precision	0.40	0.97	0.45	0.85	0.54	0.78
Recall	0.99	0.19	0.87	0.42	0.63	0.70
F1 Score	0.57	0.31	0.60	0.57	0.560	0.57
Accuracy(%)	47.16		58.19		67.89	

TABLE III
MNB, SVC, RBF COMPARISON WITH 3324 DATA

(a) Confusion Matrix

	MNB Classifier		Linear SVC		RBF SVC	
	NO	**YES**	**NO**	**YES**	**NO**	**YES**
Actual Class	0.99	0.01	0.84	0.16	0.85	0.15
	0.83	0.17	0.53	0.47	0.54	0.46
	Predicted Class					

(b) Efficiency Parameter

	MNB Classifier		Linear SVC		RBF SVC	
Class/Label	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Precision	0.40	0.97	0.46	0.84	0.46	0.84
Recall	0.99	0.17	0.84	0.47	0.85	0.46
F1 Score	0.57	0.29	0.60	0.60	0.60	0.60
Accuracy(%)	46.15		59.87		59.87	

B. CNN-LSTM Performance

From **Table VII**, we noticed that accuracy of LSTM network has increased with the increase of volume of dataset and we achieved 77.5% accuracy for 5644 data. For 3604 data we obtained 65.5% accuracy and with 4933 data we got 73% accuracy.

C. Accuracy Comparison:

Fig 6 illustrates the accuracy of each models (MNB, Linear SVC, RBF SVC and CNN-LSTM) where the X-axis indicates the dataset size and the Y-axis represents accuracy in percentage. The first two cells for CNN-LSTM are intentionally left blank because the CNN-LSTM model was not build when we had only 3324 data.

TABLE IV
MNB, SVC, RBF COMPARISON WITH 3604 DATA

(a) Confusion Matrix

	MNB Classifier		Linear SVC		RBF SVC	
	NO	**YES**	**NO**	**YES**	**NO**	**YES**
Actual Class	0.97	0.03	0.77	0.23	0.70	0.30
	0.62	0.38	0.45	0.55	0.38	0.62
	Predicted Class					

(b) Efficiency Parameter

	MNB Classifier		Linear SVC		RBF SVC	
Class/Label	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Precision	0.61	0.93	0.63	0.71	0.65	0.67
Recall	0.97	0.38	0.77	0.55	0.70	0.62
F1 Score	0.75	0.54	0.69	0.62	0.67	0.65
Accuracy(%)	67.50		66.00		66.00	

TABLE V
MNB, SVC, RBF COMPARISON WITH 3604 DATA (WITH STEMMER)

(a) Confusion Matrix

	MNB Classifier		Linear SVC		RBF SVC	
	NO	**YES**	**NO**	**YES**	**NO**	**YES**
Actual Class	0.95	0.05	0.79	0.21	0.20	0.80
	0.60	0.40	0.33	0.67	0.04	0.96
	Predicted Class					

(b) Efficiency Parameter

	MNB Classifier		Linear SVC		RBF SVC	
Class/Label	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Precision	0.61	0.89	0.71	0.77	0.83	0.55
Recall	0.95	0.40	0.79	0.67	0.20	0.96
F1 Score	0.75	0.55	0.75	0.72	0.32	0.70
Accuracy(%)	67.50		73.00		58.00	

TABLE VI
MNB, SVC, RBF COMPARISON WITH 5644 DATA

(a) Confusion Matrix

	MNB Classifier		Linear SVC		RBF SVC	
	NO	**YES**	**NO**	**YES**	**NO**	**YES**
Actual Class	0.97	0.03	0.88	0.12	0.82	0.18
	0.57	0.43	0.32	0.68	0.33	0.67
	Predicted Class					

(b) Efficiency Parameter

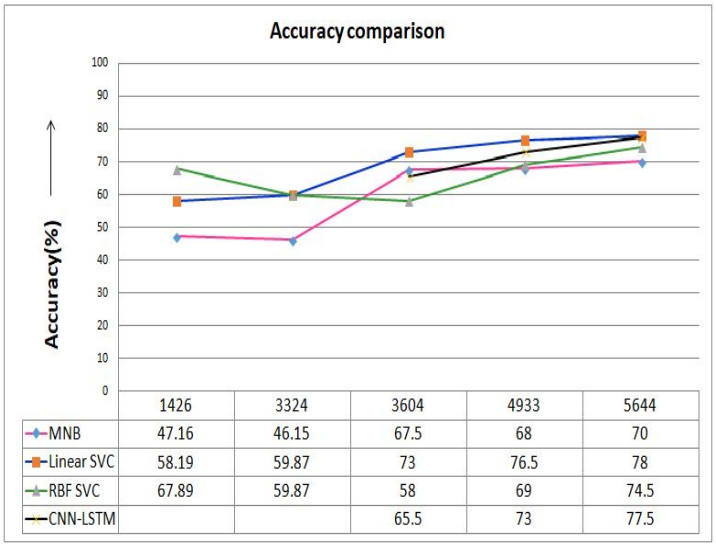
	MNB Classifier		Linear SVC		RBF SVC	
Class/Label	**NO**	**YES**	**NO**	**YES**	**NO**	**YES**
Precision	0.63	0.94	0.73	0.85	0.71	0.21
Recall	0.97	0.43	0.88	0.68	0.82	0.67
F1 Score	0.76	0.59	0.80	0.76	0.76	0.32
Accuracy(%)	70.00		78.00		74.50	

From **Fig 6** it is evident that the accuracy of each classifier, except than RBF SVC, has increased with the increased amount of data. If we notice the rate of change in accuracy, we see that when total data increased from 4933 to 5644 the

TABLE VII
CNN-LSTM ACCURACY

Data	Accuracy %
3604	65.5
4933	73.0
5644	77.5

accuracy for both MNB and linear SVC classifier has increased about 2% but CNN-LSTM model has highest rate of change in accuracy from 73% to 77.5%. Though linear SVC has highest accuracy for 5644 data, CNN-LSTM model has almost reached to the accuracy of linear SVC for the same amount of data.



	1426	3324	3604	4933	5644
MNB	47.16	46.15	67.5	68	70
Linear SVC	58.19	59.87	73	76.5	78
RBF SVC	67.89	59.87	58	69	74.5
CNN-LSTM			65.5	73	77.5

Fig. 6. Accuracy of models

VI. CONCLUSION AND FUTURE WORK

From the result of our proposed system we observed that SVM with Linear kernel performed better than other algorithms when the data is balanced. With biased data MNB classifier was able to classify only one type data more accurately. We noticed that for most of the cases, MNB failed to identify threat and abusive type data. Rather, it predicted most of the data as non-abusive. We noticed that the result of SVM classifier with RBF kernel fluctuated to a large degree. The result of SVM classifier with linear kernel was more consistent as its performance improved as we balanced our dataset and its accuracy peaked at 78%.

In future, We are going to research about other features to improve our system. Besides, we will expand our dataset. We noticed that the accuracy of CNN-LSTM increased at a large rate with larger dataset. Since CNN-LSTM model is able to focus on both local information and long dependencies, we will try to find whether CNN-LSTM can perform better or not than other algorithms with huge amount of data. We will implement other feature extraction techniques (for example word2Vec etc.) with CNN-LSTM and will compare the result so that we can improve our system.

REFERENCES

[1] "Cyberbullying statistics," last accessed 01 February,2019. [Online]. Available: http://www.bullyingstatistics.org/content/bullying-and-suicide.html

[2] "Number of facebook users worldwide till december,2017," last accessed 01 February,2019. [Online]. Available: https://wearesocial.com/blog/2018/01/global-digital-report-2018

[3] "Number of internet user in bangladesh till december,2018," last accessed 29 January,2019. [Online]. Available: http://www.btrc.gov.bd/content/internet-subscribers-bangladesh-december-2018

[4] T. Davidson, D. Warmsley, M. Macy, and I. Weber, "Automated hate speech detection and the problem of offensive language," *International AAAI Conference on Web and Social Media (ICWSM)*, 2017.

[5] Z. Waseem and D. Hovy, "Hateful symbols or hateful people? predictive features for hate speech detection on twitter," *Association for Computational Linguistics*, 2016.

[6] Dadvar, Maral, de Jong, and Franciska, "Cyberbullying detection; a step toward a safer internet yard," *WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web Companion*, 2012.

[7] W. William and H. Julia, "Detecting hate speech on the world wide web," *LSM 12 Proceedings of the Second Workshop on Language in Social Media*, 2012.

[8] R. Zhao, A. Zhou, and K. Mao, "Automatic detection of cyberbullying on social networks based on bullying features," *Proceedings of the 17th International Conference on Distributed Computing and Networking*, 2016.

[9] S. C. Eshan and M. S. Hasan, "An application of machine learning to detect abusive bengali text," *2017 20th International Conference of Computer and Information Technology (ICCIT)*, 2017.

[10] "Facebook page of daily prothom alo newspaper," last accessed 02 February,2019. [Online]. Available: https://www.facebook.com/DailyProthomAlo

[11] "Facebook page of bbc news bangla," last accessed 02 February,2019. [Online]. Available: https://www.facebook.com/BBCBengaliService

[12] "Statistics of most popular facebook page of bangladesh," last accessed 02 February,2019. [Online]. Available: https://www.socialbakers.com/statistics/facebook/pages/total/bangladesh

[13] J. Wang, L.-C. Yu, K. R. Lai, and X. Zhang1, "Dimensional sentiment analysis using a regional cnn-lstm model," *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 2016.

[14] M. H. Seddiqui, A. A. M. Maruf, and A. N. Chy, "Recursive suffix stripping to augment bangla stemmer," *International Conference Advanced Information and Communication Technology (ICAICT)*, 2016.

APPENDIX

Our collected dataset contains 5644 Facebook comments and status in Bengali language. The dataset can be provided on request. Anyone can use the dataset for their research purpose as well.

Adheetee: A Comprehensive Bangla Virtual Assistant

Syed Mohidul Islam, Most Fowziya Akther Houya, Syed Mynul Islam, Salekul Islam, Nahid Hossain
Department of Computer Science and Engineering
United International University
Dhaka, Bangladesh
Email: isyedmohidul@gmail.com, fowziyaboshra@gmail.com, syed.mynul1@gmail.com, salekul@cse.uiu.ac.bd, nahid@cse.uiu.ac.bd

***Abstract*— A Virtual Assistant is an application program that can understand natural language (audio and text) commands and perform tasks for the user. The importance and usage of the virtual assistant are growing exponentially worldwide. In this paper, we have demonstrated a comprehensive design and development methodology of a Bangla virtual AI assistant 'Adheetee (Erudite)' for smart devices such as smartphones and personal computers. Bangla is the native language of Bangladesh and some states of India with millions of native speakers worldwide. There are some popular virtual assistants available in English and other languages but not in Bangla. Therefore, using a virtual assistant become very difficult for Bangla speaking people around the world. Moreover, a large portion of Bengali speaking people is illiterate. This makes using a virtual assistant more difficult. This gives us the motivation to build a comprehensive virtual assistant for Bangla language. We have covered vast varieties of most frequent commands in our system with the accuracy of 94.065%.**

Keywords—Adheetee, Virtual Assistant, Bangla, Commands, API

I. INTRODUCTION

A Virtual assistant makes tasks of a person in a smartphone or a personal computer easier by just giving a voice command or a written command. The usage of a virtual assistant is presumed to hit 1.8 billion worldwide in 2021 [1][2]. People can get benefits of 21st century's technology by using a virtual assistant as for using it no extra knowledge is required. A user can use it just like they communicate with a human being. Intelligent natural language processing makes the system understand the natural language commands of a human being [3]. A virtual assistant can be used to set a reminder, sending email, know weather information, search information in Google, update to-do list etc. Bangla is the 6th most widely spoken language with approximately 250 million total speakers worldwide [6]. However, there is no virtual assistant yet developed for Bangla language. Euromonitor International reports only 18% of the total population in Bangladesh can understand and speak English [7]. In addition to that 27.2% of the total population is illiterate who cannot even read or write Bangla language [8]. Due to the low literacy rate and ineptness of English knowledge among the Bangla speaking people we are not being benefited properly by the English virtual assistants. Therefore, the authors decided to propose a comprehensive Bangla virtual assistant which takes both verbal and text Bangla input as a command. According to Newzoo's Global Mobile Market Report 2018 [11], there are around 27 million Smartphone users in Bangladesh. Our proposed system can help these enormous users to be able to utilize features of their smartphones more easier and faster way in their mother tongue.

Previously, a considerable amount of work had been done in virtual assistant in English and other western languages such as Google's Google Assistant, Apple's Siri, Microsoft's Cortana, and Samsung's Bixby etc. Although most of them have the same functionalities, each of them has some new features and abilities [5]. However, no virtual assistant has been designed or developed in Bangla language yet. Note that there are some chatbots available in Bangla. T.D. Orin proposed Bangla chatbot named Golpo in 2017, which can have basic text conversations [10]. In 2018, Anirudha Paul et al. proposed a closed domain contextual chatbot framework for resource-poor languages including Bangla [9]. However, these are only text chatbots for general conversations and thus have very limited functionalities.

In this paper, we have proposed a comprehensive Bangla virtual assistant for smart devices. We have named our system 'Adheetee' which is a Bangla word, meaning 'Erudite (having or showing great knowledge or learning.)'. It contains a large set of most frequent commands and variations of those commands. Our system is consists of a large database containing several types of information about different commands and their potential responses. We have distributed our commands as basic commands and core commands. Basic commands handle a basic conversation like a chatbot. However, our basic conversations are not just introductory communications; we have stored a user's important information through the conversation for future references, make the system more customized and to provide responses based on context. In the core commands, the system handles more complex device oriented activity. Handling varieties of the same command is a difficult task, however, we have used Named Entity Recognition (NER), Keyword Extraction (KE) and Cosine Similarity (CSM) algorithms to identify and respond accordingly for command variations. We have built our system to be able to respond by itself using its intelligence and the information stored in its database. However, to handle unknown commands and fetch the latest information the system sometimes depends on external API calls.

The paper is organized as follows: In section II, we describe the proposed system and step by step demonstration of the system diagram with proper examples and algorithm. The paper demonstrates experimental results in section III with examples. Section IV interprets performance from the experiments, while section V concludes the paper with limitations of our system and future work.

II. PROPOSED METHOD

Our system takes both voice command and text command as input. A person can wake up our system just by calling "হে অধীতী (Hey Adheetee)" or "ওহে অধীতী (Hello Adheetee)" or simply "অধীতী (Adheetee)". On the other hand, a user can put written text as a command. For voice command, we translate the speech into text by using a Speech to Text (STT) system. We have used Google's STT

API to do this work [12]. Then, we extract keywords from the text command. We need to extract the keywords from command due to the fact that the commands are given in natural language. Therefore, a different user can speak or write the same command in different ways. Moreover, even the same user can speak or write differently in different time periods. For example, if we observe the following commands:

a. *আমাকে আবহাওয়ার খবর বলুন?* (Give me weather update.)

b. *আজকের আবহাওয়া কেমন?* (How is the weather today?)

c. *বাহিরের তাপমাত্রা কত?* (What is the temperature outside?)

d. *অধীতী, আজ বাহিরের তাপমাত্রা কত?* (What is the temperature outside today?)

e. *কাল আমাদের এলাকার তাপমাত্রা কেমন থাকবে?* (What will be the temperature tomorrow in our area?)

We can see that all of these different commands actually mean the same thing and it is requesting for the weather update. Thus, here, our keyword is আবহাওয়া (weather) for the examples a and b, and তাপমাত্রা (temperature) for the examples c, d and e. Based on these keywords, we determine the type of commands such as whether it is a basic command or a core command. We have divided commands into two different phases. These are,

1. Basic Commands
2. Core Commands

In basic commands, we handle commands with basic conversations. Such as asking for a name and storing the name. Having basic knowledge question answering conversation goes to basic commands. In this case, we generate a cosine similarity measure [14][15] between the keywords and commands stored in our knowledge base to determine the command category. This helps us to determine commands category of the same command in different variations. On the other hand, more complex commands such as a command that need actions goes to core commands such as opening an app, setting up an alarm, searching specific information on the Internet etc. Figure 1 demonstrates the system diagram of our proposed system.

On receiving a command, the system determines whether it can respond by itself or we need to call an external API to fetch the information from other sources. In some commands, we do not depend on external APIs such as, all basic commands and commands that the system can handle itself. In this case, our system collects data from our knowledge base and/or store data from the user for future references. On the other hand, in some commands, we need to call an external API to do our work. We also determine based on keywords whether we need to translate the keyword into English since some external APIs do not support Bangla words as input parameters. In such cases, we apply machine translation API to translate the Bangla text to equivalent English text and pass the English text as a parameter of the external API. For machine translation, we have used the Google Translate API [13]. In some cases, Google Translate API provides some garbage data such as '#?' with the result. Thus, we have developed a program based on several regular expressions to filter out these unwanted garbage data.

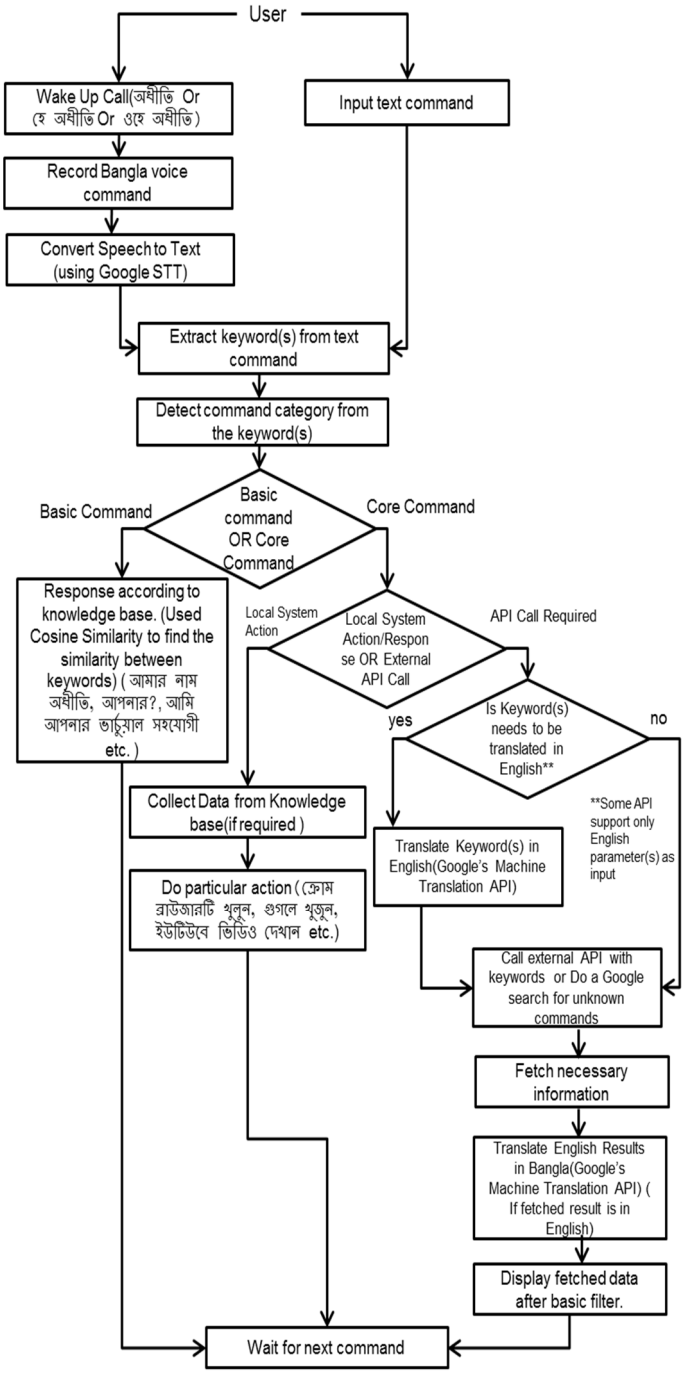


Fig.1 System Diagram of Proposed System

The algorithm of our proposed system is given below.

Algorithm 1: Algorithm of proposed system

```
IF wakeup call OR press the voice button THEN
    RV ← record voice command
    IN_TXT ← convert speech command to text(using Google's STT)
ELSE
    IN_TXT ← read text command
END IF
KEY ← extract keyword(s) by analyzing IN_TXT
CAT ← detect command category using KEY
IF CAT is a basic command THEN
    CS ← measure cosine similarity between CAT and keywords from knowledge base.
    based on CS determine closest command and give response accordingly.
```

```
KB ← Store basic user information in knowledge base for future reference (such as user's name, birthdate, favorite color)
ELSE
    LORAPI← detect whether the command can be handle by our system or we need external API call
    IF LORAPI indicate our system call handle THEN
        collect data from knowledge if required
        do particular action based on command
    ELSE
        ENG← detect whether translation of keyword(s) to English required or not
        IF ENG indicate translation required THEN
            translate keyword(s) to English (using Google's Machine Translation)
        END IF
        call external API with the keyword(s) as parameter
        JSON←fetch necessary information and display accordingly
        IF JSON returns English texts THEN
            convert JSON to Bangla (using Google's Machine Translation)
        END IF
        Extract necessary information from JSON
        Display JSON
    END IF
END IF
```

We have created a dataset of keywords and their corresponding actions. However, if someone put command with a keyword outside of our dataset of keywords, we search the whole command in Google and show the result accordingly. This dataset also let us know whether we need to call an external API to fetch information or our knowledge base is sufficient to respond to the command. In our knowledge base, we have developed an SQLite database and we have several JSON and text files which contain different types of data from different domains and functionalities. The reason behind this distributed knowledge base is to ease of access and to make modification easier. We have an index JSON file that contains commands, keywords of commands, flag that indicates whether it is a basic command or core command, flag that indicates whether we candle this by our system or we need an external API call, flag that indicates whether we need to translate the command keyword(s) in English or not and some other necessary information. This file makes our work easier to decide what to do next. On each external API call, we get JSON formatted data as a result and we keep a local log of the command and fetched data to enrich our system. However, we also store the latest data whenever we find new data for the same command to keep the system up to date.

III. Experimental Results and Analysis

In this section, we have demonstrated the output of different commands in our proposed system and analysis of the results. We have distributed the commands into Basic and Core commands. In total, we have processed 221 commands.

a. Basic Commands

In basic commands, we have processed commands that people use for an initial conversation, to know basic information about our system and some basic mathematics. We have designed the responses as natural as human conversation and our system stores important details for contextual responses. At the very beginning of a conversation with our system, if a user says introductory words such as 'Hi', 'Hello' or 'Hello Adheetee', our system responds mentioning the user's name. Knowing the name of a user is very important for natural friendly conversation. Thus, we address the user as '(বন্ধু) friend' until we know the name of the user. After storing the user's name, we always call the user with his/her name. We have processed 101 basic commands and variations of these commands so that it responds in a natural way just like a human does. Examples of basic commands are:

a. আপনার নাম কি? (What is your name?)

b. আপনাকে কে সৃষ্টি করেছে? (Who created you?)

c. আপনি ছেলে না মেয়ে? (Are you a boy or a girl?)

d. আমার নাম করিম, আপনার নাম কি? (My name is Karim, What is your name?)

e. অধীতী, আমার প্রিয় রঙ লাল এটা মনে রাখুন? (Adheetee, Remember my favourite colour is Blue.)

f. অধীতী, আমার নাম কি বলুনতো? (Adheetee, Tell me my name?)

g. আপনার প্রিয় রঙ কি? (What is your favourite colour?)

h. ২৫ কে ৩ দিয়ে ভাগ করলে কত হয়? (What is 25 divides by 3?)

Figure 2 shows a very initial conversation with a user.

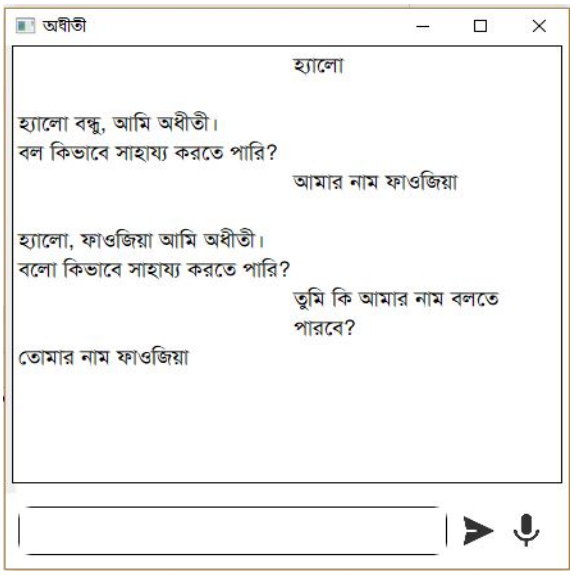


Fig.2 Initial conversation using basic commands

b. Core Commands

In core commands, we have processed commands that need particular actions such as opening a browser, setting an alarm, opening a file, playing a song etc. Core commands are mostly action based with very little conversation. We have processed 120 core commands in our system. Examples of core commands are,

a. অধীতী, দুপুর দুইটার জন্য অ্যালার্ম সেট করুন। (Adheetee, Set an alarm for 2 PM.)

b. অধীতী, ক্রোম ব্রাউজারটি খুলুন। (Adheetee, Open the Chrome browser.)

c. আমার প্রিয় গানটি বাজান। (Play my favourite song.)

d. ইউটিউবে ইংরেজি শেখার ভিডিও চালান। (Play video tutorial of English in Youtube.)
e. মোহাম্মদপুর থেকে মতিঝিল যাব কিভাবে? (How to visit Motijheel from Mohammadpur?)
f. অধীতী, রাশিয়ার রাজধানীর নাম কি? (Adheetee, what is the capital of Russia?)
g. আজ আমার এলাকার তাপমাত্রা কত? (What is the temperature today in my area?)
h. করিম কে ইমেল কর যে আমি আজ অফিসে যাব না। (Send an email to Karim saying I will not go office today.)

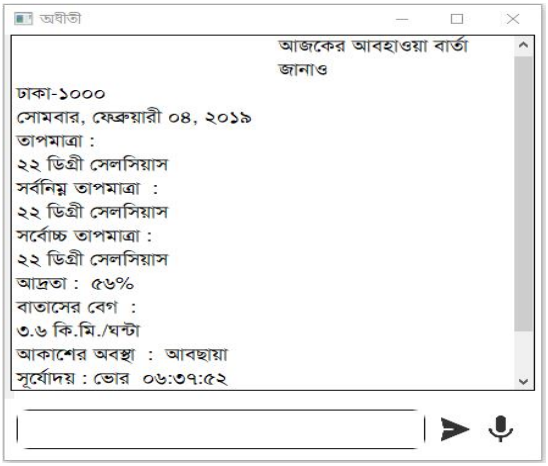


Fig.3 Example of weather command

Figure 3 shows how our system responds for a weather command. We fetch information from internet based on GPS (Global Positioning System) location. As weather information, we fetch different information including temperature, sunset, sunrise, and humidity. Our system shows all this information for a generic weather command. However, if a user asks for specific information the system shows only that specific information.

Sending an email is one of the most important tasks in our daily life. Thus, our system makes it easier to send an email using simple Bangla command. A user needs to say the intended email address or just the name of a recipient with the text.

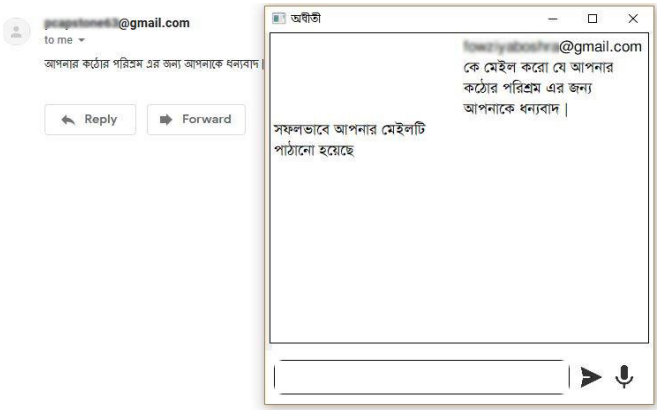


Fig.4 Example of sending email command

In order to send an email by just a name of a recipient, a user must have to give a command anytime earlier so that our system store the email address corresponds to that name. However, if our system does not have an email address for a name it automatically asks for the email address later. Figure 4 shows the response to an email command.

Travelling in a large city such as Dhaka is difficult. In order to go to a new place we often depend on the map. Our system makes it easier to get a map direction from one place to another in the Bangla language. We have used Google Map to display the direction from a source to a destination. For a simple Bangla command such as 'আমাকে মিরপুর ১২ থেকে উত্তরা যাওয়ার রাস্তা দেখাও। (Show me the direction from Mirpur 12 to Uttara.)'. In this case, we have used an NER (Named Entity Recognition) algorithm [4] and gets the name of source and destination. After that, we call the Google Map API using these names in appropriate parameters and finally display the route from the source to the destination. If the user says a command with no source but only destination such as 'আমি বাড্ডা যাব কিভাবে? (How to go Badda?)', we find the source location using the GPS. Figure 5 shows a direction command.

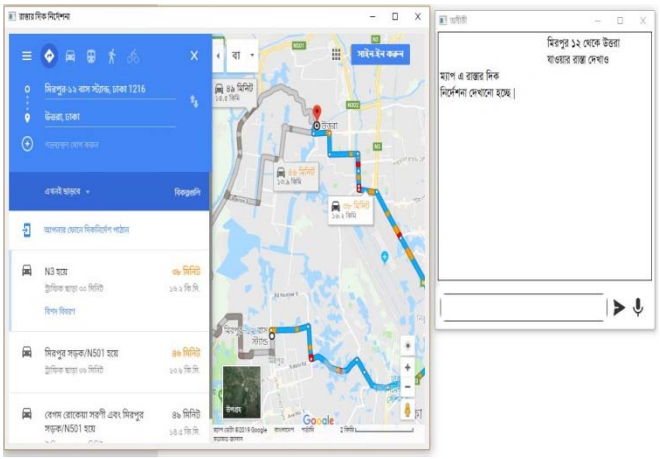


Fig.5 Example of seeking direction in map command

To open a browser, a user needs to mention the name of the browser. Our system opens the browser immediately if the browser is available in the device. Figure 6 shows a command 'অপেরা ব্রাউজারটি ওপেন কর। (Open the Opera browser.)' that open the Opera browser.

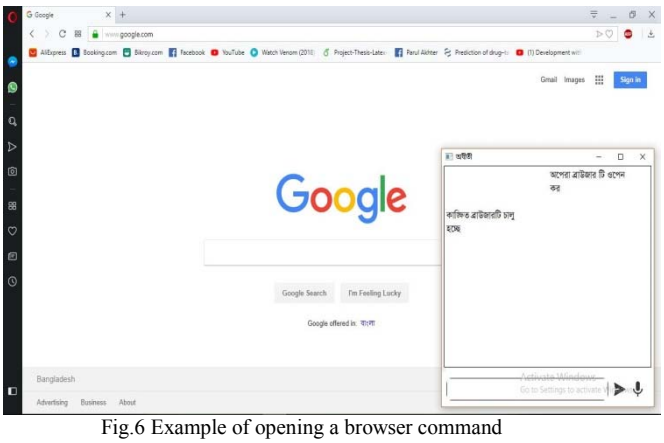


Fig.6 Example of opening a browser command

If a user searches for a video or wants to play a song he can command with appropriate name and source such as Youtube, Amazon Prime, Vimeo etc. If a user asks for a video without mentioning the source of the video, then by default we search the video on Youtube. In Figure 7, we

have shown an example of command seeking for a video on Rubik's cube without mentioning the source.



Fig.7 Example of video command

Our system can fetch any general information from different sources. It can give answers to almost any general knowledge questions. We have an enormous database that contains all the basic general information of all the countries in the word. This contains basic information such as counties name, capital, the biggest city, size, celebrities, poets, the name of the presidents, national sports etc. If a user asks for such basic commands, our system first tries to give answers from its database. However, if there is no data available for specific countries our system fetches the information from the internet. Figure 8 shows some basic general knowledge commands.

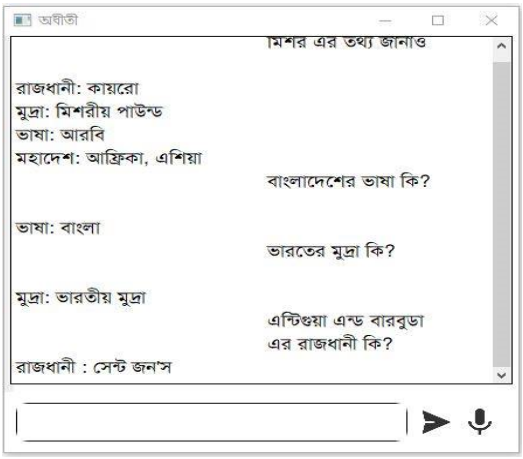


Fig.8 Example of responses to general knowledge commands

A user can search a term in Google using our system with Bangla command. Moreover, if a user asks something that cannot be answered by our system itself, then our system search the command or term in Google directly and show the search result of Google. Figure 9 shows a search result of Google by our system.

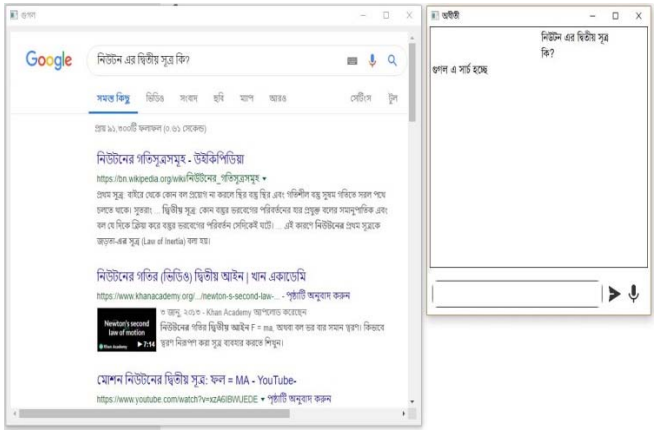


Fig.9 Example of Google search for unknown term by our system

IV. PERFORMANCE ANALYSIS

In order to calculate the performance of our system, we have tested our system in two phases. In the first phase, we have tested our system with some basic commands known and unknown to the system such as, "আপনার নাম কি? (What is your name?)" "আপনি কি কি করতে পারেন? (What can you do?)" etc. In this phase, we have determined our system's ability to have a basic conversation with a user. In the later phase, we have applied our core commands such as "অধীতী, ক্রোম ব্রাউজারটি খুলুন (Adheetee, open the Chrome browser.)", "অধীতী, ধানমন্ডি থেকে বাড্ডা যাওয়ার ম্যাপ দেখান (Adheetee, show the direction in map from Dhanmondi to Badda)" etc. to test our system's ability to do complex things with a device. In this case, we have provided all possible variations of a command a user can ask to our system. We have considered a test success if our system correctly identified our command and correctly provided the expected output. On the other hand, if our system failed to detect our command correctly or it provided incorrect or irrelevant output then we considered the test as a failure. Table I demonstrated the test summary of our system.

TABLE I. PERFORMANCE SUMMARY OF OUR SYSTEM

Phase	**1**	**2**
Name of Phase	Basic Commands	Core Commands
Total Commands Tested (including all variations)	102	112
No. of Correctly Identified and Responded Commands	99	102
No. of Incorrectly Identified or Responded Commands	3	10
Error Rate	2.94%	8.93%
Accuracy Rate	97.06%	91.07%

TABLE II. AVERAGE ERROR AND ACCURACY RATE

Average Error Rate	$\frac{2.94\% + 8.93\%}{2}$	= 5.935%
Average Error Rate	$\frac{97.06\% + 91.07\%}{2}$	= 94.065%

In Table I, we have demonstrated the test results of both phases and Table II demonstrates the average error rate and accuracy rate of the system. The average accuracy rate of the system is 94.065% with the error rate of 5.935%. Since our proposed virtual assistant is the first Bangla virtual assistant

and there is no Bangla virtual assistant available to compare with, we have compared our system with existing English virtual assistant. Figure 10 shows a comparison of the accuracy rate between English virtual assistants and our system 'Adheetee'. The accuracy rate of these English virtual assistants was reported in an article of Business Insider [16].

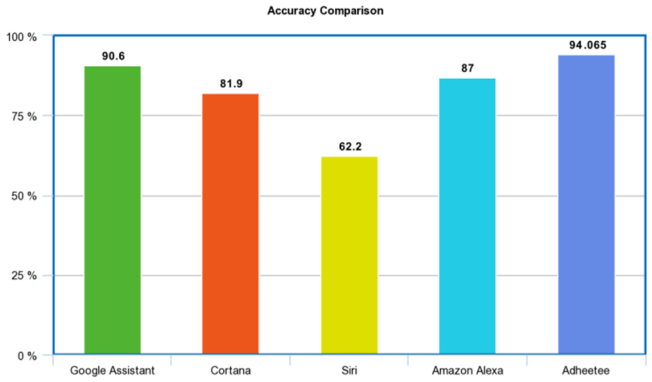


Fig.10 Accuracy comparison between English virtual assistants and Adheetee.

V. Conclusion and Future Work

This paper demonstrates a comprehensive Bangla virtual assistant. We have demonstrated each and every step of building the system. Moreover, we have shown different experimental results and analysis of the results and finally, we have demonstrated the test results. We find our system very impressive according to our test results.

Although we have reached our goal, the system has some limitations. Such as, our system mostly depends on external APIs to fetch information which makes the system dependent on other systems. Since the system depends on the internet to fetch most of the information internet speed may affect the performance of the system. Finally, we have covered a limited number of commands at this stage which makes the system less versatile.

In future, we are planning to do a user experience survey according to HCI discipline based on a robust online and offline questionnaire. We are developing some APIs to reduce the dependency on external APIs. Moreover, we are developing a mobile app for both Android and IOS platforms. Handling more frequent commands is our most important future task. We are planning to test the system with more complex commands and hence increase the accuracy rate of the system.

References

[1] Hoy, Matthew B. (2018). "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants". Medical Reference Services Quarterly. 37 (1): 81–88. doi:10.1080/02763869.2018.1404391

[2] Rudolph, Stacey. The Growth in Usage of Virtual Digital Assistants [Infographic]. Retrieved from https://www.business2community.com/infographics/the-growth-in-usage-of-virtual-digital-assistants-infographic-02056086

[3] A. Bodart and C. Condon, "Virtual assistant," Feb. 22 2007, uS Patent App.11/206,829.

[4] Chowdhury, Shammur & Alam, Firoj & Khan, Naira. (2018). Towards Bangla Named Entity Recognition. 10.1109/ICCITECHN.2018.8631931.

[5] Varma, S. (2018, June 14). Siri vs Google Assistant vs Alexa: Which Is the Smartest Virtual Assistant in 2018? Retrieved from https://gadgets.ndtv.com/apps/features/siri-vs-google-assistant-vs-alexa-which-is-the-smartest-virtual-assistant-in-2018-1867292

[6] Hays, Jeffrey. BENGALIS. Retrieved from http://factsanddetails.com/india/Minorities_Castes_and_Regions_in_India/sub7_4b/entry-4198.html

[7] Pinon, Robert; Haydon, Jon (December 2010), The Benefits of the English Language for Individuals and Societies: Quantitative Indicators from Cameroon, Nigeria, Rwanda, Bangladesh and Pakistan, Euromonitor International Ltd

[8] "Bangladesh - Literacy rate 2017." [online] countryeconomy.com. Available at: https://countryeconomy.com/demography/literacy-rate/bangladesh

[9] Paul, Anirudha & Latif, Asiful & Amin Adnan, Foysal & Rahman, Mohammad. (2018). Focused domain contextual AI chatbot framework for resource poor languages. Journal of Information and Telecommunication. 1-22. 10.1080/24751839.2018.1558378.

[10] Orin , T. D. (2017). Implementation of a Bangla Chatbot (Unpublished Bachelor Thesis). BRAC University, Dhaka, Bangladesh.

[11] "Top Countries/Markets by Smartphone Penetration & Users. 2018." [online] newzoo.com. Available at: https://newzoo.com/insights/rankings/top-50-countries-by-smartphone-penetration-and-users/

[12] "Cloud Speech-to-Text - Speech Recognition | Cloud Speech-to-Text API | Google Cloud." Google, Google, cloud.google.com/speech-to-text/.

[13] "Cloud Translation API Documentation | Cloud Translation API | Google Cloud." Google, Google, cloud.google.com/translate/docs/.

[14] A. Dhar, N. Dash and K. Roy, "Classification of text documents through distance measurement: An experiment with multi-domain Bangla text documents," *2017 3rd International Conference on Advances in Computing,Communication & Automation (ICACCA) (Fall)*, Dehradun, 2017, pp. 1-6. doi: 10.1109/ICACCAF.2017.8344721

[15] The Cosine Similarity algorithm. [online] www.neo4j.com. Available at: https://neo4j.com/docs/graph-algorithms/current/algorithms/similarity -cosine/

[16] Dunn, Jeff. "It Looks like Apple Has Some Work to Do If It Wants Siri to Be as Smart as Google Assistant." Business Insider, Business Insider, 2 June 2017, www.businessinsider.com/siri-vs-google-assistant-cortana-alexa-knowledge-study-chart-2017-6.

A Bangla Spell Checking Technique to Facilitate Error Correction in Text Entry Environment

Tanni Mittra[1], Sadia Nowrin [2], Linta Islam[3] and Deepak Chandra Roy[4]
Dept. of CSE, East West University, Dhaka, Bangladesh
([1]tanni@ewubd.edu)

***Abstract*—Spell checker is a common tool used in different text entry techniques for error free writing. Recent advances in spell checking techniques in different languages significantly decreases word and sentence level typos from writing text. But developing an optimized spell checker in Bangla language is still a great research challenge due to the complex rules of Bangla spelling. To overcome the challenge we propose a new technique of spell checking considering both word and sentence level errors to facilitate the text correction techniques of Bangla language. We use a hybrid approach by combining edit distance algorithm and probability based N-gram language model for the error detection and correction tasks. A large corpus is built containing almost 0.5 million words and 50,000 n-gram sentences to detect and correct the errors. We evaluate the performance of our proposed method with different test data collected from online sources. Again, the efficiency of proposed technique is estimated by conducting a pilot study involving users of different ages. The accuracy rate of our proposed spell checker is approximately 97%. The validity of our claims have been proved through comparison with other Bangla spell checkers.**

Keywords*—Bangla spell checker, Real-word error, Natural Language Processing, N-gram

I. Introduction

Spell checker is an optimized application which determines misspelled words and provides most probable suggestions based on the wrong words. Usually spell checker is integrated with the text entry techniques of different word processing software like Microsoft Word, Writer, Ispell, Aspell4, LibreOffice, etc. [1]. It helps users to write error-free textual documents by detecting spelling errors.

Spelling errors are classified into two broad categories [2]. One is non-word error, which occurs when an invalid word is present in the text and not found in the word dictionary. Another is the real-word error, which occurs if the word is valid but not syntactically and semantically correct [1], [3]. Any spell checking technique is evaluated by its capability of handling both non-word and real-word errors. As a result, detecting both types of error and automatically correcting them is still a major research challenge. Plenty of research works have been done on automating the process of spell checking in English and non-English language like Bangla, Hindi, French, Spanish etc.

The evaluation in recent years indicates that, spell checking system of Bangla language has made a measurable improvement [3]. Considerable works have been done on Bangla spell checker. For example, phonetically similar word grouping based spell-checker [4], phonetic encoding for better spelling suggestions [5], double metaphone encoding technique [6], a light weight stemmer for bengali spell checker [7] are some of the automated processes for correcting misspelled words in Bangla. However, there exists some more challenges to improve the performance of Bangla spell checker [3]. Among them, handling both types of error together with great efficiency is a big challenge. Because for any spell checker, increasing the efficiency of error detection as well as improving its quality has great importance.

According to our best knowledge, some of the spell checkers concentrate on structural similarity of words while other focus on syntactic similarity of words in a sentence. A very few of them focus on both structural and syntactic similarity. On the other hand, current spell checkers are not optimized for efficiency handling and memory consumption. So in overall, the main challenge is to emerge an optimized solution that can detect errors by handling all complexities and provide justifiable suggestions based on the wrong words. To overcome these challenges, we present a new smart Bangla spell checking technique with two features. Firstly, it can handle both types of error by considering the complex rules of Bangla spelling and provides justifiable suggestions. Secondly, it improves the performance of spell checking by reducing search time and space. The method described in this paper constructs a group based dictionary, where each of the words having equal length is grouped together. Thus the issue of time and space complexity is minimized. A simple edit distance algorithm is used to check the structural similarity of words. The syntactic similarity is calculated using probability based N-gram model. In the result analysis section, we compare our method with a well established spell checker to analyze the performance of our method.

The rest of the paper is organized as follows. Section II covers background study and section III discusses about some related research works. Section IV contains brief descriptions of the proposed method, section V describes the evaluation procedure and results attained by different experiments. Finally, Section VI discusses the future plan with conclusions.

II. Background Study

Bengali also known as Bangla is an Indo-Aryan language spoken in South Asia. It is the official language of Bangladesh and second most widely spoken language of India. Bangla is the seventh most used native language in the world and fifth most widely used writing system in the world[1]. It has 49 letters in it's alphabet and 10 digits in decimal number system [8]. Among the set of letters, it has some complex set

[1]https:/en.wikipedia.org/wiki/Bengali_language

of characters which have made it's writing very difficult rather than the verbal communication [3]. They are: 1) Phonetically similar characters example:ন,ণ,শ and স 2) Consonant cluster example:ক্ষ ,ল্ম and ন্ম 3) Phalas such as YA-phala, RA-phala and LA-phala example: ব্য,ক্ষ 4) Matra example: পর্ব ,কর্ম 5) vowel characters example: ◌া, ◌ো, ◌ী .

The complex rules of using above-mentioned characters in writing script have emerged different types of the spelling error. According to the classification of spelling error Kukich [2] classified non-word error as word level error and real-word error as sentence level error. In context of Bengali writing, most of the non-word error occurs because of less knowledge of the spelling rules [9] and due to inaccurate insertion, deletion of characters while typing. Example: চুল, তালী, উপককরণ,মানুস all are the example of non-word errors because they are not the valid words. Contrarily most of the sentence level error occurs due to the inappropriate use of phonetic characters, Phalas and Matras. Example: if someone write a sentence আমার মণ খারাপ where all of the three words are valid dictionary word but the word মণ is inappropriate in the context of the sentence. The correct sentence would be আমার মন খারাপ . Another example এখানে অন্ন কেউ , where the word অন্ন is invalid in this sentence.

III. Related Works

There exists some well established algorithms for spell checking such as double metaphone encoding [10], metaphone encoding [11],soundex algorithm [12] and phonix algorithm [13]etc. They used phonetic codes for detecting misspelled words and generating suggestions. But, all of them are used for developing spell checker in the English language. To our best knowledge, a small number of researchers worked on these concepts to implement spell checker in the Bangla language [3]. For example, UzZaman et al. provided Bangla phonetic encoding based on soundex algorithm [5]. They modified existing soundex algorithm for Bangla and proposed a new encoding technique. It considered only phonetic problems and used the phonetic encoding to provide suggestions. A double metaphone technique for Bangla [6] is used to detect spelling errors and improve suggestions. They considered the phonetic problem and used traditional edit distance based method to produce suggestions.

A light weight stemmer algorithm [7] is used to detect misspelled word. By using this stemming algorithm they reduced dictionary size and improved efficiency of spell checking. They used simple minimum edit distance method to find suggestions. UzZaman et al. used a direct dictionary look up method for determining errors [14]. They handled both phonetic and real-word error by considering edit distance and double metaphone encoding.

A reversed dictionary [4] was built by reversing the characters of each word and also an original dictionary where phonetic characters were grouped together. A simple string matching algorithm is used to detect the phonetic error. Both dictionaries were used to provide suggestions. It considered the phonetic problem and can correct a single error.

Finite state automation [15] is used to determine the spelling error and probability is used to generate suggestions. Khan et al. [16] used n-gram model for determining the error in Bangla text. To check the error, the probability was calculated using the N-gram model. Statistical n-gram model is used to recognize real-word error in the Bangla sentence in [17]. On the other hand, clustering approach [18] is used to oraganize the dictionary. This organization reduces the search time and space of the algorithm. But this algorithm worked only for non-word error.

Although there exists some well-established methods for Bangla spell checker, most of them focused on the non-word error. A very little amount of work focused on the real-word error. These error detection and suggestion generation operations require more efficiency. As a result, we focus on the design of an optimized spell checking solution that can handle quality with proper efficiency of the error detecting task.

IV. Optimized Bangla Spell Chekcing Technique for Error Correction

The spell checking technique of our proposed method detects all types of erroneous words and provides most probable corrections based on the wrong input. The main goal of our work is to detect misspelled words efficiently by minimizing the search time and space. A relatively straight-forward algorithm is used to determine the correctness of the word and generate most likely candidate words as suggestions. In the following subsections, we discuss the implementation details of our proposed method.

A. Determining Erroneous Word

Determining erroneous word is the first step of any spell checker. The error detection process of our proposed method is subdivided into two basic steps. They are: 1) non-word and 2) real-word error detection.

1) Non-word Error Detection: When a user enters a word which does not exist in the dictionary is known as non-word error [2]. To check non-word error, we build a Bangla word dictionary with approximately 6,00,000 words. All words are collected from different online repository of Bangla words[2], newspapers, social networking sites and Bangla blogs. All words of the dictionary are organized according to the length of the words i.e. all words having equal length are grouped together. Then each of the words is subdivided into a set of trigrams which are also stored in the dictionary. For example: trigram set of the word মানুষ is <ম ◌া ন >, <◌া ন ◌ু> , <ন ◌ু ষ>. The main reason behind this organization is to minimize the searching time. So to determine the non-word error at first we determine the length of the inputted word and divide the word into trigrams. Then the word is searched on the group having the same length in the dictionary. Comparison starts with the trigrams of the trigram set. If the trigrams are not matched with any other words of the dictionary, then the word is treated as non-word error. For example: if a user types সে মানুস হিসেবে ভালো in his typing episode, our system takes each word sequentially and verifies with the dictionary. From the above example our system determines মানুস as incorrect word.

[2]http://www.english-Bangla.com/bntobn



(a) 1a (b) 1b

Fig. 1: Non-word and real-word error detection

In some cases, words may contain suffixes. For example, the word করেছিলাম contains the suffix িলাম.For this type of word, our system finds no exact matching in the dictionary. In order to reduce this type of problems, we use the concept of the stemming algorithm [7] that only strips the root word and the suffix. We consider some frequent suffixes for stemming: ই,ব, িলাম,ি, ী, টি ,সব etc. and order the suffixes in descending order according to the length of the suffix. We match the inputted word from the right with the suffix list. If any matches found, then we remove the suffix from the actual word. Then we use the same procedure as discussed above to determine the non-word error.

If any word is detected as an incorrect word, then our system generates suggestions based on the wrong word. For this, we match the first trigram of the entered word with the first trigrams of other words from the dictionary. If any matches found, we select them as candidate words. Then we calculate the length difference between the inputted and the candidate words. If the length difference is between zero to five, then the words are selected as suggestions. Continuing the previous example, suggestions generated for the word মানুস are মানুষ , মানুষকে , মানুষের which is illustrated in Figure 1a.

2) real-word Error Detection: The next step is to check whether the word is syntactically correct or not. For example, in text "`his presence ink my room`", the word "`ink`" is a real-word error. In Bangla, as discussed above this type of error generally occurs due to substitution of a phonetically equivalent letter. Based on Hoque and Kaykobads Soundex type encoding [19] and Zaman and Khans Soundex type encoding [5] we have created some phonetic groups containing similar phonetic characters, which are shown in Table I

TABLE I: Phonetically Similar Characters Grouping

Group Number	Members
1	ন, ণ
2	স, শ, ষ
3	র, ড়, ঢ়, ঋ
4	খ, ক্ষ
5	ে , ে,ঊ , উ
6	ী , ি,ই ,ঈ
7	জ, য
8	ব, ভ

If a word contains any of the above mentioned letters, then our method searches for the candidate words containing other phonetic words of the same group having Levenshtein distance [20]equal to 1 from the word dictionary. Then it calculates the probability of the candidate set along with the inputted word. Based on that, it detects real-word error and suggests some relevant words from the candidate set simultaneously.

a) Candidate Set Construction: The candidate set of a word W_k is a set of words which needs to perform single edit operation to generate W_k. An edit operation is either insertion, deletion or substitution. The candidate set C_{W_k} is represented as:

$$C_{W_k} : \{W_1, W_2, W_3,...,W_i\}$$

Here k is the position number of the word in the typed sentence and i is the word number of the candidate set. For example: if a user types আমার মণ খারাপ then our system finds that মণ contains phonetic letter from above mentioned group and generates মন, বন,সন,শন as its candidate set. For better understanding, we have used the above mentioned example in the following subsections.

b) N-Gram Model Generation of Candidate Set: After generating the candidate set the next step is to form bigram (N=2) and trigram(N=3) for each member of the candidate set C_{W_k}. For bigram, we generate both left and right bigram by considering left word and right word and for trigram we use both of them. Following model (taken from [1])is used to generate bigram and trigram model for each member of the candidate set.

$$\begin{aligned}\text{Left Bigram} &: \{W_{k-1},W_k\}\\ \text{Right Bigram} &: \{W_k,W_{k+1}\}\\ \text{Trigram} &: \{W_{k-1},W_k,W_{k+1}\}\end{aligned}$$

Table II shows bigram and trigram models of each candidate word along with the original word of the above mentioned example.

TABLE II: N-gram Models of Candidate Set

Candidate words	right bigram	left bigram	trigram
মণ	আমার মণ	মণ খারাপ	আমার মণ খারাপ
মন	আমার মন	মন খারাপ	আমার মন খারাপ
বন	আমার বন	বন খারাপ	আমার বন খারাপ
সন	আমার সন	সন খারাপ	আমার সন খারাপ

The occurrence of these bigram and trigram models is calculated from an experimental Bangla corpus containing 50000 sample text collected from different newspapers, blogs and from an informal survey.

c) Probability Calculation of N-Gram Model: The bigram and trigram probabilities of each candidate word is calculated using the following maximum likelihood (ML) estimation equation.

$$\text{P1}(W_k|W_{k-1}) = \frac{\text{Count}(W_{k-1}W_k)}{\sum_0^i \text{Count}(W_{k-1}W_i)} \quad (1)$$

$$\text{P2}(W_k|W_{k+1}) = \frac{\text{Count}(W_kW_{k+1})}{\sum_0^i \text{Count}(W_iW_{k+1})} \quad (2)$$

$$P3(W_k|W_{k-1},W_{k+1}) = \frac{\text{Count}(W_{k-1}W_kW_{k+1})}{\sum_0^i \text{Count}(W_{k-1}W_iW_{k+1})} \quad (3)$$

Here, equation 1 estimates left bigram probability by counting total number of left bigram W_{k-1},W_k. The denominator is the summation of the total count of all bigrams, containing previous word W_{k-1} of the typed sentence and one word W_i from the candidate set. In the same way, both P2 and P3 estimates right bigram and trigram probability of the candidate set. These three probabilities are combined to count a weight of each candidate word using the principle of Linear Interpolation.

$$\begin{aligned} W(W_k) = \lambda_1 P1(W_k|W_{k-1}) + \lambda_2 P2(W_k|W_{k+1}) \\ +\lambda_3 P3(W_k|W_{k-1},W_{k+1}) \end{aligned} \quad (4)$$

Here, λ_1+λ_2+λ_3= 1. To estimate the value of λ a small training set is built containing 20% of the experimental corpus. We noticed that, the results are best if we estimate λ_1 = λ_2 = 0. 3 and λ_3= 0.4. As trigram models are low biased and have a robust count on given sufficient training data [1], so we maximize λ_3.

d) Error Detection and Suggestion Generation : The above equation is used to calculate the weight of each candidate set. After generating the weight we sort the weight in descending order. If the weight of the actual word that was typed is zero then it is considered as real-word error. In some cases,the weight of each candidate set is nonzero, in this regard we check if the weight of the actual word is greater than other words of the candidate set then it is also considered as real-word error. Continuing the previous example, we have the candidate set মন, বন,সন,শন. For each of the candidate set we calculate the weight as shown in Table III. There is no count for bigram আমার মণ and মণ খারাপ as well as trigramআমার মণ খারাপ of actual word . So the total weight for the actual inputted word মণ is zero. So it can be considered as real-word error. Figure 1b illustrates the process.

TABLE III: Candidate Set With Weight

Candidate Word	Weight
মন	3.75
সন	0
মণ	0
বন	0

Let us consider another example আমার আসা নেই. Here ,the word আসা is a phonetic word and generates আসা , আশা ,বাসা as candidate set. In this case, the weight of each candidate word is nonzero, so our system rearranges the weight in descending order as shown in Table IV. However, the total weight for the actual inputted word আসা is less than the weight count of word আশা. So it can be considered as real-word error.

Again weight zero may not always mean real-word error. If a particular bigram or trigram is absent in the corpus then the weight may be zero. This type of problem arises if the inputted phonetic word contains suffixes. For example, if the given text is তোমার ফোনের আশাই করেছিলাম, our system finds no bigram and trigram count generated by the word আশাই. In order to reduce this type of problem we use the same concept of stemming described in section IV-A1. After removing the suffix from the actual word we use the same procedure as discussed above to determine the real-word error.

TABLE IV: Candidate Set with weight

Candidate Word	Weight
আশা	0.875
আসা	0.5
বাসা	0.25

If an inputted phonetic word is considered as real-word error then we select the ordered candidate set and provide them as suggestions. Because we rearrange the candidate set according to the weight in descending order. If the weight of a candidate word is greater than a threshold value 0.5 then the word is treated as a suggestion. We choose this value as threshold value because we perform different experiments on the weight of different candidate words and notify that most of the candidate words having a weight greater than 0.5 are considered as a good candidate for suggestions.

V. Experiments and Evaluation

We developed an android application to implement our proposed method. To develop the application we used android studio(Version 3.3.2), an official IDE for android application development[3] which uses Java language to write codes. We tested our application with 50,000 words and 3000 sentences collected from different sources such as Bangla newspapers, social networking sites and Bangla blogs. Again, we tested our application by ten users who used our application for their text writing purpose. For a better evaluation, we compared our spell checking technique with another familiar spell checking method [16] which used N-gram model to check spelling errors. The following subsections contain the discussion of results attained from different experiments in terms of error detection accuracy and suggestions generation efficiency.

A. Evaluation Dataset

As there is no specific dataset available for Bangla language, so in this experiment we created our own dataset. As discussed above we collected 50,000 structurally correct words from different sources to determine non-word error detection accuracy of our application. Among them, 27000 words were collected from two prominent newspapers i.e. Prothomalo[4] and Kalerkontho [5], 10200 words from social networking sites(i.e. facebook) and 5400 words from Bangla blogs[6]. Rest of the 7400 words were tested by the application users. Again, among the 3000 correct sentences as stated above almost 1000 texts were collected from online sources and other 1600 were from application users.

[3]https://developer.android.com/studio/
[4]https://www.prothomalo.com
[5]https://www.kalerkantho.com/
[6]http://www.sachalayatan.com/

B. Experimental Methodology and Result

We conducted two types of experiments in order to figure out, (i) the correctness of error detection and (ii) correctness of suggestions . The first experiment was done to determine error detection quality considering both non-word and real-word error. Because error detection quality reflects the effectiveness of the detection algorithm and a good detection algorithm increases the possibility of correct detection. At first, we tested our application with 50000 correct words to determine non-word error detection quality. Then we manually modified these words by performing single edit operation i.e. insertion or deletion operation to generate 50000 incorrect word. The result generated for both correct and incorrect words are shown in Table V, which shows that the average error correction rate of our proposed method is 96.32% for non-word error.

TABLE V: Result Generated for Non-word Error

Type	Total Words	Correct Detection	Efficiency
Correct	50000	47640	95.28%
Incorrect	50000	48681	97.36%
Average			96.32%

Again to determine the real-word error, at first we tested 1000 correct sentences containing approximately 1990 words and 50 phonetic words as stated above. Among them, 600 sentences were from newspapers, 300 from blogs and 100 were from social networks. The result generated by the experiment with correct sentences is shown in Table VI.

TABLE VI: Result Generated from Correct Sentences

Source	Total Sentences	Correct Detection	Efficiency
Newspaper	600	578	96.83%
Blogs	300	285	95%
Social Networks	100	97	97%
Average			96.27%

As all the sentences contain correct words, so for another experiment we prepared an erroneous test file containing corrupted words to test our system. To make such an erroneous file, we replaced the correct phonetic word with erroneous word by performing single edit operation including insertion, deletion and substitution at phonetic letter position. In this way, we generated 3000 erroneous test cases of real-word error. Then we computed the actual detection of real-word error R and the erroneous detection rate E of the proposed method. Let consider e_1 represents the total number of error, e_2 represents the total number of error detection and e_3 represents total number of correct real-word error detection. Where,

$$R = \frac{e_3}{e_1}$$

$$E = \frac{e_2 - e_3}{e_1} * 100\%$$

Table VII represents the percentages of incorrect identification of real-word error and erroneous detection rate of our proposed system.

TABLE VII: Efficiency Calculation of Erroneous Sentence Detection

Correct Detection	Erroneous Detection Rate
94.37%	3%

As described above, for a better evaluation we also evaluated our application by conducting a pilot study involving 10 participants aged from 18 to 40. Participants used sentences containing different length of words. According to the length, we classified these sentences as simple, complex and compound sentences. For this experiments, participants used almost 1600 number of sentences of different length. Table VIII indicates that the accuracy of our proposed method is on an average of 96.68%. Therefore, the average real-word correction rate of our developed application is 96.48%.

TABLE VIII: Performance Evaluation from User Study

Type	Word Length	No. of Sentences	Correct Detection	Accuracy
Simple	1-4	680	678	99.70%
Complex	5-7	532	512	96.24%
Compound	8-10	356	335	94.10%
Average				96.68%

Our second experiment was for determining correct suggestion generation rate. If any word selected as non-word error, then our system can generate suggestions for approximately 95% misspelled word. As discussed in IV-A1, suggestions are generated if the entered and the candidate word from the dictionary having a minimal length difference. Depending on the procedure, we divided the generated suggestions into two categories. Category 1 contains words having length difference between 0 to 3 and Category 2 contains words having a difference from 4 to 5. Table IX represents the result, which shows that 90% suggestions are of a length distance between 0 to 4.

TABLE IX: Results Based on Suggestions for Non-word Error

Length Distance	Selected as Suggestions(%)
0-3	90%
4-5	5%

Again for real-word error, our system can generate approximately 90% correct suggestions treated as top-ranked suggestion. If a word is treated as real-word error, then the candidate suggestions are generated based on the sorted score of each candidate set. Based on the sorted score we divided suggestions into relative ranks. The relative ranks of the suggestions are shown in figure 2. It is observed that almost in 85% cases Rank1 suggestions were correct and 8% correct suggestions were from Rank2. This results show that our proposed method generates correct suggestions for almost 93% of the cases with a failure rate of 3%.

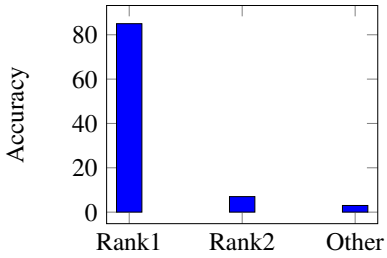


Fig. 2: Evaluation of ranked suggestions

Again we compared our proposed method with another well-known spell checking algorithm [16]. We selected this paper for comparison because it used the N-gram model to detect error correction. Table X represents the result obtained from above-mentioned comparison. The result showed that the error detection correctness is 96.65% while for another method the correctness is 94%.

TABLE X: Result Obtained from Comparison with Other Method

	Proposed Method	**Compared Method**
No. of words	2540	2540
Correct Detection	2455	2389
Accuracy	96.65%	94.05%

VI. Conclusions and Future Work

In this paper, we presented a new spell checker to facilitate the error correction technique of Bangla text. The proposed method can effectively detect word and sentence level errors and provide suggestions based on the errors. We used the concept of edit distance algorithm in our own way to detect the word-level error. For the sentence level error, the probability of the N-gram models are calculated using the principle of linear interpolation model. We evaluate our proposed spell checking technique with test words and sentences collected from different online sources and by conducting a pilot study. The evaluation results showed that this technique significantly detects both errors and provides relevant corrections based on the error. Results also revealed that our proposed method works better than other spell checking technique of Bangla language.

The future work may involve creating a spell checker that can handle multiple non-word and real-word errors in sentences. In the future, we will develop a complete spell checking application for Microsoft word. Moreover, our future target is to apply the spell checker for patient complaint system in the medical sector of Bangladesh.

References

[1] P. Samanta and B. B. Chaudhuri, "A simple real-word error detection and correction using local word bigram and trigram," in *Proceedings of the 25th Conference on Computational Linguistics and Speech Processing (ROCLING 2013)*, 2013, pp. 211–220.

[2] K. Kukich, "Techniques for automatically correcting words in text," *Acm Computing Surveys (CSUR)*, vol. 24, no. 4, pp. 377–439, 1992.

[3] P. Mandal and B. M. Hossain, "A systematic literature review on spell checkers for bangla language," *International Journal of Modern Education and Computer Science*, vol. 9, no. 6, p. 40, 2017.

[4] B. B. Chaudhuri, "Reversed word dictionary and phonetically similar word grouping based spell-checker to bangla text," in *Proc. LESAL Workshop, Mumbai*, 2001.

[5] N. UzZaman and M. Khan, "A bangla phonetic encoding for better spelling suggesions," BRAC University, Tech. Rep., 2004.

[6] N. UzZaman and M. Khan, "A double metaphone encoding for bangla and its application in spelling checker."

[7] M. Islam, M. Uddin, M. Khan *et al.*, "A light weight stemmer for bengali and its use in spelling checker," 2007.

[8] M. Aktaruzzaman and M. F. Khan, "A new technique for segmentation of handwritten numerical strings of bangla," 2013.

[9] P. Kundu and B. Chaudhuri, "Error pattern in bangla text," *IJDL. International journal of Dravidian linguistics*, vol. 28, no. 2, pp. 49–88, 1999.

[10] Philips, Lawrence, "The double metaphone search algorithm," *C/C++ users journal*, vol. 18, no. 6, pp. 38–43, 2000.

[11] L. Philips, "Hanging on the metaphone," *Computer Language*, vol. 7, no. 12 (December), 1990.

[12] J. L. Bentley, *Writing efficient programs*. Prentice-Hall, Inc., 1982.

[13] T. Gadd, "Phonix: The algorithm," *Program*, vol. 24, no. 4, pp. 363–366, 1990.

[14] N. UzZaman and M. Khan, "A comprehensive bangla spelling checker," 2006.

[15] M. M. Abdullah, M. Z. Islam, and M. Khan, "Error-tolerant finite-state recognizer and string pattern similarity based spelling-checker for bangla," in *Proceeding of 5th International Conference on Natural Language Processing (ICON)*, 2007.

[16] N. H. Khan, G. C. Saha, B. Sarker, and M. H. Rahman, "Checking the correctness of bangla words using n-gram," *International Journal of Computer Application*, vol. 89, no. 11, 2014.

[17] M. Nur Hossain Khan, F. Khan, M. M. Islam, M. H. Rahman, and B. Sarke, "Verification of bangla sentence structure using n-gram," *Global Journal of Computer Science and Technology*, 2014.

[18] P. Mandal and B. M. Hossain, "Clustering-based bangla spell checker," in *Imaging, Vision & Pattern Recognition (icIVPR), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–6.

[19] M. T. Hoque and M. Kaykobad, "Coding system for bangla spell checker," in *5th International Conference on Computer and Information Technology*, 2002, pp. 186–190.

[20] V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions, and reversals," in *Soviet physics doklady*, vol. 10, no. 8, 1966, pp. 707–710.

Question Bank Similarity Searching System (QB3S) Using NLP and Information Retrieval Technique

*

Md. Raihan Mia
Dept. of Computer Science and Engineering (CSE)
Bangladesh University of Engineering and Technology(BUET)
Dhaka, Bangladesh
1305116.mrm@ugrad.cse.buet.ac.bd

Abu Sayed Md. Latiful Hoque
Dept. of Computer Science and Engineering (CSE)
Bangladesh University of Engineering and Technology(BUET)
Dhaka, Bangladesh
asmlatifulhoque@cse.buet.ac.bd

***Abstract*—Problem Based e-learning(PBeL) in bangla language is one of the most progressing areas of the use of ICT in education. Question Bank(QB) is the main component of any PBeL system. Searching similarity in the complex structure of QB is a challenging task in the development of PBeL system. We have been developed an efficient Question Bank Similarity Searching System(QB3S) to find similar questions, handle duplicate question and rank search result of a query input based on NLP and Information Retrieval techniques. QB3S has four modules: bangla documents processing, question structure analysis and clustered indexing by B+ tree , word-net construction and Information retrieval module. Lexical analysis, stemming by finite automata rules and stopwords removing have been used for bangla document processing. The most challenging procedures of QB3S were Analyzing the structure of data for clustered indexing in the sorted sequential file of the QB database with a B+ tree data structure and improved TF-IDF algorithm with weighted functionality. A Word-net has been used for handling synonyms. Vector Space Model(VSM) has been designed from the value of TF-IDF weighted matrix. By using cosine similarity product rule, we have been Calculated the similarity value between the query input and all mcq of DB from VSM. QB3S has been evaluated in some experimental dataset to find results by imposing different test cases. The accuracy of searching performance which has found to be satisfactory.**

***Index Terms*—Tokenization, Stopwords Removing, Stemming, Clustered Indexing, B+ tree, Word-Net, TF-IDF, VSM, Cosine Similarity**

I. Introduction

Problem-Based eLearning(PBeL) systems have been proved to be effective in blended learning in classroom and also in asynchronous learning. The PBeL system [1] developed for ICT course in Higher Secondary School(HSC) level contains a rich question bank of different types and complexities. It is utmost requirement of the QB to contain only distinct questions in terms of title, contents and answers. This requires a rigorous similarity search to remove the possible duplicate questions. In this research, we have developed a similarity searching system on complex structures of QB using NLP and IR technique. Beside duplicate handling, it also can be ranked searching results of a query input based on similarity values.

Information Retrieval (IR) [2] system can be construed as a set of techniques and tools dealing with access to information from a set of structured, semi-structured or unstructured documents database in order to allow the user to retrieve the correct and relevant information. IR algorithms for fetching relevant information from large amount of structured data of QB database where data is represented in tabular form, need to strategically better searching techniques.

We have used NLP tools [3] like, Tokenization or lexical analysis of bangla text corpus [4], StopWord removing, Rule-based sementic analysis or stemming to modify inflected word and analyzed the structured data to build a dynamic B+ tree clustered indexing to speed up the retrieval of records. To build the generative laexicon [5] for bangla language, many paper had been already published in the field of computational linguistics morphology. The stem need not be identical to the morphological root of the word; it is usually sufficient that related words map to the same stem, even if this stem is not in itself a valid root. Algorithms for stemming have been studied in computer science since the 1960s. In [6], A rule-based approach of stemming bangla word had been found an excellence solution. Sometimes Bangla words get more complex morphological structure in addition of inflectional suffix, derivational suffix and agglutination . There are existing efforts at building a complete morphological parser for Bangla [7], where experiments have been carried out with two types of algorithms: simple suffix stripping algorithm and score based stemming cluster identification algorithm. Another rule-based intelligent Bangla parser is to ease the task of handling semantic issues in the subsequent stages in machine translation [8].

B+ tree data structure for indexing Object Oriented Database has been found time efficient searching, inserting and deleting algorithms in [9]. Clustered Metric Tree(CM-Tree) had been constructed for a dynamic clustered index in unstructured metric database for searching similarities [10]. Variations of the TF-IDF(Term Frequency - Inverse Document Frequency)

weighting scheme are often used by IR techniques as a central tool in scoring and ranking a documents relevance given a user query [11]. Many TF-IDF weights function has been developed for verious IR applications like, an improved TF-IDF weights function is proposed which uses the distribution information among classes and inside a class [12]. Improved version of a well functional weighting scheme of TF-IDF algorithm based on question title or option or answer terms, Vector Space Model [13] for representing documents as a vector in t-dimension (If the document contain t no. of terms) and find the similarity with query vector using cosine similarity product rules are the main area of IR techniques used in QB3S. A hash-mapping WordNet is used to handle same meaning words or synonyms. Main objective of our research showing below:

- To improve in a certain level of Bangla documents processing tools like tokenization, stopwords removing, stemming to accomplish the processing of QB data-set.
- To analyze the structure of data and indexing in sorted sequential records(Clustered Indexing) of QB database for faster access by using B+ tree data structure.
- To construct a hash mapping word-net to handle synonyms.
- To improve a weighted functional TF-IDF algorithm.

The rest of the paper has been organized as follows. Section II describes the structure of QB[14-17]. The system architecture for QB3S has been given in section III. Section IV shows the results obtained after the application of QB3S in the existing eLearning system [1]. A detailed discussions on the results has also been given in this section. Finally the conclusion is given in section V.

II. Question bank Structure and e-learning of ICT course

A sequence of papers [14-17] were published in journals or conference towards improving PBeL in ICT course of secondary and higher secondary level of Bangladesh. Problem based e-learning system (http://epbl.org) for interactive learning of ICT course, practicing and model examination from question bank and C programming, HTML, SQL, ERD learning based on higher secondary book have already been implemented successfully. Question bank data are important resource for Problem Based e-Learning(PBeL) which requires a search engine to retrieve query information from the large scale of data and this can be used by the admin of PBeL system to find the duplicate question, user can find out searching result of the relevant query input.
Let us take a closer look at the information and utilities available in a bangla question bank database. It can be Categorized the Multiple choice questions in three classes:

1) Cognitive class:
   The simplest structure where there are a question title and 4 options.Like,
   আধুনিক কম্পিউটারের জনক কে?
   ক) অ্যাডা লাভলেস
   খ) চার্লস ব্যাবেজ
   গ) বিল গেটস
   ঘ) জেমস ক্লার্ক ম্যাক্সওয়েল
2) Analytical class:
   It contains three option and the scope of answer is the four combinations of these options.As example,
   কম্পিউটারের সময় অপচয়কারী ব্যবহার –
   i) প্রোগ্রামিং শিক্ষা
   ii) ভিডিও গেম
   iii) সামাজিক যোগাযোগ
   নিচের কোনটি সঠিক ?
   ক) ii ও iii খ) i ও iii
   গ) i ও ii ঘ) i. ii ও iii
3) Higher Ability class: One of the most complex structure and scenario based questions.There are may be two or three Cognitive or Analytical mcq question based on a scenario. As example,
   নিচের অনুচ্ছেদটি পড় এবং 1-2 নং প্রশ্নগুলোর উত্তর দাও:
   বর্তমান আমাদের দেশে বেকার সমস্যা মানাতœক সামাজিক ব্যাধি। প্রায় প্রতিটি সমাজেই এ ব্যাধিতে আক্রান্ত। বেকার সমস্যা দূর করতে বর্তমানে HTML, CSS, JAVA Script সহ বিভিন্ন ধরনের CMS সফটওয়্যার এর মাধ্যমে ওয়েব ডিজাইন করে বেশ অর্থ উপার্জন করছে।
   1. CSS- এর পূর্ণরূপ কী?
   ক. Caseding Style Sheet
   খ. Compund Style Sheet
   গ. Compund Shorting System
   ঘ. Cacacing Simple Sheet
   2. কত সালে তাদের প্রতিযোগিতামূলক ব্রাউজার রিলিজ করে?
   ক. ১৯৯৫ সালে
   খ. ১৯৯৬ সালে
   গ. ১৯৯৭ সালে
   ঘ. ১৯৯৮ সালে

Orgatization of stroed question in database schema is showing in Fig. 1. We need to analysis the structure of question and cluster it with three decision boundaries of three different classes will be described in sec. III.

	mcqid	qtitle	op1	op2	op3	op4	op5	image	answer
	c110010030051	ফুল অ্যাডার সার্কিটের সরাক সারণিতে করটি সা…	3	8	4	16			8
	c110010030052	অ্যাডার এবং রেজিস্টার ব্যবহৃত হয় কম্পিউটারে…	প্যারিটিক …	ষ্যাটাডিক	মেমোরি	মনিটরে			প্যারিটিক ও থৌক্তিক ই…
	c110010030053	শিফট রেজিস্টারে ইনপুট থাকে-	ঘর বিটের …	1 টি	ঘর বিটের রে…	কোনোটি নয়			1 টি
	c110010030054	একটি 4 বিট রেজিস্টারের 0000 সংরক্ষণ করার উ…	ইনপুটে 00…	Reset ইন…	ইনপুটে 0000 …	ইনপুটে 0000 দেওয়া, …			ইনপুটে 0000 দেওয়া, …
	c110010030055	D ফ্লিপ ফ্লপ এর ক্ষেত্রে আউটপুট 0 থেকে 1 করতে…	0, 0 থেকে 1	1, 1 থেকে 0	1, 0 থেকে 1	1, 1 থেকে 1			1, 0 থেকে 1
	c110010030056	ক্লকের Edge কোনটি?	0 থেকে 1 …	1 থাকা অ…	0 থাকা অবস্থান…	যে কোনোটি			0 থেকে 1 রূপান্তরের স
	c110010030057	নিচের কোনটি কাউন্টারের ব্যবহার নয়-	CPU-G প্রো…	সার্কিস সে…	ঘড়িতে	মনিটরে			মনিটরে
	c110010030058	কাউন্টারের একটি বিটের আউটপুট পরবর্তী বিটে…	অ্যাসিনক্র…	বাইনারি …	BCD কাউন্টার	সিনক্রোনাস কাউন্টার			অ্যাসিনক্রোনাস কাউন্টার
	c110010040001	ওয়েব সার্ভারে থাকা ওয়েব সাইটটি দেখার জন্য …	ওয়েব ব্রাই	ওয়েব ব্রা…	ডোমেইন ব্রাউ	সার্ভার ব্রাউজার			ওয়েব ব্রাউজার
	c110010040002	WWW এর পূর্ণ রূপ হলো -	World We…	Website …	World Wide Web	Web Wide World			World Wide Web

Fig. 1. Organization of questions in database

III. QB3S system architecture

Theoretical analysis of the body of methods and principles associated with branches of procedure. QB3S has taken an input mcq or part of a mcq, processing data and performing search for finding the maximum similar mcq from entire QB DB. See Fig. 2, showing the architecture of system in terms of process flow. The whole process can be split into four modules:

1) Question structure analysis and clustering module
2) Document processing module
3) WordNet module
4) IR module

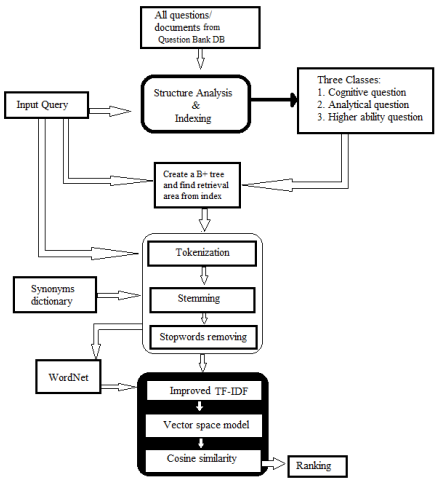


Fig. 2. System Architecture of QB3S

A. Question Structure analysis and clustering

In sec. II, we have been already discussed about the structure of question bank data and categorized them in three classes. Criteria of classification based on structure of question is showing in Fig. 3.

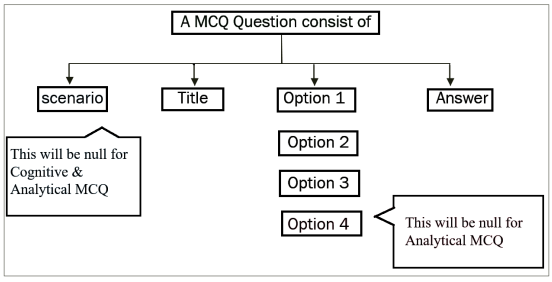
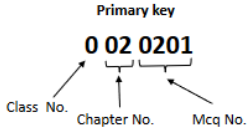


Fig. 3. Question structure analysis

PBeL system of ICT course database schema has been designed such a way that a table allocated for each class. Each table for a specific class contains identical questions of all chapters.

Clustered indexing by B+ tree:

B+ tree data structure has been used to track the clustered index(primary key on sorted order) of database table. B+ tree efficiently reduce the access time of search space in database of QB. In case of ICT course QB, it has 6 chapters and 3 classes of question. Primary key of table has been maintained as,



We had been compelled to track 18 indexes in the leaves node, each one pointed to the starting key of sorted data and degree of the B+ tree 6 in this case(see Fig. 4). So the height of tree ceil(log(18)/log(6/2)) = 2.

During the search operation, h nodes are read from the disk to the main memory where h is the height of the B+ tree and $h = log_t(n)$, where n is the number of the keys stored in the tree and t is size of a block or node . In addition of the disk reads, B+ tree searching algorithm performs a linear search in every node read from the disk. The time complexity of each linear search is O(t) . Thus, the total time complexity of the B + tree search operation is $O(tlog_t n)$ and same time complexity for inserting and deleting [18].

In case of higher number of chapter, it needs higher order and for adding other courses in QB, extra level will be added dynamically by imposing precondition [19].

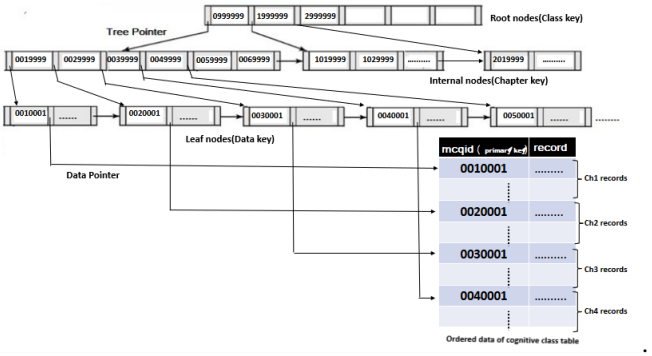


Fig. 4. B+ tree structure for clustered indexing

B. Bangla Document Processing

In the field of Bangla language processing, we have been already discussed the existing work of tokenization, Stemming and stop words removing in Sec. I. Now we are going to explained the tools of language processing which have been used in QB3S.

1) Tokenizer: Tokenization is the process of breaking up the given text into units called tokens.Tokenization refers lexical analysis in natural language processing.

We use a standard tokenizer, which splits text into terms on word boundaries using white space, as defined by the Unicode Text Segmentation algorithm [20]. It removes most punctuation symbols. Array-list data structure has been used to store the terms after tokenization.

2) Stemmer: A very simple-to-use rule based stemmer [6] for Bangla word parsing is used for semantic analysis. Which take a inflected word as input and output the corresponding root stem. Some common rules which is used for semantic analysis showing below and See examples of common rules text file contains in Fig. III-B2.

X #When X appears at the end of a word, remove it

$Y \rightarrow Z$ #When Y appears at the end of a word, replace it with Z

$Y.Z \rightarrow A.B$ #When Y, followed by some character a, followed by Z appears at the end of a word,replace it with AaB

(

ই	৯ এটাই, সেটাই
ও	৯ এটাও, সেটাও
তো	৯ হবতো, করলতো

)

(

কে	৯ এটাকে, আমাকে
তে	৯ যসতে, গাইতে
রা	৯ বহিমরা, করিমরা

)

Fig. 5. Examples of some common rules for stemming

3) Stop words Remover: 'Stop Words' usually refers to the most common words in a language.There is no single universal list of bangla stop words. We build stop words list in followed way:

1) From some pre defined words
2) From the most frequently occurring word in Question Bank database

All stop words are removed from the array-list by using UTF-8 string matching library in java [21].

C. Word-Net module

As you know, synonyms are words that have similar meanings. A synonym set, or synset, is a group of synonyms. A synset, therefore, corresponds to an abstract concept.We have been constructed a runtime hash mapping from text file which contains bangla synonym dictionary(see Fig. 6).

প্রতিস্থাপন করা;বিনিময়,স্থলাভিষিক্ত করা,আদান প্রদান করা,বিকল্প
নেটওয়ার্কিং;জালাধান
রক্ষণাবেক্ষণ;যত্ন
জড়ো হওয়া;জমান,সমবেত হওয়া,সঙ্কলন করা,সংগ্রহ করা
সুবিচার;যথার্থতা,বিচারক,
ভংগুর;তুচ্ছ,শিষ্ট,ভঙ্গুর,ক্ষীণ হওয়া
ভান্ডার;জমা হওয়া,ন্যাসরক্ষক,স্মৃতিসৌধ,সচিব
হেলিক্স;পেচানো
ডিকোডিং;ডিক্রিপ্ট করা
বিধিলঙ্ঘন;অমান্য,বিধিলঙ্ঘন,অপকর্ম

Fig. 6. Bangla synonym dictionary txt file

Step of hash-mapping word-net construction:

1) Take input from bangla dictionary text file
2) Tokeniztion
3) Stop-word removing
4) Stemming
5) Use hash function, $f(n) = charAt(0) \mod 128$ for finding hash value (total 128 bangla character,unicode value 0980H-09FFH).

Hash-map of the first line of Fig. 6 showing in Fig. 7.

Searching time complexity : O(1)



Fig. 7. Representation of dictionary hash-map

D. Information Retrieval module

1) Improved version of Weighted TF-IDF: Term Frequency(TF) measures how frequently a term occurs in a document.Suppose we have a query or input mcq consisting terms, $t_1, t_2, t_3, t_4 ..... t_n$ and there are many documents $mcq_1, mcq_2, mcq_3 ..... mcq_m$ where the tf-idf will be performed.Then,

$$TF(t, mcq_j) = \frac{f(t, mcq_j)}{\sum_{T \xi mcq_j} f(T, mcq_j)} \quad (1)$$

$Where, f(t, mcq_j) =$ Numbers of time Term t appear in mcq_j

Inverse Document Frequency(IDF) Measures how important a term is. Log base e of ratio between total number of documents and Number of documents with that term t in it.

$$IDF(t) = ln\frac{N}{n} \quad (2)$$

Where, N = Total number of documents And n = Number of documents with term t

We have been developed an improved version of this algorithm where we used some pre-defined weighting factor based on the classification of question structure.
Weighting Factor,

$$W(t) = \begin{cases} w1 & if\ t\ is\ a\ Answer \\ w2 & if\ t\ is\ an\ option \\ w3 & otherwise \end{cases}$$

Where, w1> w2>w3
So tf-idf equation with pre-define weighting factor as like,
$TF - IDF_{improved}(t, mcq_j) = \frac{f(t, mcq_j) * W(t)}{\sum_{T \xi mcq_j} f(T, mcq_j) * W(T)} * ln\frac{N}{n}$

2) Vector space Model Representation from TF-IDF: Vector space model or term vector model [22] is an algebraic model for representing text documents as vectors of identifiers like index terms. Formally, a vector space is defined by a set of linearly independent basis vectors.The basis vectors correspond to the dimensions or directions of the vector space. The basis vectors are linearly independent because knowing a vectors value on one dimension does not say anything about its value along another dimension. Documents and Query are represented as vectors.

$$D_j = (w_{1,j}, w_{2,j}, \ldots, w_{t,j})$$

$$Q = (w_{1,q}, w_{2,q}, \ldots, w_{n,q})$$

Each dimension corresponds to a separate term. On the off chance that a term happens in the document, its incentive in the vector is non-zero. Graphical representation of a vector space in Fig. 8.

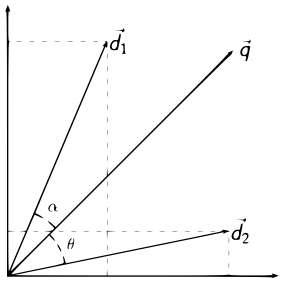


Fig. 8. Vector space model

Vector Space Model (VSM) is a common technique to depict the text feature in text classification. It adopts TF-IDF weights to compute the term weighting in each element of the text feature. Be that as it may, it just considers the connection between the term and the entire content however disregards the relationship between various terms. VSM has been used for representing tf-idf weight of a mcq as a vector in the space. Let's assume we have a query mcq '**Q**' like,

নিচের কোনটি উচ্চস্তরের ভাষা?
A. পাইথন
B. ভিজুয়াল বেসিক
C. জাভা
D. সব ক'টি
Answer: সব ক'টি

After processing it using the method described in Subsec. III-B '**Q**' will be transformed into:

নিচ উচ্চস্তর ভাষা পাইথন ভিজুয়াল বেসিক জাভা সব

Query '**Q**' can be represented in the 7-dimensional vector space. Calculated the value of TF-IDF in a wighting matrix where each row represents a vector of a document in the space of query document.After performing improvedTF-IDF algorithm the weighting matrix is showing in Table I.

TABLE I
VECTOR REPRESENTATION FROM TF-IDF VALUE IN VSM

Doc	$term_1$	$term_2$	$term_3$	$term_4$	$term_5$	$term_6$	$term_7$
mcq_1	0.0	0.05	0.05	0.05	0.05	0.05	0.05
mcq_2	0.05	0.08	0.0	0.01	0.0	0.03	0.0
mcq_3	0.0	0.12	0.22	0.33	0.44	0.55	0.83
mcq_4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mcq_5	0.0	0.0	0.03	0.03	0.03	0.03	0.03
mcq_6	0.09	0.0	0.11	0.1	0.1	0.1	0.1

3) Similarities measurement using Cosine Similarity: Cosine similarity is a measure of similarity between two non-zero vectors of an inner product space that measures the cosine of the angle between them. This metric is a measurement of orientation and not magnitude, it can be seen as a comparison between documents on a normalized space because were not taking into the consideration only the magnitude of each word count (TF-IDF) of each document, but the angle between the documents. To assign a numeric score to a document for a query, the model measures the similarity between the query vector and the document vector. The similarity between two vectors is once again not inherent in the model. Typically, the angle between two vectors is used as a measure of divergence between the vectors, and cosine of the angle is used as the numeric similarity [23]. Cosine has the nice property that it is 1.0 for identical vectors and 0.0 for orthogonal vector. If $\vec{D}$ is the document vector and $\vec{Q}$ is the query vector, then the similarity of document D to query Q (or score of D for Q) can be represented as:

$$sim(\vec{D}, \vec{Q}) = \cos\theta = \frac{\vec{D} \cdot \vec{Q}}{\|\vec{D}\|\|\vec{Q}\|} = \frac{\sum_{i=1}^{n} D_i Q_i}{\sqrt{\sum_{i=1}^{n} D_i^2}\sqrt{\sum_{i=1}^{n} Q_i^2}} \quad (3)$$

Cosine Similarity will generate a metric that says how related are two documents by looking at the angle instead of magnitude, like: Using the eq. 3, we can find out the similarity

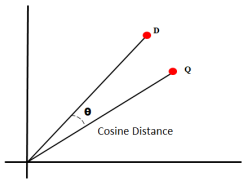


Fig. 9. Cosine distance between $\vec{D}$ and $\vec{Q}$, where $0 \leq cos\theta \leq 1$

between Query vector and other mcqs in VSM showed in table. I .The **Cos**θ values for different mcqs or docs in range of 1 (Most Similar) and 0 (less similar) and ranking(see table. II) from calculated cosine value.

TABLE II
RANKING FROM COSINE VALUE

Doc	$CosineValue$
mcq_1	0.93
mcq_3	0.81
mcq_2	0.64
mcq_5	0.32
mcq_6	0.24
mcq_4	0.00

IV. RESULT AND DISCUSSION

QB3S has been evaluated its first experiment in a multiple choice question bank(ICT) database. It contains total 1289 mcq. Information of experimental data-set is showing in table. III.

We implied two test case with present or absent of condition and analysis the result graphically. Accuracy of searching performance, Sensitivity and specificity have been calculated

TABLE III
EXPERIMENTAL DATA-SET

Chapter ID	Chapter Name	Total Question
c11001001	Information and Communication Technology	304
c11001002	Communication Systems and Networking	245
c11001003	Number system and digital device	273
c11001004	Web Design Contacts and HTML	166
c11001005	Programming language	152
c11001006	Database management system	149

from the value of True Positive(detects the condition when the condition is present), False Negative (does not detect the condition when the condition is present), True Negative(does not detect the condition when the condition is absent) and False Positive(detects the condition when the condition is absent) [24].

Test 1: Select 50 existing full or part of a mcq as query inputs from 'Communication Systems and Networking' chapter and found the maximum cosine valued mcq. Graphical result showing in Fig. 10.

Test result can be True Positive(TP) or False Negative(FN) depends on the value of cosine similarity.In case of TP, cosine value greater than 0.5, it detects the condition where query mcq belongs to this chapter is satisfied. otherwise FN.

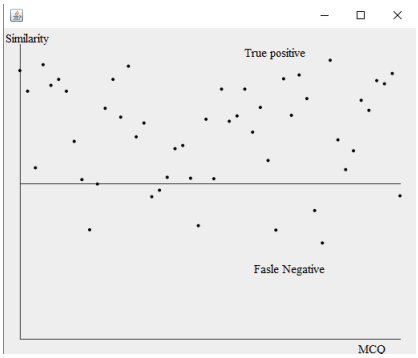


Fig. 10. Result of test 1 where 42 True Positive(TP) and 8 False Negative(FN) value was detected

Test 2: Select 50 full or part of a mcq as query inputs from other chapter except "Communication Systems and Networking" chapter(Not exist in this chapter) and find the mahapterximum cosine value holder mcq.Graphical result showing in Fig. 11.

Test result can be True Negative(TN) or False Positive(FP) depends on Cosine value. In case of TP, cosine value less than 0.5, it detects the condition where query mcq doesn't belongs to this chapter, otherwise FP.

In this experiment of QB3S, we found TP=42, FN=8, TN=43, FP=7

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} = \frac{42+43}{42+8+43+7} = 85\%$$
$$Sensitivity = \frac{TP}{TP+FN} = \frac{42}{42+8} = 84\%$$
$$Specificity = \frac{TN}{TN+FP} = \frac{43}{43+7} = 86\%$$



Fig. 11. Result of test 2 where 7 False Positive(FP) and 43 True Negative(TN) value was detected.

V. CONCLUSION

Information retrieval techniques are effective in finding similar documents from a text database. These techniques are not efficient in finding the similarity of questions in a Question Bank developed for PBeL systems for many reasons: i) the document structure is different in nature that the conventional text documents, and ii) the contents of the different parts of the documents have different weights.

In this paper, we have analyzed the structure and weight of the different types of questions in the QB and developed an improved TF-IDF algorithm based on the weight. At the initial stage we applied NLP tools and techniques like tokenization, stemming, stopword removal in the bangla text. Handling synonym is critical in any similarity searching system. We have developed a WordNet based on hash technique technique count of of synonym in TF-IDF. We have created a the Vector Space Model using the improved TF-IDF weighted matrix. For faster and efficient access to the QB DB, we have used a B+ tree index structure.

Using the above techniques, we have developed a Question Bank Similarity Searching System(QB3S). We have applied QB3S in a real life dataset emerging from the PBeL systems for ICT course of HSC level in Bangla. We have achieved an accuracy level of 85% for similarity search in the QB.

REFERENCES

[1] G. M. M. Bashir, A. Latiful Haque, and B. Chandra Dev Nath, E-learning of php based on the solutions of real-life problems, vol. 3, 12 2015.

[2] Baeza-Yates, Ricardo, and Berthier Ribeiro-Neto. Modern information retrieval. Vol. 463. New York: ACM press, 1999.

[3] Chowdhury, Gobinda G. "Natural language processing." Annual review of information science and technology 37.1 (2003): 51-89.

[4] F. Alam, S. Habib, and M. Khan, Text normalization system for bangla, tech. rep., BRAC University, 2008.

[5] J. Pustejovsky, The generative lexicon, Computational linguistics , vol. 17, no. 4, pp. 409441, 1991.

[6] S. Das and P. Mitra, A rule-based approach of stemming for inflectional and derivational words in bengali, in Students Technology Symposium (TechSym), 2011 IEEE , pp. 134 136, IEEE, 2011.

[7] A. Das and S. Bandyopadhyay, Morphological stemming cluster identification for bangla, Knowledge Sharing Event-1: Task , vol. 3, 2010.

[8] G. K. Saha, Parsing bengali text: An intelligent approach, Ubiquity , vol. 2006, no. April, p. 1, 2006.

[9] Ramaswamy, Sridhar, and Paris C. Kanellakis. "OODB indexing by class-division." ACM SIGMOD Record. Vol. 24. No. 2. ACM, 1995.

[10] Aronovich, Lior, and Israel Spiegler. "CM-tree: A dynamic clustered index for similarity search in metric databases." Data & Knowledge Engineering 63.3 (2007): 919-946.

[11] N. Wang, P. Wang, and B. Zhang, An improved tf-idf weights function based on information theory, in Computer and Communication Technologies in Agriculture Engineering (CCTAE), 2010 International Conference On , vol. 3, pp. 439441, IEEE, 2010.
[12] Ramos, Juan. "Using TF-IDF to determine word relevance in document queries." Proceedings of the first instructional conference on machine learning. Vol. 242. 2003.
[13] Turney, Peter D., and Patrick Pantel. "From frequency to meaning: Vector space models of semantics." Journal of artificial intelligence research 37 (2010): 141-188.
[14] A. Habib and A. L. Hoque, Towards mobile based e-learning in bangladesh: A frame- work, in Computer and Information Technology (ICCIT), 2010 13th International Con- ference on , pp. 300305, IEEE, 2010.
[15] A. Hoque, M. M. Islam, M. I. Hossain, and M. F. Ahmed, Problem-based e-learning and eval-uation system for database design and programming in sql, International Journal of E-Education, EBusiness, E-Management and E-Learning-IC4E , pp. 537542, 2013.
[16] A. S. M. L. Hoque, G. M. M. Bashir, and M. R. Uddin, Equivalence of problems in problem based e-learning of database, in Technology for Education (T4E), 2014 IEEE Sixth International Conference on , pp. 106109, IEEE, 2014.
[17] G. M. M. Bashir and A. S. M. L. Hoque, An effective learning and teaching model for programming languages, Journal of Computers in Education , vol. 3, no. 4, pp. 413437, 2016.
[18] Pollari-Malmi, Kerttu, and Eljas Soisalon-Soininen. "Concurrency control and i/o-optimality in bulk insertion." International Symposium on String Processing and Information Retrieval. Springer, Berlin, Heidelberg, 2004.
[19] Bruso, Kelsey L., and James M. Plasek. "Dynamic preconditioning of A B+ tree." U.S. Patent No. 7,809,759. 5 Oct. 2010.
[20] Davis, Mark, and L. Iancu. "Unicode text segmentation." Unicode Standard Annex 29 (2012).
[21] Duerst, Martin. "The properties and promises of UTF-8." Proc. 11th International Unicode Conference, San Jose. 1997.
[22] Lee, Dik L., Huei Chuang, and Kent Seamons. "Document ranking and the vector-space model." IEEE software 14.2 (1997): 67-75.
[23] A. Singhal et al. , Modern information retrieval: A brief overview, IEEE Data Eng. Bull. , vol. 24, no. 4, pp. 3543, 2001.
[24] D. M. Powers, Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation, 2011.

Textual Question Answering for Semantic Parsing in Natural Language Processing

Jaydeb Sarker
Dept. of CSE
Rajshahi University of Engineering and Technology
Rajshahi, Bangladesh
jaydeb.ruet@gmail.com

Mustain Billah
Dept. of IT
University Of Information Technology and Sciences
Dhaka, Bangladesh
mustain.billah@uits.edu.bd

Md. Al Mamun
Dept. of CSE
Rajshahi University of Engineering and Technology
Rajshahi, Bangladesh
a.mamun@cse.ruet.ac.bd

***Abstract*—Question answering (QA) problem is one of the most frequent tasks in natural language processing (NLP) over language input. To understand the natural language, semantic parsing is an extensively used process. For deriving syntactic and semantic structures of texts, synchronous frameworks are being used frequently. With the aim of answering questions efficiently, in this paper, an approach by semantic parsing of texts with a view to leveraging semantic information is introduced. The proposed approach uses the method based on lambda calculus for semantic parsing to derive logical forms of sentences. The questions are analyzed by collecting significant features to find out correct answers from the facts. Unlike traditional approaches, this paper extends lambda calculus based dependency parser methodology for question analysis. This proposed approach for answering questions is based on generalized searching techniques. Proposed system achieves 95% accuracy for Yes/No questions with an overall 83% mean accuracy for the five tasks of bAbI-10k question answering dataset surpassing the existing approach by about 11%.**

***Index Terms*—Natural Language Processing, Semantic Parsing, Lambda Calculus, Question Answering, Logical Forms, Text Parsing.**

I. Introduction

A complicated task in natural language processing [1]–[4] is question answering which requires the understanding of the meaning of a text and the ability to reason over relevant facts. To represent the meaning of sentences of the human language, semantic parsing is the best-defined tasks that support automated inference for knowledge representation [5], [6]. Relevant works includes Transforming Dependency Structures where they use lambda calculus for parsing method to find dependency of sentences [7]. The dependency of words in a sentence can be found by open text semantic parsing [6].

For building a question answering system, it is important to analyze the question and input text; then the answer has been found from the relevant facts. One of the popular question answering task is factoid question answering to find the correct answer in documents [7]. Dynamic memory network, a neural network architecture has been introduced by Kumar et al. that processes input sequences and questions, forming episodic memories, and generates relevant answer [8]. Query-Regression Network (QRN) approach, a variant of recurrent neural network (RNN), is able to effectively handle long-term dependencies and is highly parallelizable [9]. This model is also suitable for end-to-end machine comprehension. Todai robot project on Japanese examination system has been learned the exam answers and test that from the student answers [10]. Question answering by using syntactic information is also enhanced in several attempts using heuristics analysis for ranking potential answers [11]. Some interesting works on question answering system are found in [14]–[16], [20]. QA module has been also designed for business module [15].

In the question answering system, syntactic information gives good accuracy than information retrieval. In proposed approach both sentences and selected questions in the top documents are parsed and stored in their corresponding database. We propose a novel technique to find out the answer to questions by a generalized searching algorithm. Our method first analyzes the input text and questions and find the relevant facts. It finds the answer by searching technique from the relevant facts. Fig.1 provides an example for question answering technique:

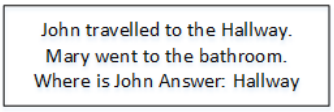


Fig. 1. Example of inputs and questions, together with answers.

This paper consists of following sections: Proposed structure and methods are discussed in section II. Section III contains

the results and discussions. Finally, the paper is concluded in section IV.

II. Proposed Structure and Methods

Proposed system introduces a searching based novel approach to find the answers from corresponding text. However, it consists of four modules: Input module, Facts finding module, Question Module and Answer module (figure 2). Proposed system is designed to map the sentences to its logical forms. Features from questions are found and the actual answers are extracted by searching from facts.

It first computes a representation for all inputs and the questions. The question representation has been triggered an iterative attention process that searches the inputs and retrieves relevant facts.

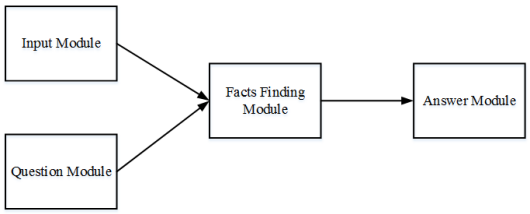


Fig. 2. Proposed Question Answering System

A. Input module

In the input module, a set of text is given. The text represents a long sequence of sentences of a story or movie. All the sentences are in simple sentence structure which have sequential meaning. Firstly, each of the sentence is being parsed from the text by nltk POS tagger [13]. Logical forms of each sentence are derived from lambda calculus based dependency parsing method [12]. Lambda calculus based dependency parsing method has been chosen for parsing because it puts the main $verb$ as root word of a sentence. The semantic form of each sentence depends upon the $verb$. The $subject$ and $object$ of a sentence have been affiliated by verb.

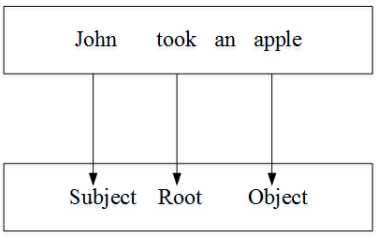


Fig. 3. Lambda calculus parsing method.

The facts are found from the text by parsing method based on lambda calculus dependency parsing [12]. Lambda calculus method has been used for making a relationship with $subject$ and $object$ of a sentence by taking the verb as a root of the sentence. An example of lambda calculus method is depicted in fig. 3.

B. Facts finding module

The given input text has been analyzed in this module. The $subject$ and $object$ of the sentence are also derived by checking tagged parts of speech of each words of that sentence. The $verb$, $subject$ and $object$ from each sentence have been taken and stored in corresponding database for easily finding the answer.

Logical forms derived from lambda calculus dependency parsing [12] will help further to find the features those eventually lead us to analyze question to find answers.

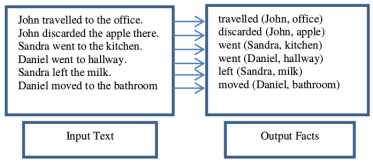


Fig. 4. Lambda calculus parsing method is used to find facts from input text.

In fig. 4, a sequence of input sentences have been parsed by lambda calculus dependency parsing method. Three types of feature are found from each sentence. The extracted features are stored in their corresponding databases, like verb database (VB_i), subject database (SB_i) and object database (OB_i). It assumes that if there are n number of sentences in a text, then the number of $verb$, $subject$ and $object$ are also n numbers. Here *i* represents the position of sentences in each text. Answer of the question has been extracted from these databases. The database of $verb$, $subject$ and $object$ extracted from fig. 4 has been shown in fig. 5, 6, 7.

Travelled, discarded, went, went, left, moved

Fig. 5. Verb database (VB_i) for corresponding input text.

John, John, Sandra, Daniel, Sandra, Daniel

Fig. 6. Subject database (SB_i) for corresponding input text.

Office, apple, kitchen, hallway, milk, bathroom

Fig. 7. Object database (OB_i) for corresponding input text.

C. Question Module

The corresponding question is also expressed in natural languages which is related to the input text. We mainly work on analyzing two types of question. They are 'W/H' question and 'Yes/No' question. 'W/H' questions are started with 'WH' tagged word by using the natural language toolkit (NLTK) [13]. Example:

$$Where\ is\ Daniel?$$

The answer of the question which starts with $Where$, indicates that the answer should be a place. Similarly, 'W/H' question extracts the factual answers from the text. On the other hand, 'Yes/No' question starts with a auxiliary verb (like, am, is, are etc). This type of question only requires 'Yes' or 'No' as answers.

After parsing and tagging each word of a question, we find two features from that sentence.

- Question Word (QW)
- Object Word (OW)

Question represents a sentence according to the input text. For example $Where\ is\ Milk?$ is an input question. Question should be corresponding to the input sequence. The important features from the question have been derived by extracting the features which is depicted in fig. 8. The two important features (w/h question word and object word) have been derived from each question.

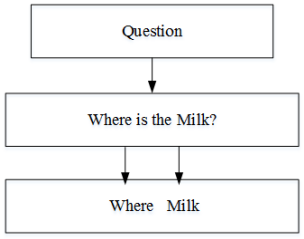


Fig. 8. Feature finding from W/H question.

Let we take two main features in QW(question word) and OW(object word). Fig. 8 shows that
QW = 'where'
OW = 'milk'
QW is also known as W/H feature. W/H feature is 'where' which indicates to find the location of the milk.

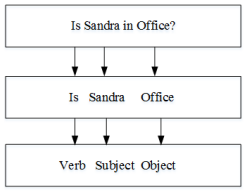


Fig. 9. Feature finding from Yes/No question.

Similarly, Figure 9 shows the feature finding from 'Yes/No' question word. Two features ($Subject$ and $Object$) are collected from this 'Yes/No' type question.

D. Answer module

This module generates the answer analyzing by the facts those are found from input module and question module. We have found two types of question in our test from the dataset. Firstly, we made a database of verb which indicates the position of that subject. For example, $Daniel\ Moved\ to\ office$. In this case, $moved$ indicates the location of the subject of this sentence. The positional verb database T_i is shown in figure 10.

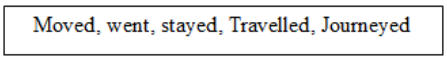


Fig. 10. Positional verb database T_i.

Storage of positional verb helps us to find the answer of a question which starts with 'where'. For finding the last position of the given Object Word (OW) in the question, it is important to search from the last position of the database and find the matches. The question which is started with 'where', our proposed system finds the answer by the given algorithm:

Algorithm 1: Finding answers of Questions starting with "Where"

Input: OW
Output: Position of SBi (where i is the position of sentence)

```
Take OW from the question
Check OW with SBi and VBi with Ti,
  answer = OBi
if OW matches with OBi then
    position of SBi of corresponding OBi is found by
      previous step.;
end
```

To test our algorithm, we give two examples here. For instance, the question is $Where\ is\ Sandra?$. Here, the OW = 'Sandra' and by using our algorithm the answer is 'kitchen'. It generates 'kitchen' as answer because the last position of Sandra is in 'kitchen'.

We also design an algorithm for 'Yes/No' question answering. It is known to us that two features ($Subject$ and $Object$) are collected from the corresponding question for 'Yes/No' answers. The algorithm is as following:

Algorithm 2: Finding answers of Yes/No Questions

Input: Subject, Object
Output: Answer

```
Take Subject and Object from the question
if Subject = SBi and Object = OBi (where i is the
  position of sentence) then
    answer = 'Yes';
else
    answer = 'No'
end
```

For example, the question is $Is\ Sandra\ in\ Office$, where $Subject$ = 'Sandra' and $Object$ = 'Office'. By using our proposed algorithm, it finds no matches where 'Sandra' and 'Office' occurs in the same sentence. So, the system gives 'No'

as it's answer. Moreover, we also make the similar matching algorithm for finding the answer of question which starts with $what$.

III. Results And Discussion

A. Dataset Collection

To analyze proposed model, bAbI-10k English question answering dataset [9] has been used. It's a synthetic dataset featuring 20 different tasks. The dataset comes in two sizes: bAbI-1k and bAbI-10k, referring to the number of training examples of each task. Proposed system is tested by 5 different tasks of bAbI-10k. Five tasks are listed with their corresponding class name in table I.

TABLE I
THE TASK IN BABI/TASKS CORRESPOND TO THOSE FROM THE ORIGINAL DATASET AS FOLLOWS:

	Task	Class Name
1	QA with single supporting fact	WhereIsActor
2	QA with two supporting fact	WhereIsObject
3	QA with three supporting fact	WhereWasObject
4	Yes/No questions	IsActorThere
5	List/Sets	Listing

In case of single supporting facts, only the questions about the position of the $Subject$ are asked. Tasks with two supporting facts have class name $Where\ Is\ Object$ which indicates to find out the location of the object. But, three supporting facts uses the sequential sentences and the question is about the position of object using the word either $before$ or $after$. On the other hand, Yes/No questions just justify the truth value of the question. Another task is about Lists/Sets, where more than one answer would be generated from the question.

B. Experimental Analysis

Analyzing each questions using proposed algorithm 1 and algorithm 2, answers are found for each question. For Parts of Speech Tagging, proposed system uses the NLT POS tagger [13]. As there is no need to update memory like dynamic memory network [8] or QRN [9] , it is easy to find the relevant facts without any complexity. It can retrieve answers easily by searching from the databases of facts.

C. Performance Evaluation

Accuracy of proposed Question Answering system for different tasks are shown in TABLE II. Better accuracy have been gained for each tasks excepts Lists/Sets. TABLE II also shows the number of input sentences and questions for testing the system. The proposed system gives the highest accuracy (95%) for Yes/No questions. Without the Lists/Sets task, the system gives more than 80% correct answer for given questions. The performance degrades when new type of tasks are included as input in the system. Though it gives high performance for sequence of simple sentences but for compound and complex sentences, it will be tough to find the features by lambda calculus based dependency parsing method.

TABLE II
ACCURACY OF QA FOR DIFFERENT TASKS

	Task	No of Sentences	No Questions	System Accuracy
1	Single Supporting Facts	13	10	90%
2	Two Supporting Facts	25	10	90%
3	Three Supporting Facts	37	10	80%
4	Yes/No questions	40	20	95%
5	List/Sets	60	10	60%

Table III and Figure 11 show the comparison of Structured Support Vector Machine(SSVM) [14] and proposed system. Our proposed QA system gives a better accuracy than SSVM for Two and Three supporting facts. To be precise, our system gives 80% accuracy for Three Supporting facts whereas SSVM gives only 17%. The overall mean accuracy of our system is 83% for these five types of tasks which surpasses the SSVM. The mean accuracy of SSVM is 71.8% for this five types of tasks.

TABLE III
COMPARISON OF STRUCTURED SVM AND PROPOSED SYSTEM

	Task	Structured SVM	Proposed System
1	Single Supporting Facts	99%	90%
2	Two Supporting Facts	74%	**90%**
3	Three Supporting Facts	17%	**80%**
4	Yes/No questions	99%	95%
5	List/Sets	70%	60%
	Mean Accuracy	71.8%	**83%**

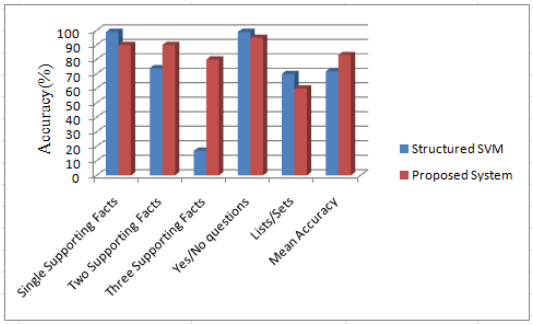


Fig. 11. Accuracy of Structured SVM vs Proposed System

We have considered 5 specific tasks for our work. For these tasks, our system shows better accuracy than some other existing methods. Most of the existing methods used machine learning approach which requires a large data sets. On the other hand, proposed method performs well for small dataset which resembles real life problems. Performance comparison with existing methods [14] is shown in Table IV.

IV. CONCLUSION AND FUTURE WORK

In this paper, a new methodology of question answering system has been introduced. POS tagging and semantic parsing techniques have been used for finding the answers from a given text. Proposed system is easy and less complex than existing systems in case of simple sentences. Moreover, it gains a better accuracy for five types of tasks in question answering

TABLE IV
PERFORMANCE COMPARISION WITH EXISTING METHODS

Methods	Accuracy
N-gram	26.8%
LSTM	36.6%
MemNN	68.8%
Proposed Method	**83%**

systems than the existing system. However, proposed question answering system is implemented only for simple sentences. In future, we will work with compound and complex sentences and long sequence of text. Machine learning technique with position vector will be used to learn the input text and retrieve the answer.

ACKNOWLEDGEMENT

We are very grateful to Biprodip Pal, Assistant Professor, Dept of CSE, Rajshahi University of Engineering Technology (RUET), Md Al Shayokh, Assistant Professor, Dept of IT, University of Information Technology and Sciences (UITS) and Syntax and Semantic Research Group (SySeReG) for their kind cooperation and contribution.

REFERENCES

[1] Fonferko-Shadrach, Beata, et al. "Using natural language processing to extract structured epilepsy data from unstructured clinic letters." International Journal of Population Data Science 3.4 (2018).
[2] Weston, Jason E., et al. "Techniques to predictively respond to user requests using natural language processing." U.S. Patent Application No. 15/077,814.
[3] Nie, Pengyu, et al. "Natural language processing and program analysis for supporting todo comments as software evolves." Workshops at the Thirty-Second AAAI Conference on Artificial Intelligence. 2018.
[4] Weston, Jason E., et al. "Techniques to predictively respond to user requests using natural language processing." U.S. Patent Application No. 15/077,814.
[5] Dreisbach, Caitlin, et al. "A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data." International Journal of Medical Informatics (2019).
[6] Kaur, Parneet, Vandana Pushe, and Rupinderdeep Kaur. "A Review of NLP Based Upon Semantic Analysis." (2014).
[7] Wang, M., "A survey of answer extraction techniques in factoid question answering," Computational Linguistics 1.1, vol. 9, 2006.
[8] Xiong, C., S. Merity, and R. Socher, "Dynamic memory networks for visual and textual question answering ," arXiv preprint arXiv:1603.01417, vol. 534, 2016.
[9] Seo, M., H. Hajishirzi, and A. Farhadi, "Query-Regression Networks for Machine Comprehension," arXiv preprint arXiv:1606.04582, vol. 534, 2016.
[10] Fujita, Akira, et al. "Overview of todai robot project and evaluation framework of its nlp-based problem solving." World History 36 (2014): 36.
[11] Li, X.,"Syntactic features in question answering, " Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval. ACM, vol. 38, 2003.
[12] Reddy, Siva, O. Tckstrm, M. Collins, T. Kwiatkowski, D. Das, M. Steedman, and M. Lapata. "Transforming Dependency Structures to Logical Forms for Semantic Parsing," Vol.: 38, 2016.
[13] Loper, E., and S. Bird, "NLTK:The natural language toolkit." Proceedings of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics-Volume 1. Association for Computational Linguistics, 2002.
[14] Weston, Jason, et al. "Towards ai-complete question answering: A set of prerequisite toy tasks." arXiv preprint arXiv:1502.05698 (2015).
[15] Curtis, Jon, Gavin Matthews, and David Baxter. "On the effective use of Cyc in a question answering system." Proc Workshop on Knowledge and Reasoning for Answering Questions. 2005.
[16] Sukhbaatar, Sainbayar, Jason Weston, and Rob Fergus. "End-to-end memory networks." Advances in neural information processing systems. 2015.
[17] Chung, Junyoung, et al. "Empirical evaluation of gated recurrent neural networks on sequence modeling." arXiv preprint arXiv:1412.3555 (2014).
[18] Graves, Alex. "Generating sequences with recurrent neural networks." arXiv preprint arXiv:1308.0850 (2013).
[19] Rajpurkar, Pranav, et al. "Squad: 100,000+ questions for machine comprehension of text." arXiv preprint arXiv:1606.05250 (2016).
[20] Rahman, Habibur, Julia Rahman, and Asmaul Husna. "WALGES: Weighted Probability Based Scoring Approach for Solving Algebraic Word Problems using Semantic Parsing." 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE). IEEE, 2019.

Information Retrieval System in Bangla Document Ranking using Latent Semantic Indexing

Md. Nesarul Hoque*, Rabiul Islam*, Md. Sajidul Karim*

*Dept. of Computer Science & Engineering
Bangabandhu Sheikh Mujibur Rahman
Science and Technology University
Gopalganj - 8100, Bangladesh

Email: mnhshisir@gmail.com, rabiul_islam_cse@yahoo.com, biplobsajidul@gmail.com

***Abstract*—Nowadays, like the English and other languages, Bangla also plays a significant role to strengthen the web repository. The storing rate of Bangla information is augmented day-by-day. Because of the numerous documents in the World Wide Web, it is very difficult for a user to retrieve the desired information. Furthermore, finding the useful documents tends to be more time spending as well as an annoying job. These demands emerge to develop an Information Retrieval (IR) system to document ranking for Bangla language. In this paper, we have built such a retrieval system where users can find their needed documents which correspond to their own query strings throughout the ranking index. Although a lot of works have been done for English and other languages to rank the documents, unfortunately, we have found a very negligible amount of contributions in Bangla Language. Many methods such as – Boolean model, Maximal Marginal Relevance (MMR), Portfolio Theory (PR), Quantum Probability Ranking Principle (QPRP), Query Directed Clustering (QDC), Vector-based TFIDF and so on, have been proposed to implement the document ranking system. Here, we have applied a new approach, called Latent Semantic Indexing (LSI) to do the same task for Bangla documents. LSI uses the mathematical method called Singular Value Decomposition (SVD). After that, we have applied the cosine similarity to rank all the documents. We believe that the performance result of our proposed system has reached the trustworthy level.**

Keywords*—IR, Bangla document ranking, LSI, SVD, Cosine similarity.

I. Introduction

At the present time, Bangla is considered the sixth popular language in the world and almost 230 million people use Bangla as their first and second language [1]. Furthermore, the use of Bangla language now increasing at a dramatic rate in the web repository. Within this vast information which is located in different documents, it is very tough to find the expected information from the desired documents. Again the online Bangla documents are highly categorized and there are a huge number of documents that carry the similar type of information. Readers are really confused which documents are more related that fulfill their demand. For these reasons, it is needed to develop a document ranking system for Bangla language based on the user query to get the expected information from the document based on the priority. There have been very few works in Bangla document ranking system. In this paper, for ranking the documents based on the more relativeness with the query, we have used a new approach, called Latent Semantic Indexing (LSI) [1] [2] [3]. LSI is an indexing and retrieval method that uses a mathematical technique called Singular Value Decomposition (SVD) to identify patterns in the relationships between the terms and the documents from an unstructured collection of text. A key feature of this technique is that it has the ability to produce the conceptual contents from a body of the document and each one is converted into a vector. Before applying the LSI, we need to generate a term-document matrix and a query matrix. Then we vectorize all the documents and the query string, using the formula of SVD which comes from the field of linear algebra. After that, we find the relativeness between the query vector and the document vectors through the method of cosine similarity. In that way we score all the documents and rank them based on the scores.

The rest of the paper has organized as follows: in **Section II**, we describe the related works. **Section III** conveys our proposed system and implementations details for ranking the document. Pre-processing for LSI technique has also been described here. **Section IV** involves the establishment of a baseline level of performance, and statistically investigates the importance of factors. The conclusion and the future direction of the paper are present in **Section V**.

II. Related works

In English and some other European languages, there are many document ranking techniques available.

Carbonell and Goldstein [4] proposed query rele-

vance based document ranking method named Maximal Marginal Relevance (MMR). However, there is a scope to increase the performance of the ranking system.

Shen and Zhai [5] combined the query history to enhance the retrieval performance in the document ranking process. However, this strategy always does not give desired outcomes. It performs inconsistent improvement in the retrieval performance.

Zuccon and Azzopardi [6] applied Quantum Probability Ranking Principle (QPRP) to improve the perfection of the document ranking for subtopic retrieval. They showed that the performance of QPRP takes advantages against the other document ranking techniques such as – Probability Ranking Principle (PRP) [7], Maximum Marginal Relevance (MMR) [4] and Portfolis Theory (PT) [8].

Zuccon et al. [9] proposed Interactive Probability Ranking Principle (iPRP) for document ranking which showed empirical effectiveness against to others ranking techniques. Here, they only worked on one possible interpretation. Interaction of other possible instantiations is not explored here.

Mitra et al. [10] proposed a method called Dual Embedding Space Model (DESM) which ranked the document according to the semantic relatedness between the query terms and all the document terms. They showed that this method performs better than the traditional term frequency based technique.

In contrast to the other languages, we get a bit of contribution in Bangla language to build up a document ranking system. Islam and Masum [11] first developed a system to rank the Bangla documents. They used term frequency with the help of a Unicode document corpus, named Bhasa. In this paper, they do not clearly specify the accuracy level of this system. Thakkar and Shrawankar [12] developed text categorization and summarization process. Here, manual selection is used for selecting a document from a set of categorized documents.

In this paper, to rank the Bangla documents, our proposed approach does these things in an effective way with more reliable compared to the existing systems.

III. Implementation Details

The main purpose of document ranking is to order the similar documents which correspond to the user query. To implement the whole process, we have divided this into two segments - pre-processing and final processing.

A. Pre-processing

In the pre-processing part, there are four tasks - removing the stop words, stemming the words, calculating terms-weight of each document and finally generating both term-document matrix and query matrix. Fig. 1 shows how pre-processing part works step by step in the document ranking system.

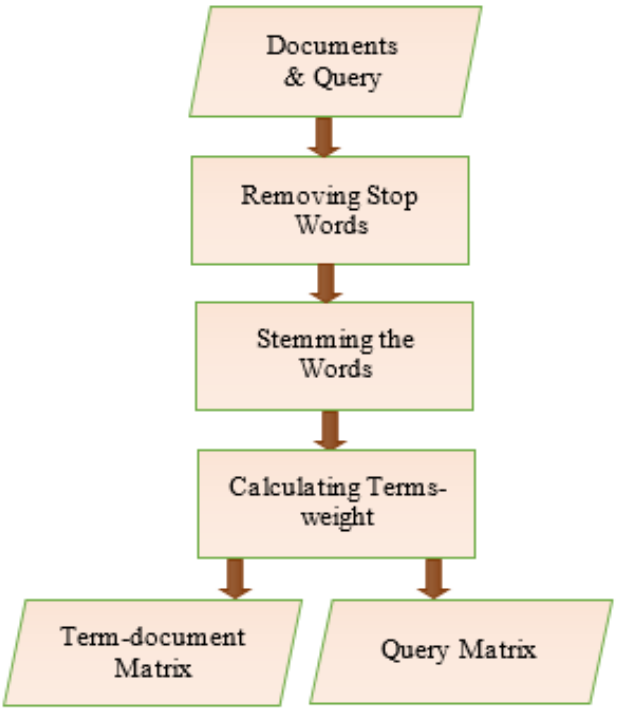


Fig. 1. Preprocessing part for generating term-document matrix and query matrix.

1) Removing stop words: Throughout the statistical analysis of the documents we have seen that some words have very low frequencies, while the others show just the opposite scenario. The most common thing of these words is that they do not carry any significant information about the document. Actually, they are used to maintain only the grammatical structure of the sentences. These words are known as the stop words [13]. From [14], we generate the stop words list. Here, we get 300 possible words. After inspecting on those words, we remove all the nouns and adjectives and add some non-information bearing words. Finally, we get 174 words of various types of articles, pronouns, prepositions and conjunctions, as the Bangla stop words.

Since the stop words of the documents do not play any role to carry out the actual meaning of the documents, at the first stage in the pre-processing, we filter all the stop words. It reduces the complexity of the overall ranking process.

2) Stemming: After removing the stop words, we apply the stemming on each word all over the documents. Stemming is the process which tries to convert the word into its root word by eliminating suffices. A suffix is considered as an affix which is embedded at the end of the root word according to the tense, numbers, persons, inflections, etc. Consider a word "বাংলাদেশ". It has various forms such as "বাংলাদেশের", "বাংলাদেশকে", "বাংলাদেশি", etc. In this paper, we have used a stemmer [15] which is applied all over the documents to get the root word from the suffix-word.

3) Calculating terms-weight: In this step, we calculate the terms-weight of each document. It is measured by counting how many times a word occurs in a particular document. Here, terms-weight will be used as the element of the term-document matrix and query matrix.

4) Generating a term-document matrix and a query matrix: Using the values of the terms-weight we build term-document matrix A, where terms represent the rows and each document corresponds to a column. By following the same procedure, a query matrix Q has been constructed.

B. Final Processing

Two matrices - terms-document matrix and query matrix, generated in the pre-processing part which are used in this segment. At first, we construct the document vector for each document and generate a new query vector by using the method of LSI [1] [2] [3]. After that, we apply the cosine similarity for each document vector with the query vector to rank the documents. All the steps in the final processing part are shown in Fig. 2.

1) Apply LSI: LSI is a good model which uses multidimensional vector space to establish a relationship between terms and documents. From each dimension in that space we get a concept which is called the topic layer or latent layer. Thus we get more topics of a document and these are encoded to vectors. Now, we generate the document vectors and the query vector.

a) Generate documents vectors:
LSI uses the mathematical technique, called Singular Value Decomposition (SVD) to get the document vectors from the term-document matrix A [Fig. 2]. SVD uses the formula comes from the linear algebra to decompose the matrix A into three matrices – two of them are orthogonal matrices V and U, where U will be transposed and the third one is diagonal matrix S, which entries are the singular values of A. The decomposition formula of matrix A can be written as, $A = USV^T$, where, U^TU and V^TV both will be the identical matrices.

To calculate the SVD we first find the eigenvalues and eigenvectors of AA^T and A^TA respectfully. The column of V is made up by the orthonormal eigenvectors of AA^T and the column of U is made up by the orthonormal eigenvectors of A^TA. The singular values in S are calculated by the square roots of eigenvalues of AA^T or A^TA in descending order. We can calculate the eigenvectors and eigenvalues using the following formula:

$$W\vec{x} = \lambda\vec{x} \tag{1}$$

Where, W is a square matrix comes from the matrix A, $\vec{x}$ is a nonzero eigenvector of W and λ represents a scalar value. Here, λ is said to be an eigenvalue of A and $\vec{x}$ is an eigenvector of A associated to λ. Using the value of λ we can determine the values of three matrices - U, S and V. Now we deduct the irrelevant or less important topic which has a rare contribution in the document scoring. Since in the S matrix, elements are organized in descending order, the lower values indicate that corresponding topics are less important or sometimes they are irrelevant. For these reasons, by neglecting the lower values, we reduce

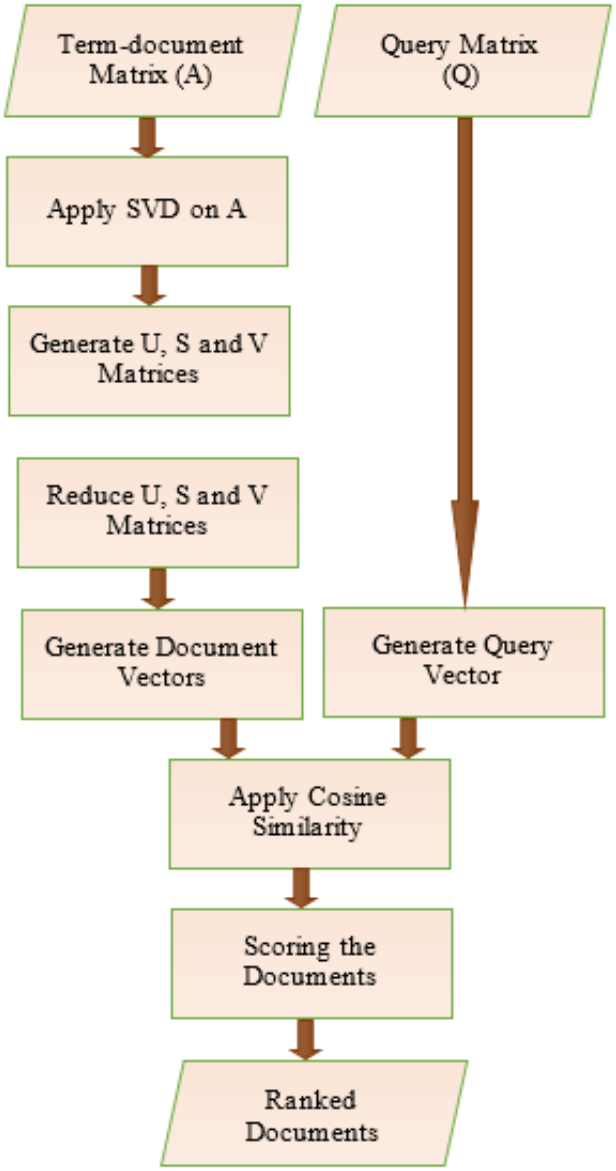


Fig. 2. Final processing part to rank the Bangla documents.

the dimension about 5% from the total dimension (n x n), then we get a new dimension i.e. k x k for the matrix S. The new column value k is used to all other matrices as $U \approx U_k$, $S \approx S_k$, $V = V_k$, and $V^T \approx V_k^T$. Here, rows of V_k holds the eigenvalues. These are the coordinates of the individual document vectors. In this way, we get $\vec{d_1}$, $\vec{d_2}$, ..., $\vec{d_n}$ as document vectors for n documents.

b) Generate query vector:
Now we generate the query vector $\vec{q}$ using the Q, U and S matrices by applying the following matrix operations:

$$\vec{q} = Q^TU_kS_k^{-1} \tag{2}$$

2) Apply cosine similarity: Cosine similarity is applied to two non-zero vectors which measure the closeness between them by measuring the angle. In our system, we consider this technique to evaluate the likeness of the query

vector with each document vector. To do this we use the following formula:

$$Sim(\vec{q}, \vec{d}) = \frac{\vec{q}.\vec{d}}{|\vec{q}||\vec{d}|} \quad (3)$$

Thus, we get the similarity score of each document. Here, the angle between the two vectors lies from 0° to 90° i.e. the similarity value is bounded between -1 and +1. The higher value means that the corresponding document is too related to the query than the other documents which carry the lower values. In that way, we rank all the documents according to the higher score to the lower score which helps the readers to choose their expected documents from the rank list.

IV. Experiment and Evaluation

To measure the performance of our system we experiment on 50 documents which are collected from different popular Bangla newspaper – the daily Prothom-alo, the daily Ittefaq, the daily Samakal, etc. We divide them into five groups. Each group contains ten news. To calculate the accuracy we use the precision and recall of this system.

A. Precision

Precision is considered a positive predictive value. It is the fraction value of relevant documents among the retrieved documents. On the other hand, we said that the percentage of selected documents that are corrected. Precision is measured by the following formula:

$$P = \frac{a}{a+b} \quad (4)$$

Where,
P = Precision.
a = Number of selected documents that are correct documents.
b = Number of selected documents that are not correct documents.

B. Recall

Recall represents the sensitivity of the documents. It is the fraction value of relevant documents that have been retrieved over total relevant documents. On the other hand, recall is the percentage of corrected documents that are selected. We can calculate the recall value from the following formula:

$$R = \frac{a}{a+c} \quad (5)$$

Where,
R = Recall.
a = Number of correct documents that are selected.
c = Number of correct documents that are not selected.

C. Accuracy

To measure the performance we now calculate the accuracy of this system. To do this we use the precision and recall value from equation (5) and (6). Now we calculate the accuracy throughout the following formula:

$$Accuracy, A = \frac{2 * P * R}{P + R} \quad (6)$$

As a sample, we show the similarity representation for a group Group-I, depicted in Fig. 3. Here, we observe that there are some negative document scores. The main reason is that when we apply the SVD technique we get some negative values for both the document and the query vectors.

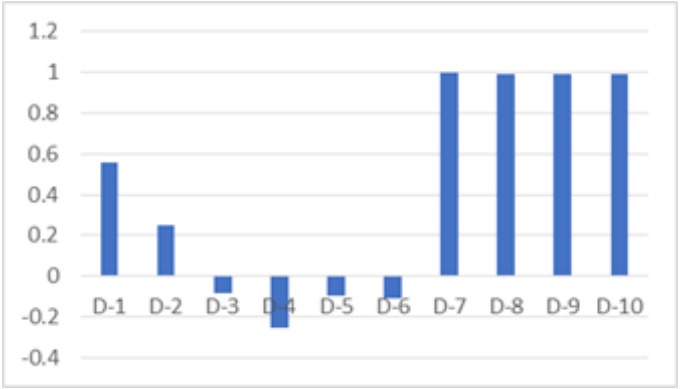


Fig. 3. Similarity representation.

To measure the accuracy, we simulate our data more times and decide the threshold value as 0.4 which mean that if the similarity value is above 0.4 then the document is selected otherwise the document is not selected. Table I gives the precision, recall and accuracy value of all five groups and shows the average value of them which are 86.22%, 92.286%, and 88.557% respectively. These values show that our system is considered as a more reliable system to rank the Bangla documents.

TABLE I. Performance evaluation of the proposed system.

	Precision	**Recall**	**Accuracy**
Group-I	83.333	71.43	76.924
Group-II	77.778	100	87.5
Group-III	100	90	94.737
Group-IV	80	100	88.889
Group-V	90	100	94.737
Average	86.22	92.286	88.557

Our system does not reach into 100% accurate level. This system will fall when the query does not fully or partially match with the searched documents. It has many reasons. If the query string carries an abstract meaning, it will produce incorrect relevancy. Again, a word has a different meaning in different contexts which causes a problem to retrieve the actual documents.

V. Conclusion and Future Direction

In our document ranking system for Bangla language, we have presented a loyal approach which shows the 88.557% accurate result. The main key point of this system is the intermediary topics layer, which is formed, based on the relative importance of the topics in the documents and we discard the less important topics which reduce the time complexity at a high rate without affecting the outcome of the system. This system gets struck when the search keywords are not matched with the documents i.e. if the keywords are contextually used or they have a different meaning in the documents. Therefore, in future, we want to try to resolve these issues with increasing the accuracy rate by linking the web data through the concept of Semantic web.

References

[1] P. Wiemer-Hastings, K. Wiemer-Hastings, and A. Graesser, "Latent semantic analysis," in *Proceedings of the 16th international joint conference on Artificial intelligence*. Citeseer, 2004, pp. 1–14.

[2] S. Deerwester, "Improving information retrieval with latent semantic indexing," 1988.

[3] G. W. Furnas, S. Deerwester, S. T. Dumais, T. K. Landauer, R. A. Harshman, L. A. Streeter, and K. E. Lochbaum, "Information retrieval using a singular value decomposition model of latent semantic structure," in *Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 1988, pp. 465–480.

[4] J. Carbonell and J. Goldstein, "The use of mmr, diversity-based reranking for reordering documents and producing summaries," in *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 1998, pp. 335–336.

[5] X. Shen and C. X. Zhai, "Exploiting query history for document ranking in interactive information retrieval," in *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*. ACM, 2003, pp. 377–378.

[6] G. Zuccon and L. Azzopardi, "Using the quantum probability ranking principle to rank interdependent documents," in *European Conference on Information Retrieval*. Springer, 2010, pp. 357–369.

[7] S. E. Robertson, "The probability ranking principle in ir," *Readings in information retrieval*, pp. 281–286, 1997.

[8] J. Wang and J. Zhu, "Portfolio theory of information retrieval," in *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2009, pp. 115–122.

[9] G. Zuccon, L. Azzopardi, and C. van Rijsbergen, "The interactive prp for diversifying document rankings," in *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. ACM, 2011, pp. 1227–1228.

[10] B. Mitra, E. Nalisnick, N. Craswell, and R. Caruana, "A dual embedding space model for document ranking," *arXiv preprint arXiv:1602.01137*, 2016.

[11] M. T. Islam and S. M. Al Masum, "Bhasa: A corpus-based information retrieval and summariser for bengali text," in *Proceedings of the 7th International Conference on Computer and Information Technology*, 2004.

[12] K. Thakkar and U. Shrawankar, "Test model for text categorization and text summarization," *arXiv preprint arXiv:1305.2831*, 2013.

[13] F. Zou, F. L. Wang, X. Deng, and S. Han, "Automatic identification of chinese stop words," *Research on Computing Science*, vol. 18, pp. 151–162, 2006.

[14] A. N. Chy, M. H. Seddiqui, and S. Das, "Bangla news classification using naive bayes classifier," in *16th Int'l Conf. Computer and Information Technology*. IEEE, 2014, pp. 366–371.

[15] M. H. Seddiqui, A. A. M. Maruf, and A. N. Chy, "Recursive suffix stripping to augment bangla stemmer."

Generating Anime from Real Human Image with Adversarial Training

Shuvendu Roy
dept. of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
bikash.shuvendu@gmail.com

***Abstract*—With growing interest in animation movie, the demand for building an automated system to convert the real video into animation is higher than ever. The frame-by-frame editing of animation generation process is very costly and time-consuming. To help to generate animation quickly and easily in an automated process we proposed a generative model that converts real-world images into corresponding-animation image without losing important details of the source image. We used a variation of the generative adversarial network as underlying architecture with custom loss function to preserve the content of source image which converting to animation image.**

***Index Terms*—Generative model, GAN, CNN, Image-to-animation**

I. Introduction

Animation movies are currently one of the most popular forms of entertainment. Nowadays, the animation has become more realistic and it is not drawn frame by frame by artist rather it is acted by actors then given animation look to the real video to generate the animation. This process lots of expert graphics a lot of time to complete one animation. This is because the animator has to go through the video frame by frame and change the character and scenarios into the animation form. To help this process we have proposed a neural network based conditional image generation model that takes a real-world image and generate a corresponding animation image. Fig. 1 shows an example of a generated animation image.

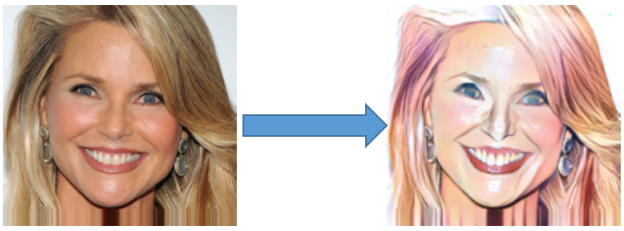


Fig. 1. Generated animated version of source image with the proposed model.

Generating animation from the real image is a conditional image generation task, where the generator model is conditioned on the source image. From another perspective, it is an unsupervised problem. We do not have the corresponding images of real and animation dataset to train a supervised model. To train this model we need a lot of real and animation images. For animation, we used the anime dataset. Anime of a form of Japanese animation which is very popular and many images of this is available. For this, we have used the word animation and anime interchangeably.

The recently proposed model called generative adversarial network(GAN) [1] has shown impressive results in generating images. Our model is built upon the conditional generative adversarial network. It takes the source image as input and generates the target image. To preserve the content of the source image which applying the style of target distribution we have combined generative loss with the content loss. In the experiment, we have shown the effect of each part of the loss and compared the result with other versions of generative models. In this work, we have focused on the dataset of the facial image only but it can be extended to build a general-purpose animation generator by changing the dataset only.

II. Related work

Generative adversarial network(GAN) has shown impressive results in generating data in different domains, especially image. Initially, GAN was used to generate random images of the target domain. Later conditional adversarial network [2] introduced control over the generated image. In conditional settings, GAN is used to generate the next frame of video, produce a photo, image from information or annotation, image inpainting, super-resolution image generation. The generator of GAN is consist of two parts-encoder and decoder. The encoder compresses the data into a low dimension matrix the decoder than generated the image of target domain from it. Generating image from given image is commonly known as image-to-image translation. The first successful image-to-image translation that generates high-resolution images was proposed by Isola et. al. [3] which is known as Pix2pix model. The pix2pix model was a supervised model. Later CycleGAN [4] proposed a method for generating images from one domain to another domain in unsupervised training.

The job of the discriminator in the GAN model is to identify the real image and generated image and help generator generate better images to fool the discriminator. The original GAN proposed a discriminator to output a single real number indicating real or fake signal but later models with high-resolution input were not able to generate good images with

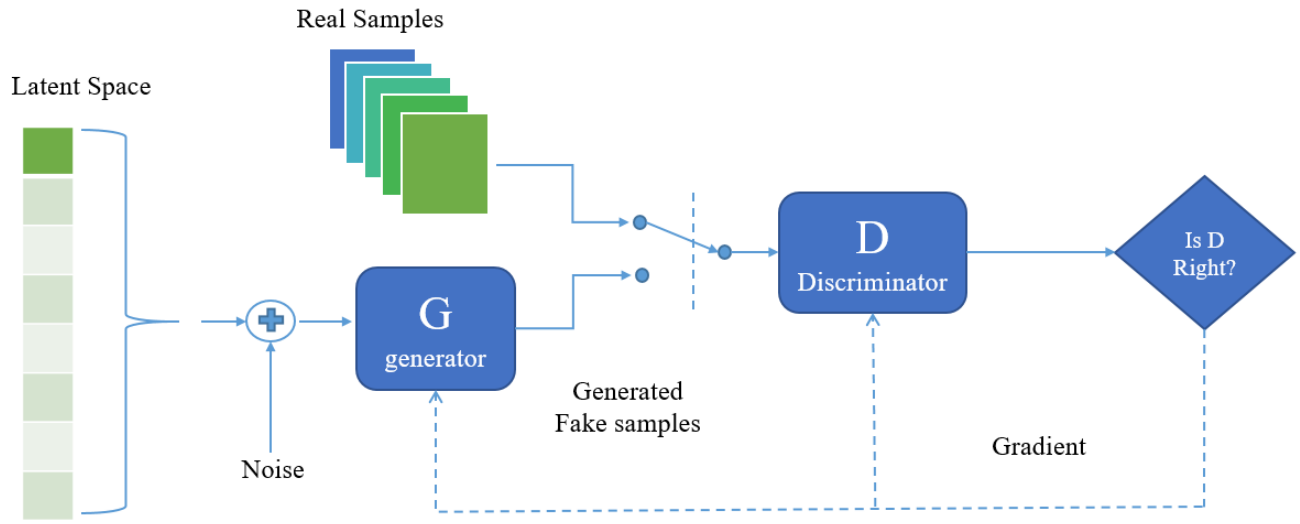


Fig. 2. Training steps of a Generative Adversarial Network.

this. Later PatchGAN [5] proposed outputting a matrix of output by the discriminator rather than single value. The patch size of 4×4 to 16×16 are common. The objective is to get all values one for real image and zero for a fake image.

Our work focuses on generating the animation from real images. Not many works had focused on this topic. Recently Chen et. al. [6] proposed CartoonGAN which focuses on generating cartoon with fewer details or share corners in images. To remove the shape edges which is not present in cartoon images, this model takes special cares by training of edge smoothed cartoon images. Our model focused on generating detail animation images and do not requires such steps.

III. Methods

We have used custom loss function of the recently proposed generative method called Generative Adversarial Network(GAN). GAN is consist of two parts-Generator and Discriminator. The job of the generator is to generate an image and the discriminator tries to separate the real image from the generated image. The model is trained as a two-player min-max game. The GAN loss helps the generator to generate a realistic looking image. In our application, the job of the generator is to generate an image that looks like real Anime. Along with the GAN loss, we have added a content loss to preserve the identity or feature of the source image.

A. Generative Adversarial Network

GAN was proposed by Goodfellow el. al.(2014) [1]. Here the generator and discriminator is trained as a two-player min-max game. Where the generator tries to fool the discriminator by generating a good looking image and the discriminator tries to distinguish the real the synthetic images.

GAN in unconditional settings just tries to generate a good looking real image. In this case, the input to the generator is a random variable draws from a normal distribution. Let x be a source image drawn from the source distribution P_X and r is a random variable from the normal distribution $\mathbb{R}^d$. We define the generator and the discriminator as G and D. The generator takes z as input and generates an image $G(z)$ of the dimension of real image x. The discriminator takes an input of the dimension of the real image and output a real number. It takes real or generated image and tried to say if it is a real image or generated image. In perfect cases, $D(x) = 1$ if $x \sim P_x$ and $D(x) = 0$ if $x \sim P_G$. The generator in the meantime tries to make the output of the discriminator 1. The generator and discriminator are trained at the same time one epoch each with opposite objective. The generator tries to minimize the following cost function while discriminator tries to maximize this.

$$\min_G \max_D V(G, D) = \mathbb{E}_{x \sim P_x}[log\, D(x)] + \mathbb{E}_{z \sim P_z}[log\,(1 - D(G(r)))] \quad (1)$$

Fig. 2 visually describe the GAN training process.

B. Image-to-Anime Generation

The GAN can be used in conditional settings to generate the target image from the source image. This is commonly known as image-to-image translation. Here the generator is given a source image as input. The generator is consist of two parts- Encoder and decoder. The encoder converts the image into some low dimensional space. The decoder than upsample it to the image dimension to generate the target image. The second input to the generator is the random variable, which is combined with the output of the encoder and passed the combined vector to the decoder. This random variable introduces some randomness in generated image and prevents generator to generate exactly the same image every time. The Objective function for cGAN is-

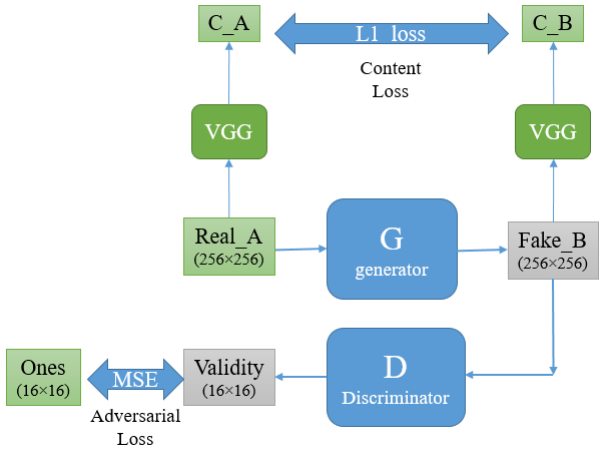


Fig. 3. Training the generator with Adversarial Loss and Content Loss.

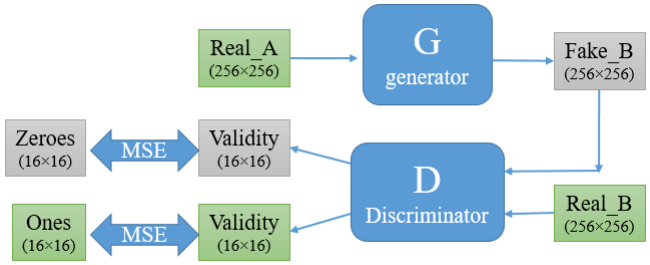


Fig. 4. Training the discriminator.

$$\mathcal{L}_{cGAN} = \mathbb{E}_y[log\, D(y)] + \mathbb{E}_{x,z}[log(1 - D(G(x, r))] \quad (2)$$

It is shown by Isole et. al. [3] that conditioning discriminator with the image helps the training. So, the objective becomes-

$$\mathcal{L}_{cGAN} = \mathbb{E}_{x,y}[log\, D(x, y)] + \mathbb{E}_{x,z}[log(1 - D(x, G(x, r))] \quad (3)$$

Redford et. al. [7] proposed a deep CNN based encoder-decoder architecture for the generator. It generates 256×256 resolution images. We followed the architecture of encoder and decoder from it and added BatchNormalization [8] both is generator and discriminator.

As the network is highly nonlinear, with random initialization, the optimizer sometimes falls into local minimum position. To solve this problem we have adopted a technique from Chen et. al. [6]. It has proposed an initialization phase to help improve the performance of the training. Here, the object of the overall model is to generate an image with the style of the target domain preserving the content of the source image. We start the adversarial training with a generator with the objective of only reconstructing the content of input images. To achieve this we pre-train the model only on the content loss for 20 epochs.

C. Loss Function

The loss function of the model consists of two parts: (1) Adversarial Loss, which drives the generator to match the generated data distribution to the real data distribution, and (2) Content Loss, which makes sure the content of the generated image match that of the source image. It has shown to be useful and helps to preserve the content in previous works [6], [9].

$$\mathcal{L}(G, D) = \mathcal{L}_{adv}(G, D) + \lambda \mathcal{L}_{con}(G, D) \quad (4)$$

Here λ is a hyperparameter that controls the importance to the content of the source and target image. The larger the value of λ is, more the target image preserves the content of the source image. The smaller value of λ results in a more cartoon-like image with less content preserved from the source image. We have conducted experiments with different values of λ.

Adversarial Loss: It is responsible for making the generated image distribution matches the target image distribution. In this application, we did not use any random variable in the generation process. During the generator training, the generator is trained to make the output of discriminator 1. For discriminator training, the discriminator is trained to output 1 for real image and 0 for generated images.

$$\mathcal{L}_{adv} = \mathbb{E}_y[log\, D(y)] + \mathbb{E}_x[log(1 - D(G(x))] \quad (5)$$

Here, x is sampled from the source data distribution, which is real human images. y is sampled from target data distribution, which is anime images. From this final GAN loss function, we have removed the source image as input to the discriminator. Also, rather than outputting as simple real number as validity we have followed the patchGAN [5] technique. It output a matrix of values as the validity. For images of real distribution, all the values of this matrix will be ones and zero otherwise.

Content Loss: It is the difference between the content of the source and the generated image. In this model, we have adopted the high-level feature represented by a VGG network [10] trained with ImageNet [11]. It is proved to have good object preservability. So, two images with the same content should have similar representation by a VGG network. For this experiment, we have used VGG19 for extracting the content of an image. The content loss is the difference between the representation of two images in the VGG network. We define the content loss as-

$$\mathcal{L}_{con}(G, D) = \mathbb{E}_x[||VGG_l(G(x)) - VGG_l(x)||_1] \quad (6)$$

Here, l represent the output of a specific layer of the VGG19 network. In this case, we have used the output of layer $conv4\_4$. We measured the difference in term of l_1 sparse regularization instead of l_2 because the l_2 loss is proved to be generating blurry images.

Fig. 3 and 4 visually shows the training steps of the generator and discriminator.

IV. Experiments

We have implemented the model with pytorch [12] int python. All experiments were done in NVIDIA 1060 GPU.

A. Dataset

As the approach of generating Anime of corresponding human image is an unsupervised learning problem, we need unpaired images of both domains. For the human image, we have used the Large-scale CelebFaces Attributes (CelebA) Dataset [13]. This is a large-scale facial dataset with attributes. It contains 202,599 images of 10,177 different peoples. It also contains 40 attributes of each image that can be used for facial attribute detection, face detection etc. For our application, we only need some plain images.

For images of Anime domain, we have used "Danbooru2017: A Large-Scale Crowdsourced and Tagged Anime Illustration Dataset" [14]. This is a large dataset with over 2.94m images. We have used a subset of this dataset with equal images that of CelebA dataset. All the images are resized to 256×256.



Fig. 5. Few sample image from both training dataset. Top rows contains image from CelebA dataset. Bottom row contains samples from anime dataset.

B. Experimental Setup

The input to the generator and discriminator of the model is 256×256. We have used Adam optimizer [15] for training both networks. It has two momentum parameter to control the learning rate decay-beta1, beta2. We applied image augmentation for images of both domains. Table I summarizes the augmentation settings.

TABLE I
Values of Augmentation Parameters

Type of Augmentation	Values
Rotation	-20 to 20
Horizontal Flip	Yes
Scaling factor	1.1

Other parameters of training the model are summarized in table II.

TABLE II
Values of Training Parameters

Parameters	Values
Image input size	256×256
lambda (λ)	10
Epochs	50
Batch size	8
Adam: beta1	0.5
Adam: beta2	0.999
Learning rate	0.0002
LR Decay rate	0.99

C. Effects of Losses

Our proposed method is trained with the loss function in Eq. 4. It is consist of two losses which is described in Section III-C. Two demonstrate the importance of both component of the loss, we have conducted experiments with only the Content Loss and with only the Adversarial Loss.

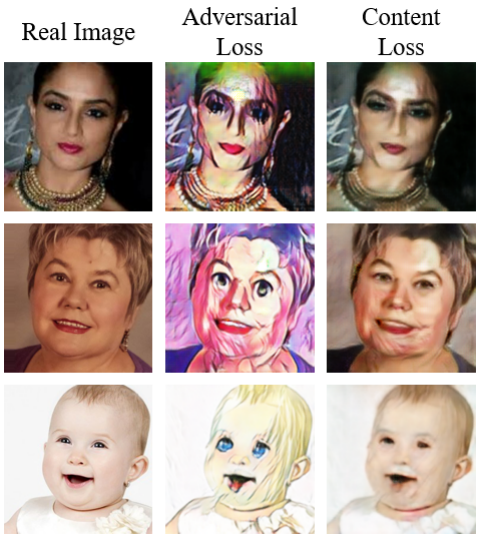


Fig. 6. Effect of only adversarial loss and only content loss.

As we can see from Fig. 6, with the adversarial loss only the generated image does look like target distribution but it losses the content of source image and some image are hard to identify as the source subject. On the other hand, content loss force the target image preserve the content of source image but it does not help to generate images with the target effect of Anime.

D. Evaluation Matrix

Evaluating generated image is an open problem and still has no general purpose solution to evaluate this rather than visual inspection. As we are using the facial image in the generation process, we have followed the evaluation matrix used by Antipov et. al. [16]. They used face recognition score for evaluating the generated image. Here a trained face recognition algorithm is used to verify two images if the FR model thinks people in this two images are the same person.

We used OpenFace [17] face recognition library for this. The reported FR score is the average of the correctly detected face similarities.

E. Result Analysis

We have conducted experiments on a few different variations of models to generate the Anime images. We trained the model with only adversarial loss. It does not keep the content of source images in the generated image, as a result, the FR score for this is very low as though the generated image looks like an anime image. Another model good for generating images with unsupervised training is CycleGAN [4]. This model is computationally expensive as it is a bidirectional model with double the number parameters to train and did not do good enough to preserve the content of source image. Table III below summarizes the results of our experiments.

TABLE III
COMPARISON OF FACIAL RECOGNITION SCORE ON COMPARISON WITH PREVIOUS METHODS

Experiment	***FR Score***
Adversarial Loss [1]	35.33%
Cycle GAN [4]	68.87%
Proposed method	76.24%

Fig. 7 shows some of the generated images of our model.

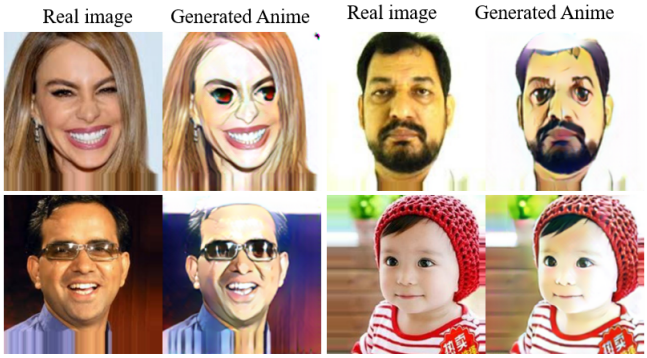


Fig. 7. Few sample image from both training dataset. Top rows contains image from CelebA dataset. Bottom row contains samples from anime dataset.

V. CONCLUSION

Generating animation from the real image with cartoon look is very important in the movie and animation industry. This is a very hard and complex job for the animator to create this animation frame by frame. In this process, the human images look too much like the real world and don't provide the cartoon-like look in this animation. To bring such look or deal with a non-existing character the animator has to go through each frame and edit. In this work, we have proposed a neural network based animation generation model that can create a nice looking animated image with preserving the content of the source image. This work will help the industry to generate animation faster and efficiently and help further research in this direction.

REFERENCES

[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[2] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.

[3] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *CVPR*, 2017.

[4] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networkss," in *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2017.

[5] C. Li and M. Wand, "Precomputed real-time texture synthesis with markovian generative adversarial networks," in *European Conference on Computer Vision*. Springer, 2016, pp. 702–716.

[6] Y. Chen, Y.-K. Lai, and Y.-J. Liu, "Cartoongan: Generative adversarial networks for photo cartoonization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 9465–9474.

[7] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

[8] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[9] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. P. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network." in *CVPR*, vol. 2, no. 3, 2017, p. 4.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[11] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[12] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," 2017.

[13] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep learning face attributes in the wild," in *Proceedings of International Conference on Computer Vision (ICCV)*, 2015.

[14] G. B. A. G. Anonymous, the Danbooru community, "Danbooru2017: A large-scale crowdsourced and tagged anime illustration dataset," January 2018, accessed: DATE.

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[16] G. Antipov, M. Baccouche, and J.-L. Dugelay, "Face aging with conditional generative adversarial networks," in *Image Processing (ICIP), 2017 IEEE International Conference on*. IEEE, 2017, pp. 2089–2093.

[17] B. Amos, B. Ludwiczuk, and M. Satyanarayanan, "Openface: A general-purpose face recognition library with mobile applications," CMU-CS-16-118, CMU School of Computer Science, Tech. Rep., 2016.

Forecast Upcoming Traffic States by Exploiting Nearest Junctions and Big Data

[1] Md. Akter Hossain, Mohammed Nazim Uddin [2]
[1] Department of Computer Science & Engineering, East Delta University, Chattogram, Bangladesh
[2] School of Science, Engineering and Technology, East Delta University, Chattogram, Bangladesh
myselfakterhossain@gmail.com [1], *nazim@eastdelta.edu.bd* [2]

***Abstract*—Like other developing country, traffic jam is one of the common problems for Bangladesh. Every day almost a hundred thousand vehicles roaming and creating a serious traffic jam in the metropolitans. In this manner, it becomes impossible for an inhabitant to reach his/her destination on time. In this paper, we have introduced a scheme, named FTCP (Future Traffic Condition Predictor). This system has carried out to examine traffic conditions by vehicles measurement and the main theme of the study is to find out that factors which describe traffic problems by exploiting big data concept. We have categorized each key factor in different classes and each class has different traffic causes along with time durations of self-solving. When a traffic seeker is requesting for the information for a specific junction by any electronic devices, the system has acquired data information in favor of two ways. The first acquisition will be completing by collects previously recorded data and on the other, current data information of target junction via images and date-times. Here, previously recorded data were analyzed according to system define format and image samples were estimates by through FCN (Fully Convolutional Network). At last, experiments all of the major factor by vector analytical method. The experimental simulation was achieved promising accuracy in terms of using sophisticated datasets.**

Keywords—vehicles, traffic jam, traffic rules, image processing;

I. INTRODUCTION

Real-time road traffic update is highly desirable both in developed and developing countries. Nowadays, providing accurate information on road traffic conditions and situations is the most challenging part of such area of research. A various study has been done by ITS (Intelligent Transport Systems) that performed quite fast and impressive [1, 2]. However, there are still many problems that need to solve and also make traffic management as comfortable and effective as possible. In every year developing countries like Bangladesh losing message amount of money, property and intellectual life due to traffic congestion and accident. Recently, a study had conducted by ARI (Accident Research Institute) of BUET (Bangladesh University of Engineering and Technology), where researchers mentioned around 1.2 million peoples of Bangladesh have died and nearly 50 million Bangladeshi were extremely injured or permanently disabled with significant losses of resources and properties by road traffic accidents that's creates enormous burden on not only society but also economy.

Citizens of Bangladesh especially locals of Dhaka and Chattogram cities constantly being angry and frustrated about the unbearable problems of traffic jam and air pollution. This thing not only occurring heavy damage of our time, health and livelihood but also affects earnings. People from subaltern economic area bear their heavy loss due to the excessive jam. Therefore, if it could possible for a citizen to get exact time-table of traffic congestion and duration of resolve then he/she would able to make their daily working plan more effectively and efficiently by taking advantage of alternative route or strategy.

In the study [3], they had introduced a forecasting algorithm based on pheromone communication model. Congestion-forecasting technology is one of the main elements of ITS. The two main methods of congestion classified mostly in many studies, long-term forecasting is a method of statistically analyzing past traffic data, and discovering a pattern where congestion has occurred. And, short-term forecasting of congestion which is a few minutes ahead by using real-time information. It could effectively make forecasts under regular-congestion conditions though that had originated from car and road situations, a large amount of past data is necessary for analysis. Besides, it has weaknesses in forecasting under irregular-congestion conditions, for example, traffic congestion that experienced from the Christmas-holiday season in the U.S.A or during the Golden-week holidays in Japan or Bengali New Year (Pôhela Boishakh) holidays in Bangladesh. Some studies used machine learning classification algorithm such as random forest to build a congestion prediction model [4]. As input variables, they had considered weather conditions, time period, special conditions of the road, road quality and holiday. Although the random forest algorithm has the characteristics of high robustness, high performance, and high practicability, the study generalized some errors explaining less applicability to the prediction of the congested condition. And, in this study [5], they had shown a report where constant civilization growth will bring an additional load of around 3 billion vehicles to road networks by 2050 that will certainly lead to increased air pollution concerns, highly congested roads.

In this study, we present a novel method for traffic condition forecasting in Bangladesh based on a machine learning approach. The proposed approach employs current data and previously recorded information for the forecast as the data source simultaneously. Firstly, the system going to acquire real-time data information from a target and nearest junctions through high definition images from finest different angles (standard decision: 0^0, 180^0, 270^0, 360^0) and current date-time statistics. Then, segmenting images to carry out assessment data of traffic transports (number of and type of transports) by applying FCN [6]. Secondly, on the other way, it will gather past several years (for this study, it was 10 years) traffic count data sets for these particular junctions and after analyzing previous traffic patterns with current traffic situation, a logistic relationship will be formed. Based on this relationship, the system makes an educated guess (prediction) on the traffic condition of the future. Both FCN and MAD are widely cited machine learning algorithms and successfully applied in many areas of science, engineering and vehicular technology.

Our contributions are explaining in follows:

- We propose a distinguish examination of road traffic, ensuring the efficiency of recognizing traffic situation for future based on enormous data analysis.
- Although, there is an abundant number of datasets of road traffic, there are lacking in terms of accuracy, efficiency and some sensitivity issues. Many researchers had been working with real-life data of traffic images, but, most of them were previously recorded [4, 7]. However, our system gathers data information based on the immediate request of the user. Besides, another deficiency was to deal with imbalance data cause of million sources. In this manner, we have collected all of the data items by following some specific data variables, which is also can be defined as the system data format.
- In the end, the system is able to generate a complete evaluation report, which had covered every scenario of our examines.

This paper has been organized as follow. Section II provides a brief review of the previous works. Section III describes the proposed approach for traffic congestion prediction and Section IV presents the experimental setup. Section V describes the evaluation methods and presents the results. Finally, Section VI concludes the paper.

II. Related Work

There is various great investigation of traffic condition improvement. That's why it's become an active area of research and innovation. In this section, some previous works been reviews in the relevant literature.

Nguyen et al. [7] proposed a traffic system for congestion monitoring, which includes data collection, the establishment of segmented structure, traffic-flow modeling, local segment traffic congestion prediction, and origin-destination traffic congestion service for drivers. They used Fuzzy rules as local segment traffic congestion prediction in their proposed approach to determining the traffic congestion state and concluded that neural network (NN) have a statistically significant impact on the field of traffic congestion.

Yan [8] employed the support vector machine (SVM) to build a method of a short-term traffic condition prediction. As they mentioned, their method can effectively deal with the negative effects of outliers that inevitably exist in collected traffic data, thus enhance the robustness of the model and improve the generalization ability. Based on simulation results, it was shown that the classification method improved better accuracy than other machine learning methods.

Jayaratne al. [9] studied the relationship of urban speed-flow in Sri Lanka. Their study was targeted at filling the gaps of urban four-lane roads with the speed-flow model. In that way, they had described the PCU (Passenger Car Unit) tool and the simulation result shown a theoretical capacity report. They found that the nature of the traffic flow is unendingly changing by the data of 2481 PCU/h/lane for 5-minute interval data.

Anggraini et al. [10] developed a new approach to resolve the lack of getting the right information problems. They design visualization of traffic conditions in the form of a map, where the information obtained from Twitter, CCTV, user and operator PRFM (the radio information about traffic conditions in Bandung). Then, another predictive system was designed to get traffic conditions based on information from Twitter, so people can predict traffic conditions for one week ahead. However, when this study was compared with NNs, the calculations showed the performance much faster.

Wan et al. [11] developed a new perspective on traffic video analysis. They have had presented a total of three contributions to their work. First, they had shown the definition of congestion by proposing a unique state of traffic. Secondly, a precise labeling method was introduced to get the ground truth of the congestion level accurately. Finally, Inverse Perspective Mapping (IPM) based algorithm was proposed to analyze the traffic situation. However, their system was becoming slow for IPM mapping and system results acknowledge the lack of greater data sets.

In studies [12-14], the authors illustrate the advantage of uses of NNs (Neural Networks) in various sectors. For example, rice yields prediction has been done by perceptron NN along with weather predicting [12]. Meanwhile, a recommendation system has made to predict movie for user satisfaction via NN [13]. Moreover, A fraud detection engine has been conducted for recognizing credit card fraud activity, which builds by the help of backpropagation feed-forward NN [14]. In this study, also, we have used sorts of NN that specialized to segment particular image objects.

The above review confirms that historical data has a significant impact on the traffic forecasting and employing this sort of data in huge amount may result from a significant improvement in prediction accuracy [e.g. 7]. In this study, we further spread our approach for traffic forecasting in a different way. Contrast to the previous studies that used current data to predict traffic condition, we use both current and previous data sets as big data concept.

III. Future Traffic Condition Predictor (FTCP)

In this prototype, FTCP (Future Traffic Condition Predictor), we have been forecasting the traffic conditions on the basis of different variables of traffic measurement (see Fig.1). A mobile application has been made to provide user to request for traffic condition report for the specific junction. Therefore, based on data requester's GPS (Global Positioning System) location, the system is able to track user junction point, which helps the system to compare two aspects: (a) distance different between seeker junction to target junction and (b) determining all of road junctions as nearest junction of target one. There are also two criteria, one is measuring from Target junction and another computing from Requester junction. From the target intersection, we have measured nearest junctions by applying Bellman Ford's algorithm where the shortest path contains at most edges. On the other hand, all intersections also are selected among Requester and Target. In this way, figure 2 demonstrates this point of view as a general model of the selection process of nearest junctions where two possibilities are analysis via "P-1 Junction B" and "P-2 Junction-B" route (could be more for another junction). Then, on depending of the shortest distance, the direct route will be selected and considering all of the intersection as the nearest junction between target and source junction

In order to do so, the first and foremost thing is acquiring each parameter values of real-time traffic statistics via images from the target and adjacent junctions which's collected or, captured from standard angles along with data-time information. Image data variables: $img_{total}^{k} = [img\_0^{360}, img\_1^{270}, img\_2^{180}, img\_3^{0}]$, where k is the number of capture objects of a junction and date variables: $dt_{img\_0}^{n} = [dt_{img\_0}, dt_{img\_1}, dt_{img\_2}, dt_{img\_3}]$, where n is the define a number of image items. Now, categories information

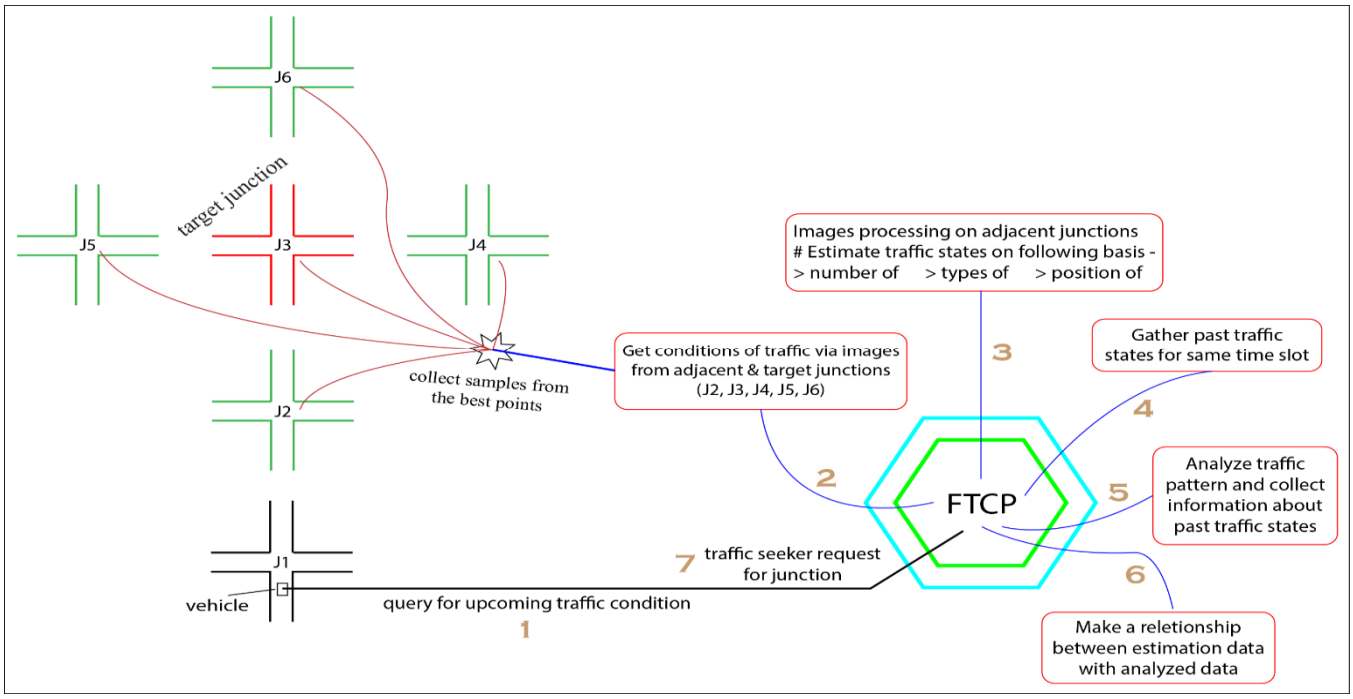


Fig. 1. FTCP System Prototype and Operational Overview.

according to the number of vehicles, types of vehicles and the position of vehicles. By mean of position is to ensure that transports are following road lane or not. It will help our system to figure out the calculation time according to measuring Signal Cycle Length (SCL) when the traffic signal is green. Then, the estimation process of image segmentation had done by using FCN NNs also known as ConvNet [15]. After that, the system will start folding data sets that gathered from the previously recorded database (see Table I) according to our system specified format. For this study, we accumulate data on about 10 years from several associations of the government of the People's Republic of Bangladesh.

Before examining the outcomes, SVM (Support Vector Machine) analytical tool have used to figure a relationship between experimental data and assessment data. As we define, we had used two sorts of datasets: (a) previous historical data and (b) currently collected data.

Finally, the system will calculate a robust measure of the variability of a univariate sample of road traffic quantitative data that explained a systematic empirical investigation of traffic condition observable via statistical statement. This process has been done by exploiting MAD (Median Absolute Deviation) method and softmax/sigmoid function.

A. Fully Convolutional Network for Image Segmentation

As 3-dimensional array layer ($x \times y \times z$) segmentation of FNN or Convnet, the first layer has considered by image samples with pixel size, x, y and d panels. From this section to the next phase known as receptive fields of path connection. Convnets are literally arranged with two common procedure: sampling and pooling by the aid of an activation function like other ANNs. The real-life implementations of these have done by CNNs (Convolutional Neural Networks). There is a subsection of machine learning technology that is inspired by biological neurons, deep neural network. It's also a subfield of ANNs (Artificial Neural Networks). CNNs is special sorts of deep NNs along with feed-forward ANNs combinedly, which means feed-forward referred to as multi-layer NNs. That defines, it's working right through the model and there is no feedback mechanism which output feedback

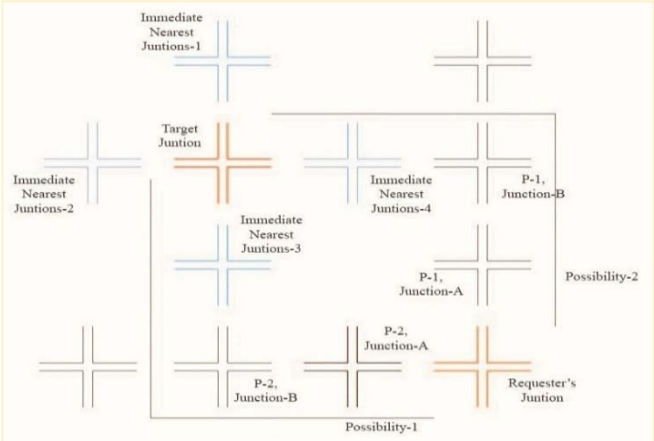


Fig. 2. Selection Procedure of Adjacent Junctions.

model comes backward by itself. The biological visual cortex has some sensitive specific visual areas that contain cells.

Figure 3 depicted the construction process of FNNs by exploiting one of our experimental data substances. This sample is collected via a mobile camera which has the standard rage of resolution. After the image leveling procedure complete, one of the segmented items picks up for the convolution process, where multiple possibilities calculated by mean of Matrix. During this process, calculated specific neurons that are connected with receptive regions of input. In the input volume, weights and a receptive field considered as a dot and then the best extraction small dot matrix go for the sampling process.

The main idea of the subsampling process is to get representation data by removing its dimensions that are actually helpful to rid of the overfitting problem. This technique is also known as pooling (max), which is able to select the maximum pixel value based on its size. The convolution and sampling process are occurring several times to detect the best-fitted data spot respectively.

Another objective of FNNs to acquiring high-range features that have learned by combining its all features.

Finally, bypassing those features to the layer of output. Then, we have used softmax or normalized exponential function to make the educational guess of input class.

In the learning process, after setup, the network by assigning input nodes, weighted (w) and bias (θ) parameter values define randomly. Then, the network function calculated as equation (1) for each neuron along with activation function (2).

$$\sum_i^n w_i \times \theta_i \tag{1}$$

$$f = (\sum_i^n w_i \times \theta_i) \tag{2}$$

B. Analytical Vector Regression

SVM (Support Vector Machine) is a supervised regression model involved with machine learning methods. Regression analysis is considered a kind of statistical process used for the accounted relationship between variables. As a notation, the following parameters are considered in SVM:

- Obscure variables are denoted as the β symbol.
- Independent variables as Y sign.
- Dependent variables as X.

By equation (3), a regression model is represented, where X and β is a function of Y. Furthermore, SVM regression means values calculated by (4) and (5) equations.

$$Y \equiv f(X, \beta) \tag{3}$$

$$Subside: {}^{1}/_{2}\,(|w^2|) \tag{4}$$

$$Related\ to: \begin{matrix} y_i - (w, \Delta_i) - \theta \leq \varphi \\ (w, \Delta_i) + \theta - y_i \leq \varphi \end{matrix} \tag{5}$$

Here, the threshold value defines by y_i and φ with the training symbol, Δ_i. In our study, input parameters considering by through scalability ration of time series, where ultimately equation is formed like (6).

$$Y = \beta_0 + \Big[\big(S_{curr_{img1}}, S_{curr_{img2}}, \dots S_{curr_{imgN}}\big) \times \big(S_{curr_{dt1}}, S_{curr_{dt2}}, \dots S_{curr_{dtN}}\big) + \big(S_{his_{img1}}, S_{his_{img2}}, \dots S_{his_{imgN}}\big) \times \big(S_{his}, S_{his_{dt2}}, \dots S_{his_{dtN}}\big)\Big] \tag{6}$$

The hyperplane formula (7) that differentiate the negative class to the positive class. And, bias update with the value of offset by (8).

$$\tilde{w} = \sum_i \beta_i S_i \tag{7}$$

$$Y = wX + \theta \tag{8}$$

IV. Design of Experiments

A. Simulation Design and Datasets

Table I illustrates all of the sources of our datasets, that publicly available records of the Government of Bangladesh. Here, some datasets are collected from Bangladesh Bureau of Statistics (BBS), which is a centralized official bureau in Bangladesh for collecting statistics on demographics. This agency is established with two other statistical organization by the Bangladesh government in 1974. In this way, it provides various statistical analysis report such as agriculture statistics, health, labor-employment, road accident, income, expenditure & poverty survey, ICI, climates GDP, trade and many more. And, another source was Transparency International Bangladesh (TIB). It's a Berlin-based International Transparency branch in Bangladesh, that's a dedicated civil society against corruption. Then, the final sources come by the annual traffic report generated various times by Dhaka Metro Police. These datasets have their own formats, for our study, we have collected and merged our required data sets on the following basis:

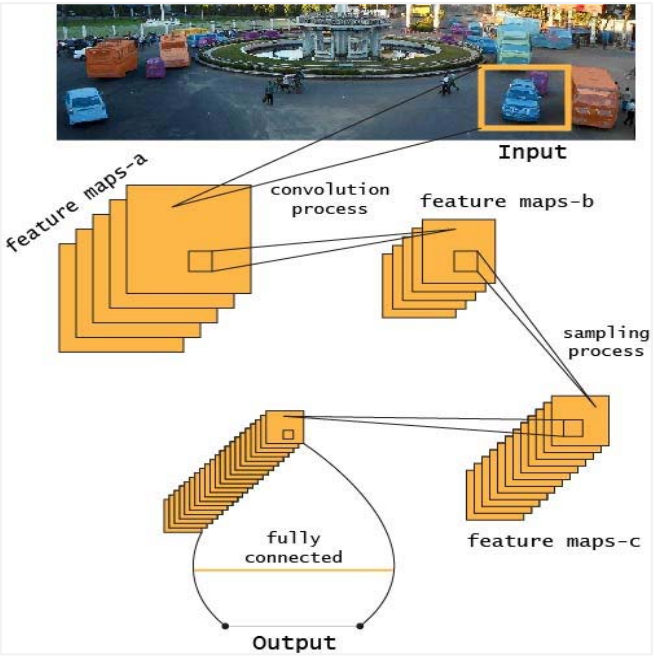


Fig. 3. The Architecture of Fully Convolutional Neural Networks.

- ***No. of transports:*** We have measured vehicles numbers of the particular junction. It helps us to calculate the estimation of time probability for the next junction.
- ***Type of transports:*** We used information of categories vehicles on the basis of the report of "Manual Classified Traffic Counts" of Bangladesh Ministry of Communications of Roads and Highways Department. The Vehicle Identification Sheet shows 15 different lists of vehicles type described based on 3 categories.
    a) Trucks and Buses.
    b) Light Motorized Vehicles.
    c) Non-Motorized Vehicles.
- ***Transports position:*** Traffic signal cycle lengths have a significant impact on the quality of the urban area. Hence, we have calculated position lengths by the following equations. In context, probabilistic cycle length calculates by (9) and TRRL model (10), which for the optimal minimum delay cycle length.

$$c_0 = \frac{1.5 \times L + 5}{1 - Y} \tag{9}$$

Where c_0 represents the optimal cycle length (sec); L represents the total lost time (sec), and Y represents the sum of the critical flow ratio of all phases.

$$C = \frac{L}{1 - [\min(CS, RS)/RS]} \tag{10}$$

Where C represents the cycle length (sec); L represents the total lost time (sec/cycle); CS

TABLE I. Data Sources

Ref.	DATA SOURCE	YEAR
[14]	Statistical Yearbook of Bangladesh 2015-2017 Statistical Yearbook of Bangladesh 2011-2014	6
[15]	Road Accident and Casualties Statistics 2009-2016	8
[16]	TIB Report on Bangladesh Road Transport Authority 2016-2017 TIB Report on Bangladesh Road Transport Authority 2015-2016 TIB Report on Bangladesh Road Transport Authority 2014-2015 TIB Report on Bangladesh Road Transport Authority 2013-2014 TIB Report on Bangladesh Road Transport Authority 2012-2013 TIB Report on Bangladesh Road Transport Authority 2011-2012	6
[17]	Dhaka Metropolitan Police - Traffic Accident Analysis - Annual Report 1996, 1998, 2001, 2002, 2004, 2006, 2009, 2010, 2015	9

represents the sum of the critical phase traffic volumes (veh/h); RS represents the reference sum flow rate $(1710 \times PHF \times fa)$, (veh/h); PHF represents the peak-hour factor; and fa represents the area type adjustment factor [0.9 if the central business district and 1.0 otherwise].

B. Methods of Evaluation

The evaluation process has been done by using two kinds of estimation methods: (a) first one is a statistical method, which we used for measuring average absolute deviation between actual and predicted values. The univariate variable (in this case, transports parameters data) have considered as quantitative data that refer to the similar parameter of MAD calculation. In this way, a univariate data set subject to real date-time define as: $T_{dt_k} = [T_{dt_1}, T_{dt_2}, \dots, T_{dt_n}]$ and average mean calculate by following equation (11).

$$MAD = (||T_{dt_k} - \overline{T_{dt_k}}||) \tag{11}$$

We have to exploit MAD as median similar in order to the consistent estimator, which the main objective is to use it as standard deviation, σ. Since we have acquired data values from images, we did measure scale factor in order to find out relational distributed data by through standard deviation formulation (12).

$$\sigma = q \times MAD \tag{12}$$

Where scalar factor, q is $\frac{1}{(\vartheta^{-1} \times 3/4)} \approx 1.4826$, ϑ denoted as quantile function.

(b) Another method, we had exploited to measure our outcome is a normalized exponential function also refereed as Sigmoid or SoftMax function. We used this method as a neural network-based classifier of the final layer of NNs. It's a probability distributor between two set of events that estimate a mean number of bits (also called cross-entropy). Over a given event sets, the cross-entropy could be defined as follows (13).

$$E_{(x,y)} = E_x(\log y) = E_x + D_{KL}(x||y) \tag{13}$$

Where E_x is entropy of x and $D_{KL}(x||y)$ is the *kullback–leibler divergence* of x and y. In this case, cross-entropy is calculated using (14) equation.

$$E_{(T,y)} = -\sum_{k=1}^{N} \times \frac{1}{N} \times log_{2 \times} y(e_k) \tag{14}$$

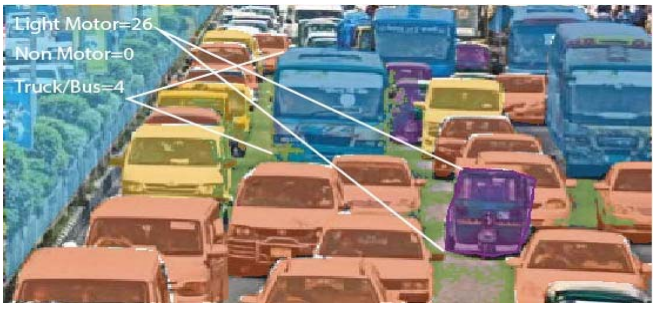


Fig. 4. One Sample Segmented Images of Chattogram Traffic.

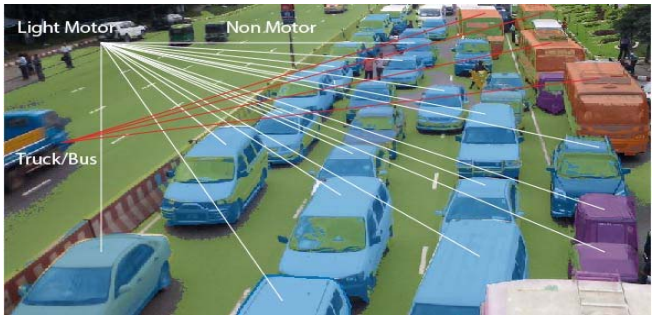


Fig. 5. One Sample Segmented Images of Dhaka Traffic.

Here, N is the test set size and the probability of the event, e calculated from a training set, T, which denoted as $y(e_k)$.

V. Results Examine

FCN is a normally consider as CNN, where the last fully connected layer is substituted by another convolution layer with a large receptive field. Figure 4 and 5 show some sample results of the segmented image, where color leveled on the vehicle based on our vehicle classification notations. There are two sorts of segmentation way applied, Edge-based and Point-based. For example, in fig. 4, some heavy and light motor-based vehicle have been selected by two portions: the upper and bottom edge. As opposed to, selecting the labeled color motors. In this study, we have found the edge-based technique much efficient than the other. It also shows a minimum error margin in term of separation of same types of motor vehicle.

By fig. 6, we showed the state of training result, which explained the fitness of our data sets by thought gradient number. It's showing, the average gradient we achieved, was almost 3.5 after 17 loops of training. The algorithm had used, "Scaled Conjugate Gradient (SCG)". It's a supervised learning algorithm for feedforward neural networks and is a member of the class of conjugate gradient methods.

Tables II and III illustrate two case studies of evaluation outcomes respectively. Here, accuracy values referred to the efficiency acknowledgment of results. Moreover, we have shown (see Fig. 7) a comparison depiction regarding other opponents. A compared method is actually a mapping data of space topology that classification leads to best accuracy with the variance networks [20].

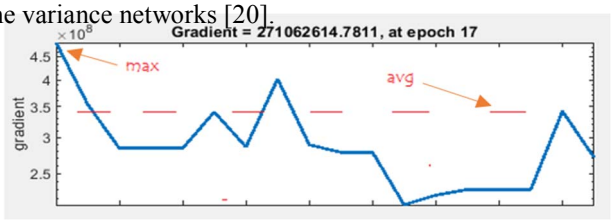


Fig. 6. Training States Outcome of Applied Datasets.

TABLE II. FIRST CASE STUDY REPORT

Min. Confidence	MAD	Entropy, *E*	Precision
0.1	3.9214	0.243	0.087
0.2	3.475	0.243	0.003
0.3	3.193	1.128	0.019
0.4	3.745	0.544	0.033
0.5	3.784	0.499	0.011

TABLE III. SECOND CASE STUDY REPORT

Min. Confidence	MAD	Entropy, *E*	Precision
0.1	2.453	2.498	0.009
0.2	2.421	3.587	0.019
0.3	2.798	1.457	0.090
0.4	3.787	2.353	0.074
0.5	1.457	2.221	0.038

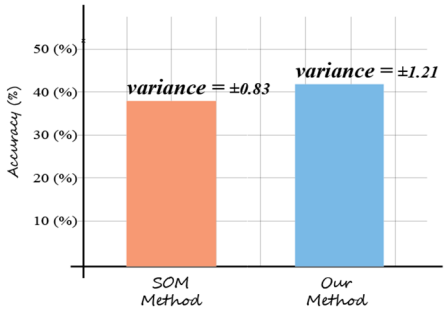


Fig. 7. Proposed Method Comparison Graph

Meanwhile, we have also calculated our proposed system by running multiple times where around 30 subject to the accuracy level. If we notice fig. 8 of our accuracy model, in which we have shown the performance criteria of our work. Here, we have presented the epoch values in five sequences after 5 loops. In this way, our accuracy level gradually increased to 86% after a slight downward on around 75%.

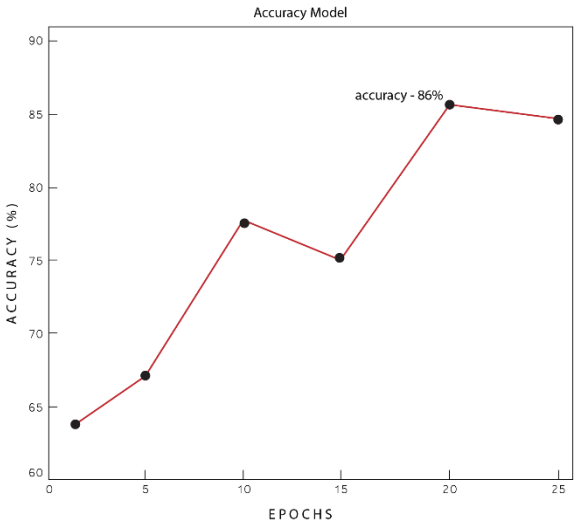


Fig. 8. Accuracy Model Generated By Matlab Statistics

VI. CONCLUSION

We have presented a new approach that predicts traffic conditions for the future by through relationship between two sorts of datasets. Our approach analysis both data sets according to time sequence that determined groups of traffic states with similar patterns. We found that working with a maximum amount of data sets was beneficial. Also, we found that the proposed ConvNet neural network was real affection for images classification or segmentation.

REFERENCES

[1] A. Kamaraj, K. Radha, M. Priyanka and M. Punitha, "Intelligent transport system using integrated GPS optimized reader," 2016 Second International Conference on Science Technology Engineering and Management (ICONSTEM), Chennai, 2016, pp. 332-336.

[2] N. Kavitha and D. N. Chandrappa, "Comparative Review on Video Based Vehicular Traffic Data Collection for Intelligent Transport System," 2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT), Bangalore, 2017, pp. 260-264.

[3] S. Kurihara, H. Tamaki, M. Numao, J. Yano, K. Kagawa and T. Morita, "Traffic congestion forecasting based on pheromone communication model for intelligent transport systems," 2009 IEEE Congress on Evolutionary Computation, Trondheim, 2009, pp. 2879-2884.

[4] Y. Liu and H. Wu, "Prediction of Road Traffic Congestion Based on Random Forest," 2017 10th International Symposium on Computational Intelligence and Design (ISCID), Hang-zhou, 2017, pp. 361-364.

[5] S. Djahel, A. Jones, Y. Hadjadj-Aoul and A. Khokhar, "CRITIC: A cognitive radio inspired road traffic congestion reduction solution," 2018 Wireless Days (WD), Dubai, 2018, pp. 151-157.

[6] E. Shelhamer, J. Long and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 4, pp. 640-651, 1 April 2017.

[7] Duc-Binh Nguyen, Chyi-Ren Dow, and Shiow-Fen Hwang, "An Efficient Traffic Congestion Monitoring System on Internet of Vehicles," Wireless Communications and Mobile Computing, vol. 2018, Article ID 9136813, 17 pages, 2018.

[8] H. Yan and D. Yu, "Short-Term traffic condition prediction of urban road network based on improved SVM," 2017 International Smart Cities Conference (ISC2), Wuxi, 2017, pp. 1-2.

[9] N. Jayaratne, C. Rathnayake and H. Pasindu, "Evaluating the Speed-Flow Relationship of Urban Fourlane Roads under Heterogeneous Traffic Conditions," 2018 Moratuwa Engineering Research Conference (MERCon), Moratuwa, 2018, pp. 455-459.

[10] D. Anggraini, W. Siswantoko, D. Henriyan, Devie Pratama Subiyanti, Mochamad Vicky Ghani Aziz and A. S. Prihatmanto, "Design and implementation of system prediction and traffic conditions visualization in two dimensional map (case study: Bandung city)," 2016 6th International Conference on System Engineering and Technology (ICSET), Bandung, 2016, pp. 87-91.

[11] J. Wan, Y. Yuan and Q. Wang, "Traffic congestion analysis: A new Perspective," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, 2017, pp. 1398-1402.

[12] M. A. Hossain, M. N. Uddin, M. A. Hossain and Y. M. Jang, "Predicting rice yield for Bangladesh by exploiting weather conditions," 2017 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, 2017, pp. 589-594.

[13] M. A. Hossain and M. N. Uddin, "A Neural Engine for Movie Recommendation System," 2018 4th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), Dhaka, 2018.

[14] M. A. Hossain and M. N. Uddin, "A Differentiate Analysis for Credit Card Fraud Detection," 2018 International Conference on Innovations in Science, Engineering and Technology (ICISET), Chattogram, 2018.

[15] A. Karpathy, [Online], Available at: <URL: github.com/karpathy/convnetjs>, ConvNetJS: CNNs using JavaScript under MIT license, [Last accessed September 8, 2018].

[16] Statistical Yearbook Report, [Online], Available at: <URL: www.bbs.gov.bd/site/page/29855dc1-f2b4-4dc0-9073-f692361112da/Statistical-Yearbook>, [Last accessed September 8, 2018].

[17] Road Accident and Casualties Statistics, [Online], Available at: <URL: da-ta.gov.bd/dataset/road-accident-and-casualties-statistics-2009-2016>, [Last accessed Sep-tember 8, 2018].

[18] TIB Report, [Online], Available at: <URL: www.ti-bangladesh.org/beta3/index.php/en/communications/reports/research-report>, [Last accessed September 8, 2018].

[19] DMP Traffic Analysis Report, [Online], Available at: <URL: www.rhd.gov.bd/DocDatabase/ViewRecords_View.asp>, [Last accessed September 8, 2018].

[20] Sam Maes, Karl Tuyls, Bram Vanschoenwinkel, Bernard Manderick, "Credit card fraud detection using bayesian and neural networks", International Naiso Congress on Neuro Fuzzy Technology, 2002.

Pest-Inspector: An Insect Detection System by Learning Data Representations

Md. Akter Hossain
Department of Computer Science and Engineering
East Delta University
Chattogram, Bangladesh
myselfakterhossain@gmail.com

A.H.M Tareq
Department of Computer Science and Engineering
East Delta University
Chattogram, Bangladesh
ahmtareq05@gmail.com

Mohammed Nazim Uddin
School of Science, Engineering and Technology
East Delta University
Chattogram, Bangladesh
nazim@eastdelta.edu.bd

***Abstract*—In this paper, we have introduced a beetle detection system of cultivation field of Bangladesh, which will recognize both harmful and useful insects on the basis of the real-time image. Based on this estimating analysis, the system is going to predict upcoming status and production quantities of target crops land. Our system has established on two scenarios, learning & classification and recall process. In the learning process, we have categorized two types of insect data items: Useful and Harmful and subclassified various types of harmful beetle in order to acquaint them by the help of ANN (Artificial Neural Network) terminology. Then, we have stored those classified datasets into a database repository. In next phase, the detection system will segment/cluster all of the insects from real life images and synchronize them with our pre-categorized data storage apropos to figure some educational guesses of the situation of agriculture ground and the ratio of the number of crop yields. Here, Fully Convolutional Network (FCN) had applied as a detection tool to estimates image samples. Finally, a regression technique, SVR (Support Vector Regression) has exploits to predict the desirable outcomes. An effective outcome has been presenting by through experimental study.**

Keywords—artificial neural network; support vector machine; image segmentation; insect detection; means absolute error; mean relative error;

I. INTRODUCTION

Agriculture production means not just produce crops, it is a technique for producing food, feed, fiber and many more. There are many more things and each has its own way of production methods. It's one of the major practices that have a serious impact on the services of the ecosystem. These eco-services are surrounded by a wide area of soil retention, pollination, water quality, carbon sequestration, biodiversity conservation, and nutrient cycling. By practicing ecosystem, agricultural productivity could far beneficial to consider it as the main point of good environmental change. This is a complex methodology relationship between agricultural practices and ecosystem impact. An ecosystem is a precise area of interaction among climate, landscape, animals, and plants. It's basically included physical and chemical components (e.g., soils, water and plants) to bacteria. In this way, people become part of ecosystems, mostly health and wellbeing of populations of the human. Also, moths are a small part of both agriculture and ecosystem, where it's very crucial aspects of the agricultural study.

In the study [1, 2], the authors had described beetles by categorizing them in gender, where they had expressed the behavioral pattern of both male and female moths. Generally, male insects are searching partners by observing females plume of its source. There are several pests, they had introduced such as pink bollworm, the tomato pinworm and, fruit moth. Moreover, they have explained the sex pheromones impacts using species-specific behavior-modifying chemicals in agriculture. Also, this study had monitored the existence and absence of insects by pheromones (hundreds of amount) and other semi-chemicals, that has been discovered to projects ecosystem against farmland beetles. There have been enormous studies done on insect classification. However, there are a few past works of literature shows by using image acquisition and features extract algorithms. Meanwhile, some of the previous methods were based on insect specimens [3-7].

In this study, we have demonstrated a differentiating process for insect detecting of farmland in Bangladesh based on the machine learning approach. Our proposed approach employs real-time data information that has collected from high-dimensional images of the different region of farmland. In this way, we have acquired around 2351 samples that contains various sizes of image. Here, we have used different types of cameras (DSLR, mobile camera) to collect our required images. That's why images contain different resolution along with dissimilar sizes. Therefore, by using MATLAB, we have resampled them into exact same resolution where pixel set on 1024×768 depth. In this way, less than half a percent of images had transformed into fixed sizes. This system has been categories into two according to our objectives. Firstly, the system going to acquire real-time data image and extract information about various sorts of moths. Secondly, the system will be learning these features through the help of Perceptron neural network. Important note, two types of beetle's data has considered – Harmful and Useful moths. As machine learning algorithms, ANN is widely used neural network technique and effectively applied to learn anything. Meanwhile, FCN also extensively applied image processing method in many areas of science, engineering, and technology.

Our contributions are explaining in follows:

- We propose a novel examination in insect detection, ensuring the efficiency of recognizing farmland's beetles and the future situation of production ratio based on cognitive data analysis.
- We have used datasets that had collected by ourselves (with the aid of high-resolution digital camera). Because of the issues of elusive and sensitivity, these types of authentic data sets has serious lacking to gather tangibly. Therefore, we had gathered our required data items (images) from different district of farmland around Bangladesh. Also, data diversity was another challenge with our datasets. We have resolved this sort of crux by using several times of collaborative filtering, where system collaborate data associated with the study [8]. In this study, the authors have explained the various features and characteristics of many harmful and useful beetles. Moreover, we did several academic surveys on some particular places that related to farmland beetle cultivation.

- Furthermore, based on our experimental results, we have also estimated the future situation of target cropland and possibilities ratio of production quantities.

The section has organized as following, where Section II builds on some literary works. And, Section III will have explained our proposed method. Section IV and V will be describing estimation methods and outcomes respectively. At last, the study concluded in Section VI.

II. Related Work

In this section, some prior experiments had been reviews in the relevant literature such as pest control and ANN.

In Study [9], where Zhong et al. proposed a flying insects recognition system with vision-based counting. Also, it's a classification system that had made on the basis of various aspects. First, we had designed a vessel for the surveillance camera to collect real-time images. After that, YOLO (You Only Look Once) object detection, which is also a course counting method and SVM classification techniques had exploited to bring out global features. At last, both insect counting and recognition system applied on Raspberry PI. Six species to major the system efficiency. For insect count, they had used a portable predefine counting system, where our proposed system is able to figure the counting by itself (with the system own counting class).

Hossain et. al. [10], authors had predicted future crops production associated with weather-based datasets, which were predicted before and uses them as inputs. By this experiment, they have shown a good work of ANN perceptron method. The presented approach (WPSRY) built a weather-based model and have got an educational guess of upcoming weather condition. Then, by using these data as parameter input values with the agricultural data and simulated for target production quantities of rice yields.

Meanwhile, in this study [11], they have built a neural engine for movie recommendation system by exploiting user's behavioral patterns. Here, two types of data were used to make a class database, which has constructed based on user interest. Authors had presented a brilliant work with Hopfield Neural Network (HNN) that were sequentially transformed the combination of primitive and user interaction data into a class dataset.

In Izumi et. al. [12], A behavior model had been built to specify the fire and gas leakage position at disaster time by using the beetle pheromone activity. Two outcomes have characterized to find out the equation of the relation of the gas and the results of a semiconductor gas sensor. By using a portable robot equipped with a gas sensor (semiconductor), they had performed discovering gas sources and simulation outcomes shows the effectiveness of the applied system. This study illustrates how important to extract behavior pattern for any simulation.

In a study [13], they had introduced a pest detection system based on insect shape, where the image processing method was exploited to extracts shapes features in the sugarcane crops field. Furthermore, a Sobel edge detection technique had used for image segmentation in opposite to background. With 9 geometric shape's features learning, the system has the ability to detect by several angles: circle, oval, rectangle and triangle angles. Authors have explained that image classification or segmentation is quite important for a study that related to farmland simulation.

Kiskin et. al. [14] had proposed a mosquito detection system using neural networks. Like our study, they also had used Convolution Neural Network (CNN), which was exploits to operate audio wavelet transformation by investigation of the network's power (predicted) and visual statistics. This statistic visualization helps the system to observe a deep insight of components missing data detection problem. In this context, they have applied CNN on another purpose however we have used another method of CNN, which is FCN for image classification and segmentation to clustered insects based on our objectives.

ANN is an effective method to study any kind of real activity or any behavioral pattern. A recent study has been published that had shown the importance of ANN to learning user's interests. In this study [15], a banking credit card fraud detection system (FDE) has been presenting along with the user's interests selection tool (FST). During transaction procedure, FDE has deployed to find out any sorts of intruders with investigating credit card holder's behavioral activity. FST is the tool that has the ability to observing user's operational movements and try to find out any unexpected activity.

Nguyen et. al. [16], introduced 3D insect detection and diagnostics system in high-throughput color ratio using a robotic-based arm to beetle control process. They had used light field cameras to gather images in depth one shot along with the aid of 3D web tools on purposes of annotation, visualization and distribution. In our study, we have also used high-resolution based digital images as our datasets. However, we didn't use any filtering tools to improve visualizations.

In the study [17], they have presented that a mobile phone can be low-cost equipment for mosquito surveillance. After studying mosquito species in medically, they had defined quantitatively acoustic supports by spatiotemporal metadata, non-invasive species and enable rapid proofs. Finally, a novel paradigm had been demonstrated for mosquito surveillance.

Grieve et. al. [18], they have shown how several multispectral sensors of the image could be effective to detect real-time whitefly insects. Simulation experiments had conducted with four types of viral vectors on both cassava and aubergine plant leaves, where outcomes indicate an efficient classification by case result among 85% to 95%. This study has described the effectiveness of using sensors images on insect detection experiments. However, it's compulsory to use image filtering tools with this method for disturbance of pixels.

By Study [19], authors have introduced a monitoring system that will constantly be observing farmland by video image processing and cloud computing combinedly with the help of robotics concepts. Here, the proposed detection system works on one of the popular fruits (tomato) farmland. Although, video processing is more effective than image segmentation but many more things are challenging to acquire efficient datasets. For instance, in video surveillance, many image frames become a disturbance for speed and resolution qualities. In this manner, highly efficient filtering tools required, which is sometimes could be a time-consuming process.

Also, Santiago et. al. [20] have proposed both detection and monitoring system for farmland insects using ANN and Mel-frequency cepstral coefficients (MFCCs). Here, the ANN method had used for classification and MFCCs as feature extraction tool. The main focus of this study was

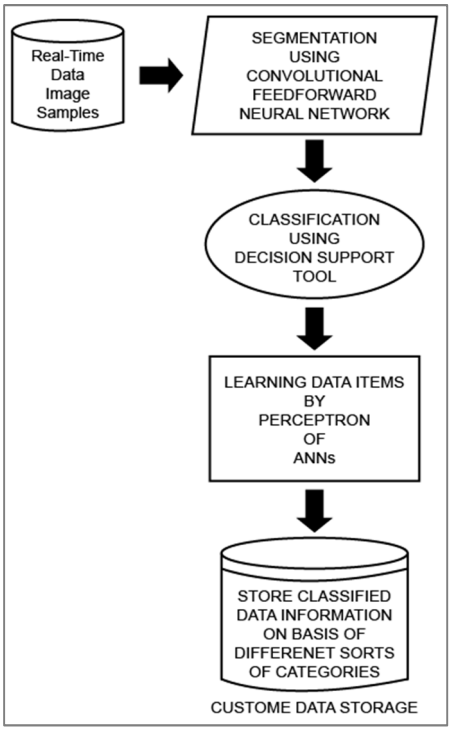


Fig. 1. System Learning Procedure

further experiments on acoustic methods of pest management. Moreover, this study had used Sitophilus oryzae (L.) sounds, which commonly known as rice weevil along with five different sensors, which will provide effective support to ANN for recognizing insects sounds.

The above review confirms that the neural network is a commonly used method for the study like insect detection or crops yields prediction [e.g., 10, 14, 16]. In this study, we further expand our approach by using useful insect. After some academic survey, we have noticed some beetles are helping the crops land to increase production quantity.

III. Proposed System (Pest-Inspector)

In this prototype, the first and foremost was designing a learning process. In order to do so, we had collected information (extract features) about several insects related to croplands by analyzing real-time segmented images (using the FCN image classification method). Here, we had extracted shapes and color feature of beetle on the basis of different angles and deepness. Both are the most challenging operation causes of similar reasons for multiple different insects. For example, according to study [8], ground beetle such Carabidae and rove beetle like Staphylinidae has various shapes similarity and both have the same color as well. On the contrary, Ladybirds and Pupa have exact shapes similarity, so, in here, color features help the system to figure out the variances. To solve these sorts of problems, we have exploited a Decision Support System (DSS) that combine with both decision tree and neural networks [21]. By using decision tree integration and backpropagation multilayer feed-forward algorithms, this method had developed to figure out a clinical decision for Eye Disease Diagnosis, which was quite related to our objectives. They had applied it on multiple sorts of syndromes of eye disease. Moreover, the uses of Object-Oriented methodology along with Object-Oriented Programming (OOP) that makes this system far effective with more than 88% success rates. Next, by perceptron multilayer feed-forward neural network, we have learned all the features to the system as per two types of insects' categories: USEFUL and HARMFUL beetles (See Fig. 1). Finally, we have stored all learning data information into data storage (p_db) similar to system requirements.

Also, for recall operation of our system, we have exploited FCN image classification and segmentation method. As per figure 2, the system has the ability to detect farmland insects keep in similarity according to our custom database (p_db). If the operational aspects of beetle symptoms have matched with our predefine prominences of insects than the next phase will start. In the following stage, system will acquire given information according to series of data information ($f = f\_1, f\_2, f\_3, \ldots\ldots, f\_n$).

At last, the system will calculate a vigorous measure based on succeeding requirement characteristics, (a) Insects genre, (b) Insects amount – (b1) Number of useful and (b2) Number of harmful. And then, SVM regression has used to predict actual insects and the require data features by making a relationship of both primitive and real-time data. Furthermore, on basis of these outcomes, we have also measured an educated guess on two future aspects of targeted farmland, where we will measure the status of cropland during a specific time of cultivation and possibilities of the number of production quantities.

Ultimately, the evaluation procedures have done by exploiting three popular estimation methods: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Relative Error (MRE).

A. Used of Custom Decision Technique

In a study [21] where Kabari el. at. had presented a Decision Tree-Neuro based model for clinical DSS of eye disease diagnosis. This model has made on basis of DSS tree and ANN feed-forward techniques combinedly with the help of OOP programming framework. We picked this method as a support for our model because of the following reasons:

- As we know, working with the DSS system environment is very competitive and stressful as well. Meanwhile, other problems like overburden information and its disturbance throughout the scope affect the potentiality of a system. Therefore, we

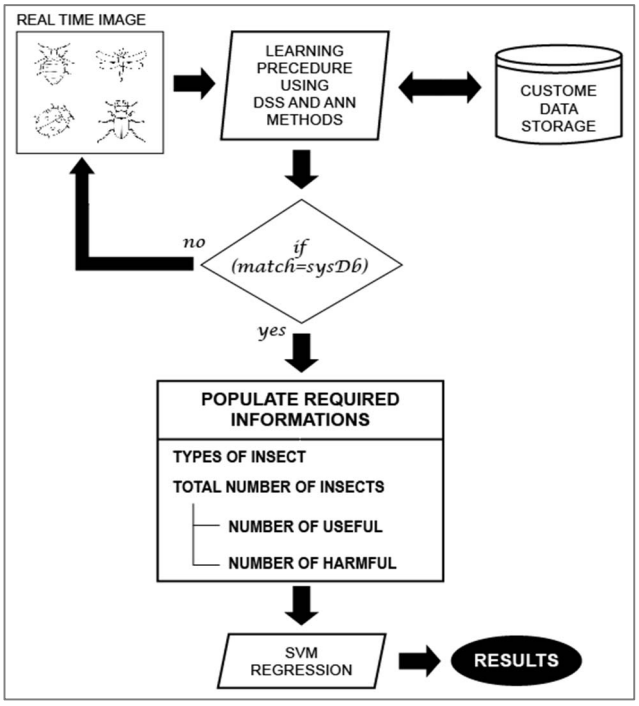


Fig. 2. System Recall Procedure

found this model more effective than other traditional DSS systems especially the uses of the ANN method associated with it

- The next reason is that this model already showed the effectiveness by through applying in diverse fields such as medical diagnosis. So, as we have explained in Sec. 1 that our datasets are diversity based (see some sample data in fig. 3). Therefore, after our analysis we have found, this hybrid approach is quite similar according to our objectives.
- The final reason was, the advantage of using the system that integrates with multiple techniques may get far better-predated results to compare to exploiting primitive method or any separate techniques.

B. *Artificial Neural Network - Perceptron*

Figure 4 illustrates the learning model that we have built for our system. This model has constructed with n input neurons for both color and shape features, U hidden neurons and V output neurons. In here, U hidden ensembles have chosen best parameter value and each and every parameter initialize by some randomly define weights value.

Perceptron ANN is one of the supervised learning algorithms which is presented by vector mathematics. Just like other feed-forward neural networks, its work on moving forward way from inputs section without maintaining any loops/cycles. There are two sorts of Perceptron ANN: Single and Multi-layer. We have exploited multi-layer feed-forward Perceptron, which is design to interconnect with the extra hidden layer.

According to the problem and its solution technique, we have set up the ANN network and its connections along with randomly initialized weights (W_{ij}) values and threshold (β) range. And, a calculative learning rate (δ). Based on our desire objectives, we did prepare the expected target model (T_j) to run the verification. Here, following formula (1) helps to calculate network function by binary outcome series as per (2) notation.

$$\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij} \times \beta \tag{1}$$

$$Y_j = \begin{bmatrix} 1 & net > 0 \\ & if \\ 0 & net \leq 0 \end{bmatrix} \tag{2}$$

Then, if this notation $(T_j - Y_j) \neq 0$ is satisfied, we are going to update connection weights and bias by following the equations (3) and (4) respectively.

$$\Delta W_{ij} = \propto (T_j - Y_i) X_i \tag{3}$$

$$\Delta \beta_j = -\propto (T_j - Y_i) \tag{4}$$



Fig. 3. Sample Images Collected From Different Farmland Region.

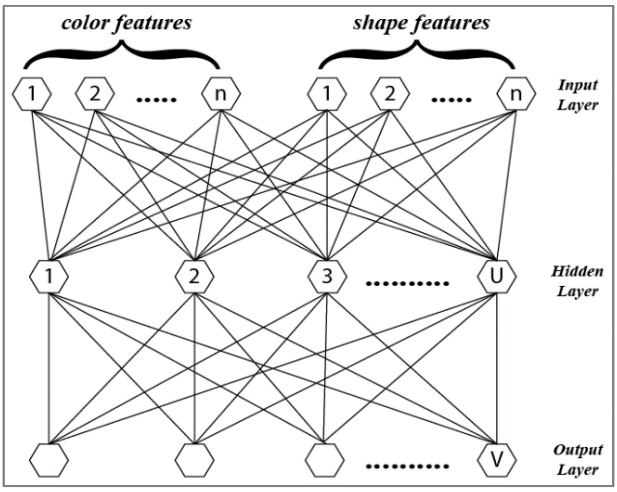


Fig. 4. Constructed ANN Model for Our System.

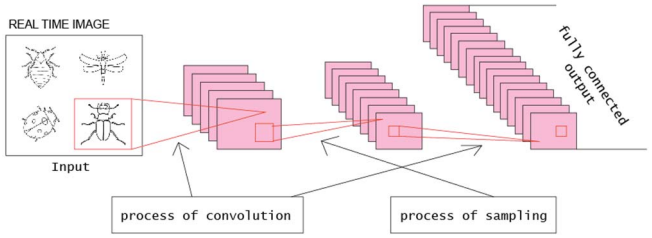


Fig. 5. Operational Procedure of FNN Convolutional Method.

Where Δ indicate the new value and is $\propto$ denoted for learning rate value. Next, after replacing old values with new, run the above operation again until following equation (5) fulfilled.

$$T_j - Y_i == 0 \tag{5}$$

C. *Convolutional Neural Network – Fully CNN Method*

As 3-dimensional array layer ($x \times y \times z$) segmentation of FNN or Convnet (see fig. 5), the first layer has considered by image samples with pixel size, x, y and d panels. From this section to the next phase known as receptive fields of path connection. Convnets are literally arranging with two common procedure: sampling and pooling by the aid of an activation function. After the multiple numbers of sampling and pooling, (until satisfied the condition) processes, input data vector (X) of a particular location (i, j) have calculated for desire output (Y_{ij}) by (6) as per study [23].

$$Y_{ij} = f_{ks}(\{X_{si+\delta i, sj+\delta j}\} 0 \leq \delta i, \delta j < k) \tag{6}$$

Where k is denoted for the kernel size, s for subsampling factor and f_{ks} determine the type of layer. And, the transformation rules (7) used when a computation has occurred with a non-linear function, which is also defined as deep filter or FNN.

$$f_{ks}\sigma g_{k`s^1} = (f\sigma g)_{k`+(k-1)s`,ss`} \tag{7}$$

IV. Evaluation Methods

A. *Applied Datasets*

By fig. 3 we have shown some sample data images of our work. As we explained before, as our dataset considered as diversity based, so, before applying CovNet operation, our system has analyzed each of them through scaling procedure where we have data of measurement that applied on CNN

```
%% Acquire reference image
ref_vid_img = snapshot(hCamera);

%% Quantize images and outputing to the screen

frames = 2000;
for i = 1:frames

% Acquire an image from the webcam
    vid_img = snapshot(hCamera);
```

Fig. 6. Operation Screenshot of Collection of Image Datasets.

technique. Meanwhile, as for color dimension, blue-gray ($b^{\emptyset}g^{\emptyset}$) color combination helps the system (CNN CovNet) discriminate the differs among them. All of these operations had conducted on MATLAB by using its own Image Processing Toolbox. It is a high-level language based interactive environment for visualization programming. It also includes predefine mathematical functions (for instance, Matrix.) to operating visualizing image data such as scaling. By through fig. 6, we have shown an operational snapshot while our system was performing the collected image data where reference data acquired into an array variable.

B. *Results Analysis*

Test efforts are not some pre-defined timeline setting but it is an approximation, which is a value used for certain objectives even though input data could be unstable. Estimation is based on several things such as historical data, both past and present knowledge, and assumptions. And, accuracy is a determination of the ratio of reality. However, some important factors always effect on every estimation: (a) input data, (b) the calculative steps, (c) in which way or how closely the previously collected data calibrate.

For estimating the accuracy of continuous parameters, both MAE and RMSE are two most popular evaluation metrics uses by many various studies. MAE estimator is counting the average magnitude errors of a series of predictions. MAE is the upgrade equation of both Mean absolute Scaled Error (MASE) and the Mean Squared Error (MSE) by the mean of using selected performance issues. It will become a slight change if we don't consider absolute value. Then, it will be estimating the average bias model as Mean Bias Error (MBE). The given equation (8) states the MAE error estimation, where Mean Magnitude of Relative Error (MRE) have using to find out the difference of MAE outcomes in percentage (9).

$$MAE = \frac{1}{n}\sum_{i=0}^{n}\left|Y_i - \hat{Y}_i\right| \quad (8)$$

$$MRE = \frac{1}{n}\sum_{i=0}^{n}\left|Y_i - \hat{Y}_i\right| \times 100\% \quad (9)$$

Where Y_i and $\hat{Y}_i$ denoted as actual and predicted data outcomes of the number of n selected objectives. RMSE is a quadratic rule for measuring the mean magnitude error (define by formula (10)).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=0}^{n}\left(Y_i - \hat{Y}_i\right)^2} \quad (10)$$

There are several differences between MAE and RMSE. Here, in RMSE, all errors are being square before figure out the average values. That's why we have to provide a high amount of weight values on larger errors. In this way, for undesired maximum errors, it could be more efficient.

V. RESULTS

The estimation results show by Table I, where the accuracy and training time describe our approach achieved satisfied accuracy. Here, 13 best exertions have clarified, where the average accuracy of sensitivity standard 91.73%, it has shown the maximum edge of the efficiency of our method. Also, the training time took less than 5 minutes for every best effort along with the exact performance of the operation. If we notice our precession chart (see Table I), our approach has gained more than 90% of actual and overall 8.5% MRE performance ratio which indicates the effectiveness and efficiency of the applied system.

According to our precision outcome, the marginal productivity has plotted (see fig. 7) by MVP (Marginal Value Product) associated with the probability ratio of upcoming production quantity. MVP measure is increasing when the number of useful insects exits on land. In the left-side, MVP margin scale indicates and right-side vertical axis presents the level of quantity. MVP margin specifies the state of cropland.

Moreover, we have analyzed our approach by comparing it with another method. Figure 8 illustrates a clear compares chart along with MRE consequences. The technique [23], we have exploited for comparing, was built by National Academy of Agricultural Science (NAAS) which is a research platform in the fields of crop & animal husbandry, fisheries & agro-forestry founded of Government of India. So, if we notice the comparative analysis between our and

TABLE I. OVERALL PRECISION RESULTS

Endeavor	RMSE	MAE	MRE (%)	Exactness (%)	Training time (min)
1	0.0423	0.465	5.20%	89.24%	3.0
2	0.8756	0.876	6.23%	91.25%	3.5
3	0.6378	0.832	6.80%	91.11%	2.5
4	0.5378	0.543	4.98%	88.79%	3.0
5	0.8634	0.358	5.96%	90.21%	2.5
6	0.7565	0.754	5.88%	89.25%	2.9
7	0.6962	0.700	6.68%	85.16%	3.1
8	0.7562	0.711	7.01%	90.97%	4.0
9	0.8612	0.810	5.12%	92.00%	4.0
10	0.9200	0.829	8.91%	91.12%	3.9
11	0.5663	0.555	5.98%	84.21%	3.4
12	0.8923	0.858	8.29%	87.15%	2.4
13	0.7236	0.711	7.15%	79.15%	3.0

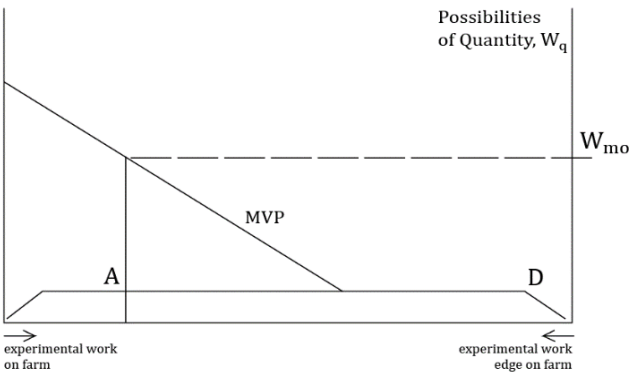


Fig. 7. MVP Response of Experimental Farmland.

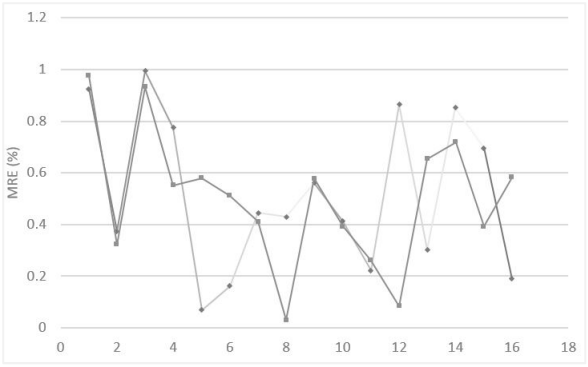


Fig. 8. Comparison Of Analysis Outcomes

another method, we can see the amount first 4 epoch, performance was average with it. Afterward, the accuracy level suddenly drops for the compared approach in minimum ratio. On the other hand, although our technique dropped on 8 epoch but it was going upward quickly in next 2 epoch at 0.6. However, the compared method staying on average MRE accuracy compare to our performance criteria.

Furthermore, we have constructed another comparison overview with previous studies [13, 14] that have already explained in section II. Here, the study [13] had exploited image processing approach to detect beetle shapes where images were segmented via Sobel based edge technique. In our study, also we have exploited image shape segmentation by using MatLab scaling tool. And, in [14], the authors explained mosquito detection operation via Artificial Neural Network. So, as we explained, CNN also an ANN concept which specifically focuses on image processing and segmentation. In this way, these two different methods have achieved fewer performance values compare to our approach. As we notice, method [13] had reached on 7.6% and [14] on 8.09% MRE where our technique had maximum MRE 9.15%.

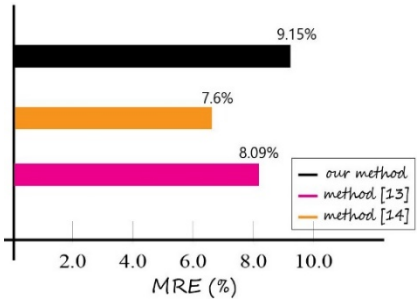


Fig. 9. Compare Results With Another Two Approaches.

VI. Conclusion

We have introduced a novel approach to insect detection for any farmland. However, our method also has capabilities to working on any study like beetle detection. Our main objective is, estimating the future condition of cropland based on the oppression of both harmful and useful insects.

References

[1] R. T. Carde, A. K. Minks, Control of moth pests by mating disruption: successes and constraints, Annual review of entomology 40 (1) (1995) 559–585.

[2] P. Witzgall, P. Kirsch, A. Cork, Sex pheromones and their impact on pest management, Journal of chemical ecology 36 (1) (2010) 80–100.

[3] S.-H. Kang, S.-H. Song, S.-H. Lee, Identification of butterfly species with a single neural network system, Journal of Asia-Pacific Entomology 15 (3) (2012) 431–435.

[4] S.-H. Kang, J.-H. Cho, S.-H. Lee, Identification of butterfly based on their shapes when viewed from different angles using an artificial neural network, Journal of Asia-Pacific Entomology 17 (2) (2014) 143–149.

[5] T. Arbuckle, S. Schroder, V. Steinhage, D. Wittmann, Biodiversity informatics in action: identification and monitoring of bee species using abis, in: Proc. 15th Int. Symp. Informatics for Environmental Protection, Vol. 1, Citeseer, 2001, pp. 425–430.

[6] P. Weeks, M. O¢¢Neill, K. Gaston, I. Gauld, Automating insect identification: exploring the limitations of a prototype system, Journal of Applied Entomology 123 (1) (1999) 1–8.

[7] J. Wang, C. Lin, L. Ji, A. Liang, A new automatic identification system of insect images at the order level, *Knowledge-Based Systems 33* (2012) 102–110.

[8] Beneficials on farmland: identification and management guidelines, [Online], Available at: <URL: *http://adlib.everysite.co.uk/adlib/defra/content.aspx?id=1QQUSGMWSS.0KK6K5GK5A81AJ*>, [Last accessed November 28, 2018].

[9] Zhong, Yuanhong et al. "A Vision-Based Counting and Recognition System for Flying Insects in Intelligent Agriculture" *Sensors (Basel, Switzerland)* vol. 18,5 1489. 9 May. 2018, doi:10.3390/s18051489.

[10] M. A. Hossain, M. N. Uddin, M. A. Hossain and Y. M. Jang, "Predicting rice yield for Bangladesh by exploiting weather conditions," *2017 International Conference on Infor-mation and Communication Technology Convergence (ICTC)*, Jeju, 2017, pp. 589-594.

[11] M. A. Hossain and M. N. Uddin, "A Neural Engine for Movie Recommendation System," *2018 4th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, Dhaka, 2018.

[12] K. Izumi, S. Ishibashi and T. Tsujimura, "Exploring of gas source location using behavioral model of insects," *2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE),* Kanazawa, 2017, pp. 180-181.

[13] K. Thenmozhi and U. S. Reddy, "Image processing techniques for insect shape detection in field crops," *2017 International Conference on Inventive Computing and Informatics (ICICI)*, Coimbatore, 2017, pp. 699-704.

[14] I. Kiskin, B. P. Orozco, T. Windebank, D. Zilli, M. Sinka, Kathy Willis, and S. Roberts, " Mosquito Detection with Neural Networks: The Buzz of Deep Learning," *Computing Research Repository (CoRR)*, 2017, volumn. abs/1705.05180.

[15] M. A. Hossain and M. N. Uddin, "A Differentiate Analysis for Credit Card Fraud Detection," *2018 International Conference on Innovations in Science, Engineering and Technology (ICISET)*, Chattogram, 2018.

[16] C. Nguyen, M. Adcock, S. Anderson, D. Lovell, N. Fisher and J. La Salle, "Towards High-Throughput 3D Insect Capture for Species Discovery and Diagnostics," *2017 IEEE 13th International Conference on e-Science (e-Science)*, Auckland, 2017, pp. 559-560.

[17] H. Mukundarajan, F. J. H. Hol, E. A. Castillo, C. Newby and M. Prakash, "Using mobile phones as acoustic sensors for high-throughput mosquito surveillance," *eLife 2017;6:e27854*, 2017, volume. 6, pages. e27854, doi: 10.7554/eLife.27854.

[18] B. Grieve, C. Veys, J. Dingle, J. Colvin and J. Nwezeobi, "Portable, in-field, multispectral imaging sensor for real-time detection of insect viral-vectors," *2017 IEEE SENSORS, Glasgow*, 2017, pp. 1-3.

[19] S. R. Rupanagudi, Ranjani B. S., P. Nagaraj, V. G. Bhat and Thippeswamy G, "A novel cloud computing based smart farming system for early detection of borer insects in tomatoes," *2015 International Conference on Communication, Information & Computing Technology (ICCICT),* Mumbai, 2015, pp. 1-6.

[20] R. M. C. Santiago, S. L. Rabano, R. K. D. Billones, E. J. Calilung, E. Sybingco and E. P. Dadios, "Insect detection and monitoring in stored grains using MFCCs and artificial neural network," *TENCON 2017 - 2017 IEEE Region 10 Conference*, Penang, 2017, pp. 2542-2547.

[21] L. G. Kabari, and E. O. Nwachukwu, "Decision Support System Using Decision Tree and Neural Networks," Computer Engineering and Intelligent Systems www.iiste.org ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online), Vol.4, No.7, 2013.

[22] E. Shelhamer, J. Long and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 4, pp. 640-651, 1 April 2017.

[23] Veena T, "Insect detection methods," in Indian Institute of Crop Processing Technology, FOOD PROCESS ENGINEERING, Graduate Student, Vol. 33, No. 3, July-September 2015.

Kinect 3D Reconstruction Fusion Mapping with C# Algorithm By RGB-D Sensor

Ashikur Jhalak[1], *Shoumik Rahman Mehedy*[2*], *Rashedur M. Rahman*[3**]
Dept. of Computer Science and Engineering, Brac University, Dhaka, Bangladesh[1]
Dept. of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh[2*,3**]
Email : ashikurjhalak@gmail.com[1]; *shoumik.rahman@northouth.edu*[2**]; *rashedur.rahman@northsouth.edu*[3**]

***Abstract*—The fundamental objective of this paper is to describe the applications of Microsoft Kinect sensor with robust algorithm and solution to present 3-D surface reconstruction mapping with different measurements. In the line of Sensor operations, there could be two possible categories into 3-D reconstructions such as exorbitant laser scanner and other is by taking photos from various angles. In this paper, we have used indoor environmental scenes by depth sensor camera Microsoft Kinect. Data Pre-operations, sensor posture evaluations, computational fusion of the depth sensor data and 3-D surface extractions are the four major part of this experiment. The algorithm of this paper has been built uniquely with C# programming.**

***Index Terms*—Microsoft Kinect sensor, 3-D reconstruction, Depth- sensor camera, image processing, Algorithms.**

1. INTRODUCTION

Microsoft Kinect sensor for X-box 360 has infrared projector and RGB camera inside the Kinect. Red, Green and Blue color defines the object distance measurement and a visual imagination. Kinect depth sensor has a wide range from 1.6 ft. to around 17 ft. and can generate up to 30 (FPS) frames/ Sec. Depth Camera generally collects the videos from RGB camera and measure the depth distance from the object to it [1-2]. The four major steps of this paper are Data pre-processing or surface measurement, sensor posture evaluation, rationalize the model and surface extraction for 3-D surface reconstructions. The graph plotting and other algorithms has been visualized through MATLAB. Our Robust Algorithm has been built with C# programming.

2. LITERATURE REVIEW

Using image processing techniques has been introduced that how Microsoft Kinect Sensor can be cost-effective to extract 3D surface with a volumetric method that fuses entire depth data using the ICP algorithm [1]. An inexpensive 3D camera like Kinect has the ability to quantify topographical surfaces with an accuracy using a digital elevation [2]. Depth data can attain surface reconstruction by enhancing ray casting of the color texture of reconstruction objects [3]. 3D reconstruction can assist a robotic vehicle to navigate visually without GPS support [4]. The history and scrutinized contemporary efforts in the research community to accentuate the fact that 3D reconstruction has significance on higher-level and mid-level of 3D understandings [5]. In [6] Nguyen, Izadi, Lovell have proposed a system that quantify lateral as well as axial noise patterns of angle and distance of a perceived surface using Kinect sensors. In paper [7], the authors introduced an approach to fetch 3D structured surface area of cultivated land with sensors of Kinect Camera. The main aim of their technique is to obtain three-dimensional land surface data that estimates wheat emergence capacity. The researchers have utilized Kinect depth-data that stacks set of 3D datasets from forest gravel roads to quantify surface roughness. The sole purpose of the research was to identify the roughness of distinct gravel roads they correlated vibration analysis with three-axis accelerometers which were placed into experimental vehicles [8]. Bokaris, Muselet, Trmeau have explained an approach to reconstruct three-dimensional indoor scenes consisting planar plots also uses RGB-D data for a single image [9].In paper [10], the authors have shown which will explore the technique of making an improvement of the Kinect fusion algorithm that enhances the precision of 3D scanning objects and lessens the noise factors. The authors in [11] have also presented a remarkable procedure by applying an upgraded version of (LEHF) Line-based Eight-directional Histogram that constructs better 3D model than original Kinect fusion. Ling, Shlizerman, Shapiro present a unique algorithm that yields high-resolution 3D meshes of face images captured by a single depth view of Kinect, this method also includes Kinect Fusion that allows the algorithm to take high depth images [12]. Calibration of depth camera and accuracy setting has been implemented on the other two papers [13-14].

In this paper, we have developed our Kinect sensor reconstruction unique algorithms in C# programming with real time fusion model of indoor environmental scenarios for 3-D reconstructions.

3. DATA ANALYSIS

Kinect has mainly 3 sensors for its operation. An RGB camera, IR light projector and CMOS sensor. Light coding technique is actually the depth measurement into the Kinect sensor. CMOS sensor operates with the light coding techniques rather than other sensors which make Kinect sensor a low cost device. The IR emitter beams a packet of light which can track the shape of the object in the basis of distance. In some past

studies, standard deviation of depth data with calibration error has been shown [13-14].

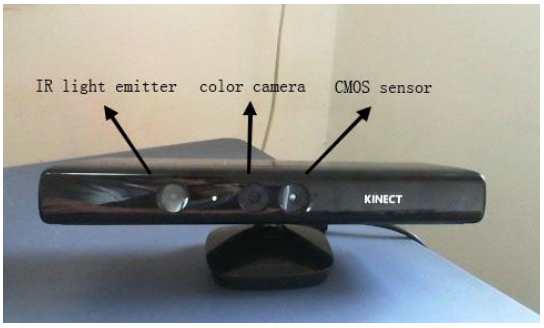


Fig. 1. The Microsoft Kinect sensor, X-box 360.

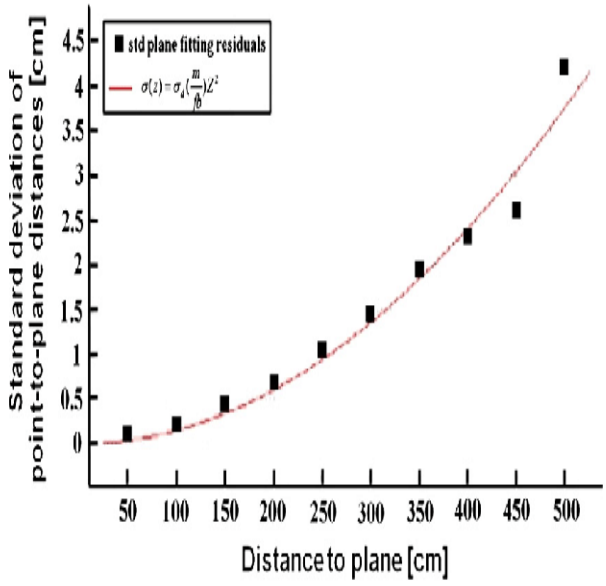


Fig. 2. standard deviation of plane fitting with different distance.

TABLE I
COLOR CAMERA INTERNAL COEFFICIENTS

f_{cx}	f_{cy}	u_{c0}	v_{c0}
532.9	531.39	318.57	262.08
0.06	0.05	0.07	0.07

k_1	k_2	k_2	k_3
0.2447	-0.5744	0.0029	0.0065
0.0004	0.0017	0.0001	0.0001

TABLE II
DEPTH CAMERA INTERNAL COEFFICIENTS

f_{dx}	f_{dy}	u_{d0}	v_{d0}
593.36	582.74	322.69	231.48
1.81	2.48	1.34	1.59

4. METHODOLOGY

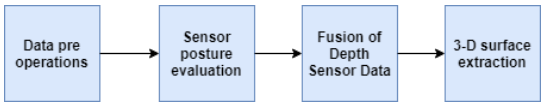


Fig. 3. 3-D surface reconstruction working flow chart.

Here, the flow chart of this paper will show the abstract of our methods.

Algorithm 1 3D Surface Reconstruction

Input: Real-time depth data

Output: 3D Surface Reconstructed Model

Clear $FusionImageFrame\ memories$
$GetImageSize(DepthImageFormat)$
$GetCameraColorImage(Internal\ Coefficients)$
$GetDepthCamera(Internal\ Coefficients)$
$Initialize$ Kinect sensors
$sensorEnableDepthStream(DepthImageResolution)$
$frameDataLength \leftarrow DepthStream(FramePixelDataLength)$
$DepthImagePixel[frameDataLength]$
$colorPixels[frameDataLength]$
$[Width, Height] =$ *[ImageWidth * ImageHeight]*
$SurfaceMeasurement \leftarrow$ *(bilateral filter)*
$colorBitmap = WriteableBitmap(Width, Height, PixelFormats(BGR))$
$ImageSource \leftarrow$ *colorBitmap*
$SensorPose \leftarrow$ *(ICP Algorithm, VolumeRepresentation, Surface extraction)*
$vol =$ *Reconstruction(VoxelsPerMeter, VoxelX, VoxelY, VoxelZ)*
$Volume =$ *FusionReconstruction(vol, worldToCameraTransform)*
$depthFloatBuffer \leftarrow$ *FusionFloatImageFrame(Width, Height)*
$pointCloudBuffer \leftarrow$ *FusionPointCloudImageFrame(Width, Height)*
$shadedSurfaceFrame \leftarrow$ *FusionColorImageFrame (Width, Height)*
if $sensor \neq$ *enable* **then**
    Stop FPS Timer
    SensorDepthFrameReady(sender, EventArgs)
    if $depthFrame \leftarrow OpenDepthImageFrame$ **then**
        if $depthFrame \neq null$ **then**
            $depthFrame(depthImagePixels)$
            $DispatcherBeginInvoke(ProcessDepthData)$
            $ProcessDepthData(volume, shadedSurfaceFrame, colorBitmap)$
            if $tracking \neq success$ **then**
                $statusBarText \leftarrow$ *CameraTrackingFailed*
            else
                $tracking \leftarrow$ *volume(ProcessFrame(depthFloatBuffer))*
                $tracking \leftarrow$ *volume(worldToCameraTransform)*
            end if
        end if
    end if
else
    Start FPS Timer
end if

4.1 Data pre-operations

This stage implies the raw data of depth sensor camera measurement. After the data measurement, we will generate the point cloud as well as figure out the map. But the point is, there are error into the depth data measurement [3]. For example: when the edge of the surface are scanned, a low cost sensor device may have some wrong information about the depth image calculation. Furthermore, little Gaussian noises exist in the depth map that will make the surface unsmooth. In fig.2, we can see that the plots are distracted from the region lines of the actual values. This happens when Kinect is in a long distance from its corresponding object. In the line of 450 cm or approximately 15 meters, the actual depth distance calculation will be wrong and far from the actual value. Some gross error may occurred3 when scanning the edge and IR may

not reflect a sharp object which will create an undependable data. For example: during scanning objects like metal glass or hardware tools, Kinect cant detect the object as these tools dont reflect the infrared light. For overcoming these problems, we need to pre-process the raw data to get rid of gross error as well as noises every time. We applied bilateral filter to reduce Gaussian noises and gross error.

4.1.1 Bilateral Filter:

Bilateral filter is nonlinear filter for image that reduce the noise and preserves the edge by restoring the intensity of each pixel with the weighted average of intensity value from nearby pixel [9-10]. The bilateral filter is defined as:

$$I^{filtered}(x) = \frac{1}{W_p}\sum_{xi\in\Omega} I(x_i) f_r(||I(x_i)-I(x)||) g_s(||x_i-x||)..(1)$$

here,the normalization term will be defined as :

$$W_p = \sum_{xi\in\Omega} f_r(||I(x_i)-I(x)||) g_s(||x_i-x||)..(2)$$

here, $I^{filtered}$ is the filtered image, I is the original input image, x is the required coordinates that is to be filtered, Ω is the window which is located in x, f_r is the range kernel for the smoothness in intensities and g_s is the spatial kernel for the smoothness in coordinates. Now the weight formula is:

$$\omega(i,j,k,l) = \exp(-\frac{(i-k)^2+(j+k)^2}{2\sigma^2} - \frac{||f(i,j)-f(k,l)||^2}{2\sigma_r^2})..(3)$$

where σ and σ_r are smoothing parameters, and I(i,j), I(k,l) are the intensity of pixels (i,j) and (k,l) accordingly. Here, let the size window σ to be 10, 30 and 50 respectively and σ_r to be 0.1, 0.3 and 0.5. The figure has shown in below fig.4.

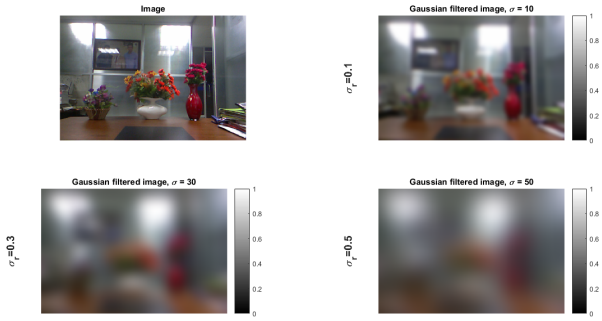


Fig. 4. Bilateral filter application for pixel intensity.

4.1.2 Point to point cloud with normal:

Let a matrix K that defines matrix of camera calibration by depth sensor camera. In time k_0, original raw depth map R_k shows the depth calculation R_k(u) at each image pixel u. showing the camera posture/ pose by transformation matrix $T_{g,k}$ and this is a 3DOF (3 degree freedom) rotation matrix. As Kinect can rotate 27 degree up to 27 degree down with total 54 degree 3-D movement at straight. 3DOF makes Kinect more reliable [4].

$$K = \begin{bmatrix} f_{dx} & 0 & u_{d0} \\ 0 & f_{dy} & v_{d0} \\ 0 & 0 & 1 \end{bmatrix} ..(4)$$

$$T_{g,k} = \begin{bmatrix} R_{g,k} & t_{g,k} \\ 0 & 1 \end{bmatrix} ..(5)$$

$$u = (u,v)^T, \dot{u} = (u^T|1)^T ..(6)$$

$\dot{u}$ is the corresponding homogeneous vector of u. now the normalize vector will be defined as: $N_k(u) = \frac{N_k(u)}{||N_k(u)||^2} ..(7)$

Now the corresponding image $\dot{u}$ is: $V_k(u) = R_k(u)K^{-1}\dot{u}$. Here, V_k is the point cloud and it is obtained by back projecting the filtered depth value into sensors scanning space.

4.2 Sensor Pose/posture Estimation:

As soon as we get the point cloud then we proceed for the 3-D surface model. For that we need to get the rigid transformation matrix of the sensor posture and reform point cloud to global model. In this stage, we need ICP algorithm to find the important point with higher accuracy and efficiency. It also decreases the number of correspondence. RGB frames optical flow calculation is important to find out point features between two closest frames. Here base alignment can be done by the previous transformation matrix equation (5). ICP algorithm then will be used forgetting the accurate sensor pose estimation [11].

4.2.1 ICP Algorithm:

Iterated closest point. The basic characteristics is to find the rotation and translation values between two points. It is really an essential algorithm to calculate the displacement between scans and determine the correct data association. The rotation valued vectors are represented as: $\vec{q}_{rot} = [q_0, q_1, q_2, q_3]..(8)$; here $q_0 \geq 0$ and $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$. The three by three rotational matrix by unit quaternion is defined as:

$$R = \begin{bmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2(q_1q_2-q_0q_3) & 2(q_1q_3+q_0q_2) \\ 2(q_1q_2+q_0q_3) & q_0^2+q_2^2-q_1^2-q_2^2 & 2(q_2q_3-q_0q_1) \\ 2(q_1q_3-q_0q_2) & 2(q_2q_3+q_0q_1) & q_0^2+q_3^2-q_1^2-q_2^2 \end{bmatrix} ..(9)$$

The translation unit vector is $\vec{q_t} = [q_4, q_5, q_6]..(10)$. By this the result of $\vec{q}$ is $\vec{q} = [\vec{q}_{rot}|\vec{q_t}]^T..(11)$. At the time of scanning by RGB-D camera, let us assume $P = \vec{p_l}$ is the calculated data set which is needed to align with point $X = \vec{X_i}$, and N_x and N_p (where N_x = N_p) be the model point amount and measured point amount respectively. The mean square object equation: $f(\vec{q}) = \frac{1}{N_p}\sum_{i=1}^{N_p} ||\vec{x_l} - R(\vec{q}_{rot})\vec{p_l} - \vec{q}_T||^2..(12)$ The translation vector will be defined as:

$\vec{q_t} = \vec{\mu}_x - R(\vec{q}_{rot})\vec{\mu}_p$ ..(13)

Here, $\vec{\mu}_x$ is the central mass of the measured point set P.

4.3 Fusion of Depth Sensor Data:

with the implementation of sensor posture estimation, every single frames of depth images get fused with each other and become one single model of frame through volumetric representation method. This also known as truncated signed distance function (TSDF) [12].

4.3.1 Volumetric representation:

it is really obvious that visuals are captured into 3-D forms into human eyes. Every surface and objects are in 3 dimensional form. So Kinect depth data frames volumetric understanding is the next step for object recognition and reconstruction. Every figure is made up by the alignment of pixels. Volumetric pixel is called voxel. Voxels represent color

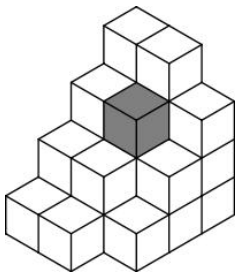


Fig. 5. 3-D format volumetric representation of Kinect sensor.

picture by aligning its position [1].

Voxel shows data point center in 3-D module. This could be opacity, various segment of data etc. This solely depend upon the user how he/she wants to use the data point. Now in truncated signed distance function, weighted signed distance function D(x) will be: $D(x) = \frac{\sum w_i(x)d(x)}{\sum w_i(x)}..(14)$ and the cumulative weight W(x) will be: $W(x) = \sum w_i(x)..(15)$ Here $d_i(x)$ is named signed distance function from i-th distance of image. Now describing the un-weighted signed distance function, the very fundamental phenomenon is adjusting some critical situation when sensor scan the object with distance, sometimes the space surface has negative distance while the behind have some positive distance. When the weight is 1, surface will be created by the averaging of two range surfaces. Now we fuse the depth frames incrementally like:

$$D_{i+1}(x) = \frac{w_i(x)D_i + w_{i+1}(x)d_{i+1}(x)}{w_i(x) + w_{i+1}(x)}..(16)$$

$$W_{i+1}(x) = w_i(x) + w_{i+1}(x)..(17)$$

4.4 Surface Extraction:

Each time we upgrade the TDSF, every voxel is targeted into image plane of the depth sensor to have the corresponding image pixel. After this the signed distance will be calculated. When the normal map is extracted into previous process stage, normal x is used to show the normal of the projected voxel using the similar formula as (16). Now about the voxel, the surface extraction will be: S(p) = p + N(p). F(p); (18) here, p is the voxel point. But we cant accept the result if surface is extracted from outside voxel.

5. EXPERIMENTAL RESULTS

Our total Kinect algorithm project is built with C# programming with MATLAB application. Starting with the Data pre-processing to Surface extraction for 3D reconstruction, every steps has been implemented accordingly and successfully. First step was surface measurement for reducing the Gaussian noise and gross error in depth data. If we project the camera calibration matrix into real time parameter co-efficient from data table, the following result could be found for 3D model Data,

$$K = \begin{bmatrix} f_{dx} & 0 & u_{d0} \\ 0 & f_{dy} & v_{d0} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 593.36 & 0 & 322.69 \\ 0 & 582.74 & 231.48 \\ 0 & 0 & 1 \end{bmatrix}$$

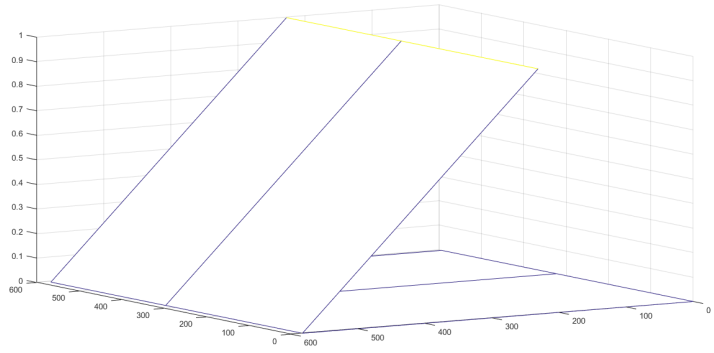


Fig. 6. 3-D model of calibration matrix of depth camera.

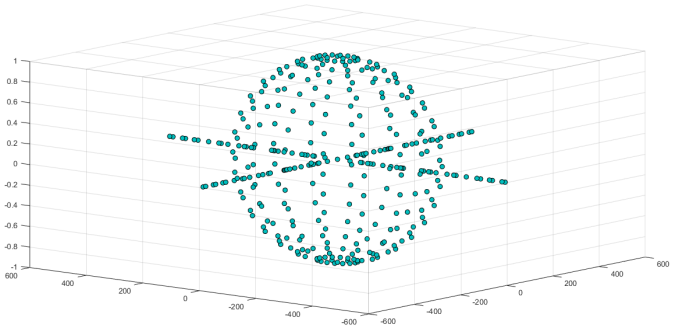


Fig. 7. Scattering 3-D model of calibration matrix of depth camera.

Bilateral filter can certainly improve the quality of pixels and recover every pixel that need to be fixed. The value is already given.

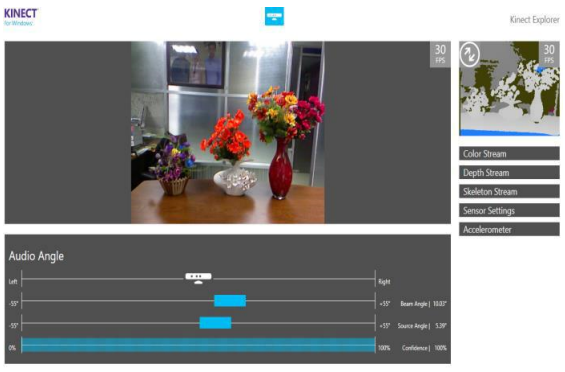


Fig. 8. Raw data image in Kinect sensor.

The basic principle of the ICP algorithm is fusion with the point clouds. Alignment of the point and give the correct data. There could be some error crossing in alignment of the data point. But ICP algorithm helps to gain a satisfactory level of accuracy.

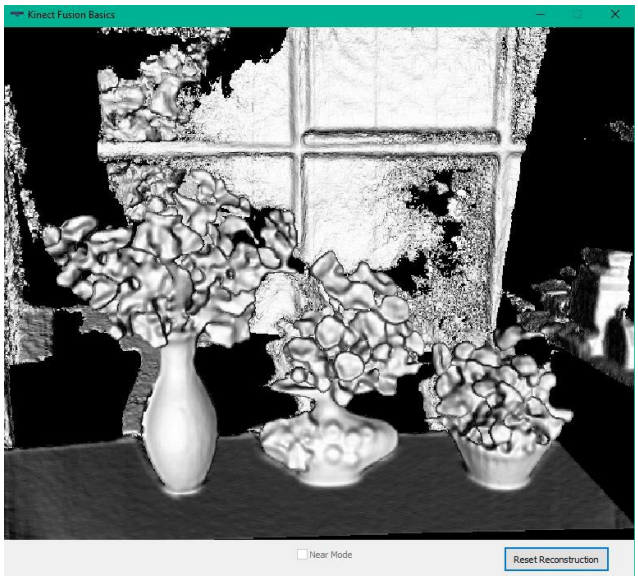


Fig. 9. Fusion point cloud with ICP algorithm.

Kinect even works in darkness with very good efficiency. There will be a big error in data point if we don't implement the ICP algorithm properly.

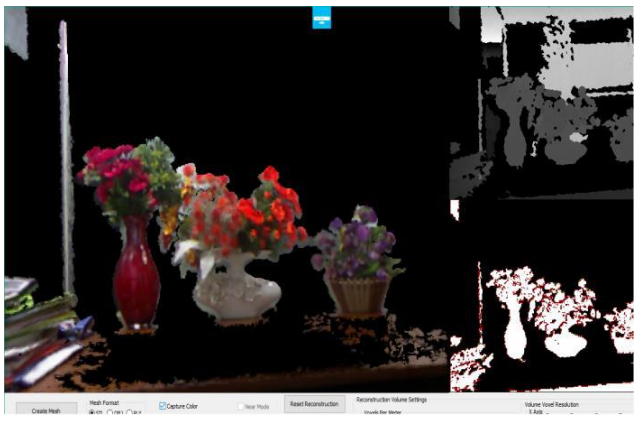


Fig. 10. Kinect real time fusion mapping.

A little bit of disadvantage that Kinect fusion has that sometimes Kinect sensor consume a little more time than usual for fusion mapping reconstruction. This happens because sometimes the aim object is not placed properly or sometime IR takes more time to scan. It is clearly visible that in fig.12 the second image is more accurate and visible. The comparative error rate of the fusion reconstruction image in fig.12 became very low in error rating comparing with the first image in fig. 12 and became more accurate.

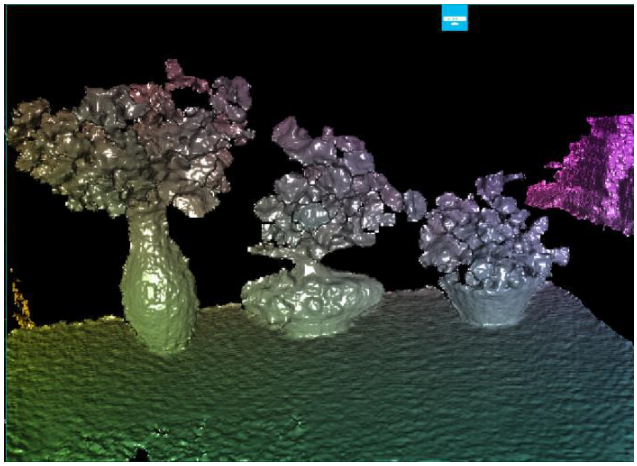


Fig. 11. Kinect fusion real time surface reconstruction.

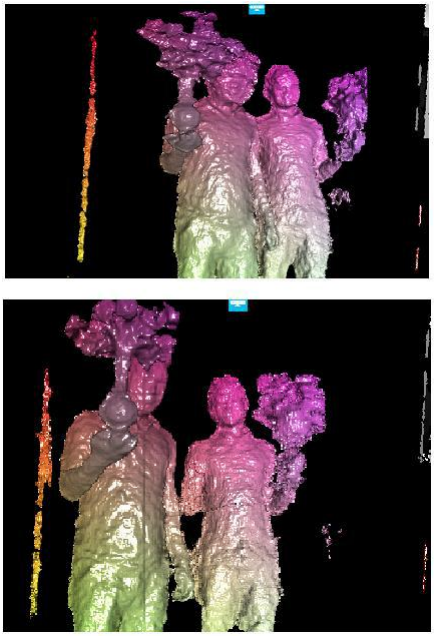


Fig. 12. Comparison of some limits of Microsoft Kinect.

6. DISCUSSIONS AND FUTURE WORK

The overall observation of the paper is clear cut. The main focus of this paper is C# algorithm structure with indoor environment Kinect fusion mapping.The execution of the research is very much cheap and adjustable in any environment. In the field of image processing, 3D reconstruction by RGB-D sensor may very much common, but apart from that similarity, Our pseudo-code is robust and very efficient to use. Some constructive error has been modified due to adjusting problem of kinect. But that could be executed by any machine and that is completely normal. In future, we will try to execute industrial security measurement by the depth sensor camera like Kinect. It will be an unique approach for the industrial security issue.

7. CONCLUSION

In this paper, we have shown the 3-D Microsoft Kinect fusion indoor environmental reconstruction mapping with accuracy and authenticity. The paper reflects the state-of-the art of different algorithms that helps Kinect to construct the work perfectly. Our robust algorithm is built with C# programming and some of the 3-D graph and calculations has been measured by MATLAB. Latex software has been used to build our robust algorithm of this paper.

REFERENCES

[1] S. Tiangang, L. Zhou, D. Xinyang, W. Yi, "3D Surface reconstruction based on Kinect" *Third International Conference on Information Science and Technology (ICIST),*China,2013,pp.986-990.

[2] K.D. Mankoff, T.A. Russo, "The Kinect: a lowcost, highresolution, shortrange 3D camera" *Earth Surface Processes and Landforms,* vol.38, no. 9, pp.926-936, 2013.

[3] L. Guo, X. Chen, Y. Chen, B. Liu, "Colored 3D surface reconstruction using Kinect sensor" *Optoelectronics Letters,* vol. 11, no.2, pp. 153-156, 2015.

[4] N. Namitha, S.M. Vaitheeswaranb, V.K. Jayasree, M.K. Bharat, "Point Cloud Mapping Measurements Using RGB-D Sensor and Kinect Fusion for Visual Odometry" *Twelfth International Multi-Conference on Information Processing-2016 (IMCIP-2016),* India, 2016,pp. 209- 212.

[5] J. Xiao, "3D reconstruction is not just a low-level task:retrospect and survey", *Computer Vision at Princeton,* USA, 2013, pp.1-8.

[6] C.V. Nguyen, S. Izadi, D. Lovell, "Modeling Kinect sensor Noise for Improved 3D Reconstruction and Tracking" *Second Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization & Transmission,* USA, 2012,pp.524-530.

[7] T. Liu, W.H. Chen, F. Li, W. Wu, C. Sun, J. Ding, X. Zhu, W. Guo, "Characterization of the 3D structure of a cultivated land surface and its influence on wheat seedlings growth using Kinect" Scientific Reports, vol.7, no.1, pp.1-10, 2017.

[8] F. Marinello, A. R. Proto, G. Zimbalatti, A. Pezzuolo, R. Cavalli, S.Grigolato, "Determination of forest road surface roughness by Kinect depth imaging" *Annals of Forest Research,* vol.60, no.2, pp.217-226, 2017.

[9] P.A. Bokaris, D. Muselet, A. Trmeau, "3D Reconstruction of Indoor Scenes using a Single RGB-D Image" *12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP),* Portugal, pp.394-401, 2017.

[10] T. Yamaguchi, T. Emaru, Y.Kobayashi, A.A. Ravankar, "3D map-building from RGB-D data considering noise characteristics of Kinect" *International Symposium on System Integration,* Japan, 2016, pp.379-384.

[11] Y. Nakayama, T. Honda, H. Saito, M. Shimizu, N. Yamaguchi, "Accurate Camera Pose Estimation for Kinect Fusion Based On Line Segment Matching by LEHF" *22nd International Conference on Pattern Recognition (ICPR),* Sweden, pp.2149-2154, 2014.

[12] S.Ling, I.Shlizerman, L.G.Shapiro, "3D face Hallucination from a single Depth Frame" *Second International Conference on 3D Vision,* Japan, pp.31-38, 2014.

[13] C.D. Herrera, J. Kannala, J. Heikkila, "Accurate and Practical calibration of a depth and color camera pair" *Lecture Notes in Computer Science,* Berlin, pp.437-435, 2011.

[14] K. Khoshelham, "Accuracy analysis of Kinect depth data" *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. 38, no.12, pp.133-138, 2011.

Detection of Lung Cancer from CT Scan Images using GLCM and SVM

Mehdi Hassan Jony[1], Fatema Tuj Johora[2], Parvin Khatun[3] and Humayan Kabir Rana[4]
[1]School of Computer Science of Technology, Hefei University of Technology, China
[2,3,4]Department of Computer Science and Engineering, Green University of Bangladesh
[1]*mehdi@mail.hfut.edu.cn,* [2]*fatema@cse.green.edu.bd,* [3]*parvin.do.cse@green.edu.bd and* [4]*humayan@cse.green.edu.bd*

***Abstract*— Lung cancer is one of the most threatening diseases and prediction of it, is challenging problem nowadays. Most of the cancer cells are overlapped with each other. It is hard to detect the cells but also essential to recognize the presence of cancer cells in the early stage. Early detection of lung cancer may reduce the death rate. In this study, we use the grey level co-occurrence technique (GLCM) to extract the feature of the cancer affected lung image and then support vector machine (SVM) to detect abnormal lung nodules. Our experimental evaluation using MATLAB demonstrates the efficient performance of the proposed system.**

***Index Terms*— Lung Cancer Detection, Gabor Filter, Computed Tomography, Marker-Controlled Watershed, GLCM, SVM.**

I. INTRODUCTION

Lung cancer is the maximum mainstream cancer that cannot be neglected. It is the second most regularly analyzed cancer in men and women. An anticipated 234,030 new instances of lung cancer are normal in 2018, representing around 14% of all cancer analyze [1]. The rate has been declining since the mid-1980s in men, however just since the mid-2000s in women [1]. Gender orientation contrasts reflect authentic examples of smoking take-up and end in the course of recent decades. During 2005 to 2014, lung cancer frequency rates decreased by 2.5% per year in men and 1.2% per year in women [1]. Lung cancer represents a larger number of deaths than some other cancer in men and women [1]. An expected 154,050 deaths are relied upon to happen in 2018, representing around 1 of every 4 cancer death. The lung cancer death rate has decayed by 45% since 1990 in men and by 19% since 2002 in women due to reductions in smoking, with the pace of decline quickening over the past decade; from 2011 to 2015, the rate decremented by 3.8% per year in men and by 2.3% per year in women [1]. Identifying of lung cancer are often done through its symptoms embrace a cough, coughing of blood, fatigue, unexplained weight loss, shortness of breath etc. Computerized tomography (CT) is that the most efficacious technique of lung nodule detection for its competency to make three-dimensional (3D) pictures of the chest, leading to more immensely colossal resolution of nodules and tumor pathology. A CT image by a computer process to avail lung nodule medicine has been widely utilized in the clinic. the method of computer availed diagnosis (CAD) of lung cancer are often divided into a detection system (often abbreviated as CADe) and diagnostic system (often abbreviated as CADx). The forsaken system divides the candidate nodules kenned within the precedent step into nodules or non-nodules (i.e., traditional anatomic structures). The goal of the CADx system is to relegate detected nodules into benign and malignant nodules [2]. Since the likelihood of malignancy is approximately associated with the geometric size, shape, and look, CADx will distinguish the benign and malignant pulmonary nodules by the efficacious options like texture, shape, and rate of magnification. Thus, the prosperity of a culled CADx system is often quantified in terms of precision of identification, haste, and automation level [3].

In this study, image enhancement, segmentation, and feature extraction and lung nodules detection procedures have been implemented using MATLAB. The paper is aiming to get the more accurate results by introducing a new feature extraction process which is much effective and less time-consuming to detect cancerous lung. In the first section of this paper, methodology has been explained and rest of the paper following contents are classification with SVM as well as result and discussions. And at the end, summary of the works and future work scope has been explained in conclusion.

II. RELATED WORKS

S. Avinash et al [4] showed that early detection of lung cancer is applicable for efficient analysis and it will help the opticians to provide initial treatment. Prior to this, researchers employed methods such as fast fourier transform (FFT) for image enhancement, thresholding method for segmentation and extrusion methods for feature extraction etc. The research has computerized these selections, enhancing databases, selecting features of lung cancer images, lung cancer detection, quality and accuracy improvement. Image processing approaches brought huge benefit for identifying Lung Cancer at an early stage. Magnetic resonance imaging (MRI), which produces computerized images of the body that are used to diagnose several different cancers, was developed by Paul Lautenberg and Peter Mansfield [5]. The imaging is based on the different signals obtained from different tissue

types that are subjected to a high magnetic field. In 2003 the developers won the Nobel Prize in medicine for their work.

Radiologists have interior relied on examining images from chest radiography and PET scans to detect lung cancer [6]. Fatema el al. [7] proposed a technique based on CT technology, the proposed system consists of a few steps like photo acquisition, preprocessing, feature extraction and categorization. At the preprocessing stage, RGB images are converted to grayscale images, used for reducing the Gabor filter, the adaptive thresholding method transforms CT scan images into binary images, and the REGIONPROPS function is used for the right body, from the binary image to the region. Feature extraction platforms, such as contrast, correlation, energy, home genetics, are characterized by a static method called Gray Level Co-occurrence Matrix (GLCM). To ameliorate the hand annotation of nodules, medical experts expand beyond an axial scan mode and rely on other techniques such as maximum intensity projections and 3D volume renderings [8]. Maximum intensity projection (MIP) is a volume rendering technique for 3D images that projects voxels with a maximum intensity of the parallel rays from a given viewpoint onto the plane [9]. This technique makes it more facile to detect denser objects like nodules since maximum projections will be concentrated in a particular area, whereas other structures like thin blood vessels will have maximum intensities more distributed throughout the lung/image. In Rashmita al. [10] are based upon gabor filter within Gaussian rules, thereafter the segmentation principles are applied over the enhanced region of the image and the input for feature extraction is obtained. In this research they have done classification based upon artificial neural networks which is used their own software.

In [11] the authors largely fixate on adequately great amelioration in contrast of masses along with the suppression of background tissues is acquired by processing the parameters of the proposed transformation function in the definite range. The manual analysis of the sputum samples is time-consuming, erroneous and needs concentrated trained person to evade diagnostic mistakes. The segmentation results will be utilized as a base for a Computer Availed Diagnosis system for early detection of cancer, which increases the chances of the patient survival. As well as Gabor filter is proposed by authors for enhancement of medical images. It is a good enhancement to implement for medical images.

Thus, Researchers are becoming more and more concerned with the elaboration of automated CAD systems for lung cancer. Many publications proposed different automated nodule recognition systems utilizing image processing, and including, different techniques for segmentation, feature extraction, and relegation. Moreover, the detection processes work in the middle stage of the lung cancer, which is very hazardous for the affected patient [12].

III. METHODOLOGY

The methodology focuses to obtain a more accurate result. In the proposed method, it is enhancing the contrast of the input image through the pre-processing method. It is done by first converting the input image to a grayscale image and then it is modified with the Gabor filter to enhance [13].

In image processing, segmentation is moreover very important part, in this section ideally enhanced images is segmented into two or more sub-segments that will be the more easier representation of the images for further extracting the elements. Segmentation could be utilized for object recognition, occlusion boundary estimation within motion or binaural systems. The goal of image segmentation is to cluster pixels into relevant image regions, corresponding to individual surfaces, objects, or natural components of objects. We mainly follow an approach like a Marker-Controlled Watershed. In feature extraction, a pixel comparing process is proposed in this paper.

We can see the various steps of lung cancer detection in the following fig 1.

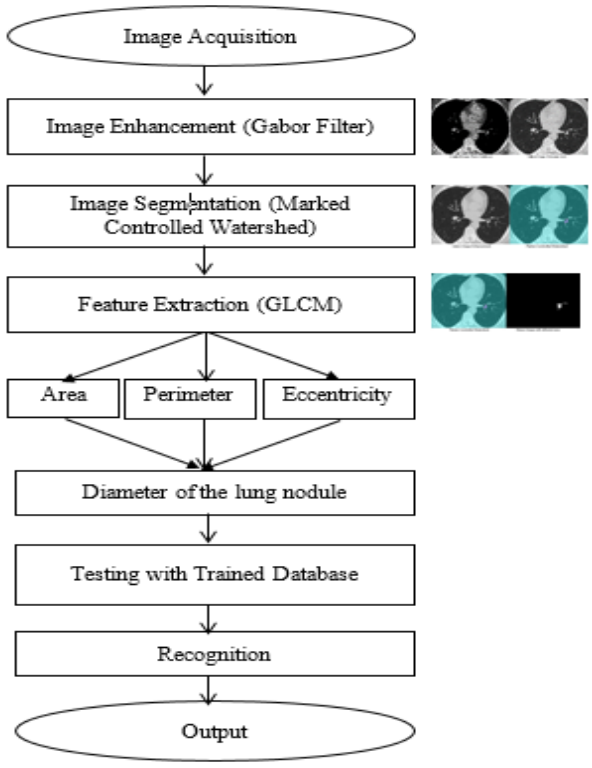


Fig. 1. Block diagram of the proposed system to detect lung cancer.

A. Image Acquisition

The first stage of methodology starts with taking a collection of CT images (normal and abnormal) from the available Database from IMBA Home (VIA-ELCAP Public Access) [12]. Lung images are very high when X-ray and MRI compared to the CT scan image. Therefore, they are considered for developing strategies. The main advantage of using calculated tomographic images gives more transparency and less distortion. In the appearance of these facts, the images collected from ELCAP public lung image database which is released December 20, 2003. The whole-lung dataset consists of 50 CT scans obtained in a single breath hold

with a 1.25 mm slice thickness. DICOM (Digital Imaging and Communication in Medicine) has become a standard for medical imaging.

The acquired image is in raw form. A lot of noise in the acquired image is observed. To improve contrast, transparency, background sound separation, it needs to be pre-processed. Therefore, different techniques like enhancement, the required form is developed from images.

B. Image Enhancement

The image pre-processing stage commences with image enhancement. The aim of image enhancement is to modernize the interpret-ability or perception of information included in the image for human viewers or to provide better input for other computerized image processing techniques. Image enhancement techniques can be divided into two categories: Spatial domain methods and frequency domain methods [14].

1) Spatial domain techniques- which operate directly on pixels.
2) Frequency domain techniques- which operate on the Fourier transform of an image.

The various image enhancement techniques can be characterized as spatial domain methods and frequency domain methods. This includes smoothing of image and removal of noises, blurring etc. There are many ways to enhance the image but Gabor filter was found to be suitable for both the CT and MRI images. In the image enhancement stage, we utilized the following Gabor filter techniques.

Gabor Filter: A Gabor filter is a linear filter whose impulse replication is defined by a harmonic function multiplied by a Gaussian function. Because of the multiplication-convolution property, the Fourier transform of a Gabor filters impulse replication is the convolution of the Fourier transform of the harmonic function and the Fourier transform of the Gaussian function [15]. It fundamentally analyses whether there is any concrete frequency content in the image in concrete directions in a localized region around the point or region of analysis. A set of Gabor filters with different frequencies and alignment may be helpful for extracting useful features from an image. In the discrete domain, 2D Gabor filters are given by,

$$G_c[i,j] = Be^{\frac{i^2+j^2}{2\alpha^2}} \cos(2\Pi f(i\cos\theta + j\sin\theta)) \quad \text{.....(1)}$$

$$G_s[i,j] = Ce^{\frac{i^2+j^2}{2\alpha^2}} \sin(2\Pi f(i\cos\theta + j\sin\theta)) \quad \text{.....(2)}$$

Where B and C are normalizing factors to be determined. 2-D Gabor filters have rich applications in image processing, especially in feature extraction for texture analysis and segmentation. f defines the frequency being looked for in the texture. By varying θ , we can look for texture oriented in a particular direction. By varying α , we change the support of the basis or the size of the image region being analyzed shown in fig 2.

C. Image Segmentation

Image segmentation is to facilitate the representation of an image into something that is more meaningful and easier

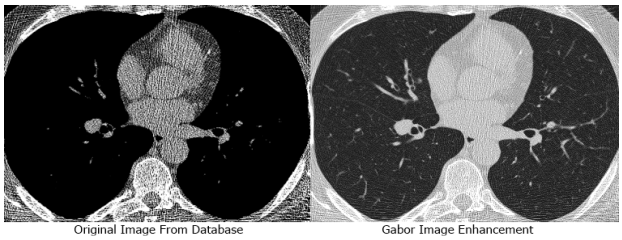


Fig. 2. Enhanced image by using Gabor filter

to analyze. Image segmentation is used to allocate objects and boundaries. It is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics [16]. The result of image segmentation is a set of segments that collectively cover up the entire image or a set of contours extracted from the image. Each pixel of the image in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture; adjacent regions are crucially different with respect to the same characteristics [17]. In this research, we used watermarked image segmentation approaches:

1) Marker-Controlled Watershed Approach: The Watershed transformation considers the gradient enormity of an image as a topographic surface. Pixels having the highest gradient enormity intensities (GMIs) match up to watershed lines, which represent the region boundaries. Water placed on any pixel roofed by a flow downwards to a common local intensity minimum (LMI). Pixels are draining to a common minimum form like a catchment basin, which signify the regions. Image segmentation result is a set of segments that cover a set of contours that are fully outlined in the whole image [18]. The marker-controlled pond section follows these basic methods and shows such fig 3.

In segmentation, image are been segmented in various sections which made easy for extraction features with GLCM process. With watershed segmentation, we detected the unwanted nodules with abnormal pixels. The outlined of abnormal pixels positions with light violet color are very similar with given position (affected cell with nodules) which is obtained from the database.

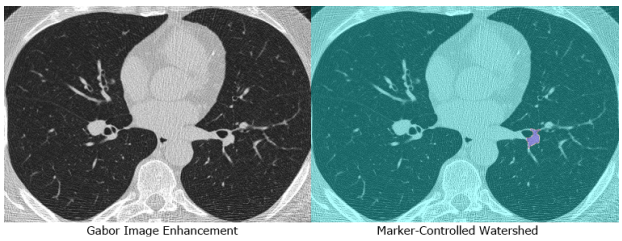


Fig. 3. Enhanced image to watershed

Algorithm 1 Marker-Controlled Watershed

1: **procedure** MARKER-CONTROLLED WATERSHED
2: Calculate a segmentation function. This is an image in which we are trying to segment them with dark areas.
3: Counting foreground markers. This is attached to the blobs of pixels in each object.
4: Counting background markers. This pixel is not part of an object.
5: Modify the segmentation function so that it only has minima at the front and background marker position.
6: Calculate the watershed transform of the revised segmentation function.

D. Feature Extraction

In image processing, the Image features extraction stage is very vital image processing techniques which utilizing algorithms and techniques to detect and isolate portions or shapes (features) of an image [19,20]. Once the segmentation is accomplished on the lung region, the features can be obtained from it and the diagnosis rule can be designed to precisely detect the cancer nodules in the lungs. This diagnosis rules can eliminate detection of cancer nodules resulted in segmentation and provides a better diagnosis [21-23]. In this stage to obtain the general features of the enhanced segmented image, GLCM is used as shown in fig 4.

Grey Level Co-Occurrence (GLCM): For extracting feature using the Gray-level co-occurrence Matrix (GLCM) to represent a set of features to reduces the misclassification of the glaucoma image. GLCM describes how many coordinates of pixel brightness values in an image illustration. GLCM represents the combined frequency of all pairs of gray levels i and j divided by a distance in the direction of nodules. GLCM Characteristics of basic character areas, perimeter, and Eccentricity this scale is measured in. These features are defined as:

A) Area: This is a scalar value that gives the actual number of actual pixels in extracted ROI (Region of interest). The Transformation function creates an array of ROI, which contains pixels with 255 values.

$$\text{Area} = A = (A_{i,j} \text{ , X ROI[Area]} = i \text{ , Y ROI [Area]} = j) \quad ..(3)$$

Where, i,j are pixels in size within the shape. ROI is Interest Area. X ROI [] is vector contain ROI X position, Y ROI [] vector retains ROI Y position [24].

B) Perimeter: This is a scalar value that actually gives nodule pixels. It extracted ROI borders length. Transform function creates an array on the edge of which the pixels contain 255 values, at least one pixel that contains 0 values [22].

$$\text{Perimeter} = P = (P_{i,j} \text{ , X edge[P]} = i, \text{ Y edge[P]} = j) \quad ..(4)$$

Where X is the edge [] and Y is the edge [] The vector represents the combination of ith and j^{th} pixels by creating curves respectively.

C) Eccentricity: This metric value is called as roundness or circularity or irregular complex (I) equal to 1 for only round and it is less than 1 for any other shape.

$$\textbf{Eccentricity} = \frac{Lengthof Major Axis}{Lengthof Minor Axis} \quad ..(5)$$

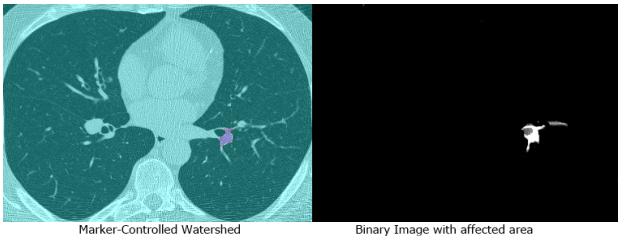


Fig. 4. Feature extraction from segmentation

After segmentation when we used GLCM on segmented images, it made a gray scale image which helped us to abstract the abnormal nodule sections with white color. Then we measured the sections with following steps. Approximate features for separate nodule of given sample images mentioned in Fig. 5 were found as follows:

1) Area: It is the most simple property and its given size. Therefore, it is the total number of white pixels in the extracted area.

2) Perimeter: Another simple property defined by the scope of this region. It's external ROI borders length.

3) Eccentricity: It is used to decide the shape or circularity of the object.

Area: 2291 Perimeter: 221 Eccentricity: 0.8289 Stage 3 Detected

IV. CLASSIFICATION

Support vectors are a learning model that is supervised by the relevant learning algorithms that can identify data analyzes and patterns for machine analysis. The basic SVM input takes a set of data and for each given input, predicts that the two classes form the input, it creates a non-possible binary linear classifier. SVM uses a kernel function that maps the given information to an individual location. Separation can be made with very complex boundaries. The various types of kernel functions, including polynomial, RBF, quadratic, Multi-Layer Perceptron (MLP). Each kernel is formulated by its own parameter γ, σ, and so on. Fig 5 shows the maximum margin hyper plane. The main hyper-plane algorithm is a way of creating non-linear classifications using the most margin hyper-plan kernel trick. In feature extraction we extracted and measured the existing nodules in the cells. Using SVM we created a hyper-plan.

V. RESULT AND DISCUSSION

A. Stage of Lung Cancer

The lung nodules are increased in small-scale in the lungs, which measure from 5 mm to 25 mm. The unusual nodule size in images are more than 25mm.

Assessing the size of cancer in staging and entering the

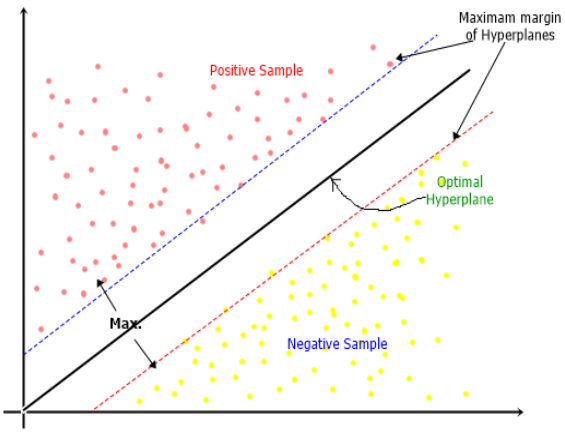


Fig. 5. SVM classifier hyperplanes

tissue surrounding it, as well as the presence or absence of metastasis in lymph nodes or other organs [25]. The stage from I to IV according to the intensity:

1) Stage I: cancer is confined to the lung
2) Stage II and III: cancer is confined to the chest
3) Stage IV: cancer has spread from the chest to other parts of the body.

This experimental information consists of 50 patients with more then 10,800 lung images. These 10,800 lung images are passed through the proposed method. Then the rules of diagnosis are generated from those images and these rules are sent to SVM classifier for the learning process. After learning, a lung image which is collected from hospital is passed through the proposed method. Then the proposed method has processed through its processing steps and finally it able to detect whether the following lung image is affected or not.

The proposed CAD system is capable of detecting nodules

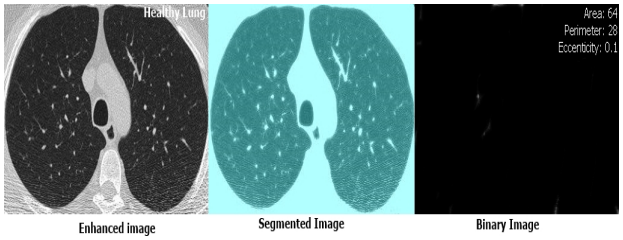


Fig. 6. A healthy lung

of lungs, larger than 25 mm of diameter, which means they are able to detect nodules in the lungs at the initial stage. In this way, diagnosing the primary disease will improve the patient's survival rate. Discrimination between cancerous and non-cancer candidates nodules which are shown in fig 6, 7 & 8. For making decision

B. Gaussian Radial Basis Function

Radial basis functions kernel have received significant attention, most commonly with a Gaussian of the form,

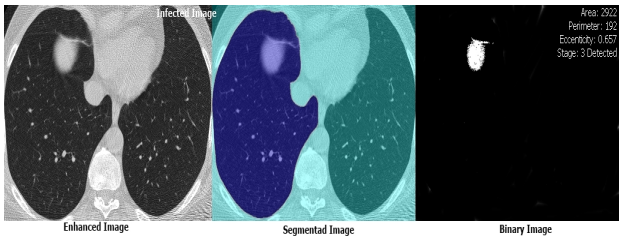


Fig. 7. A stage 3 affected lung

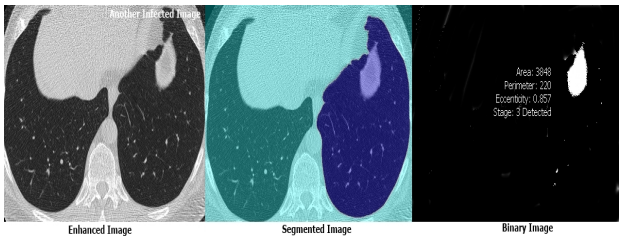


Fig. 8. Another stage 3 affected lung

$$K(a,a') = \exp(-\frac{||a+a'||^2}{2\sigma^2}) \quad .....(6)$$

Generally, a method of clustering is first employed to select a subset of the centre. An interesting feature of SVM is that this selection is perfect, each support vector centred on that data point, contributing to the local Gaussian function. The basic equation for support vector machines is:

$$C_o = H^T * F_o + b \quad .....(7)$$

Where, C_o= Output Class; H= Hyper-plane; F_o= Input Feature; b= Bias.

Now the dataset success rate is being compared with Rashmita al. [10] and Mokhled al. [13] which has same dataset with threshold binarization process. Patient one is the result with same patients' MRI analysis report is done by this paper and subject two is the result with same patient done by the following paper. Table 1 is highlighting comparison with our method to others. Confusion matrix is one of the most intuitive and easy metrics used to find the accuracy and accuracy of the any model. It is used for classification problems where the output can be two or more types of classes. Here, FP: False positive, FN: False Negative, TN: True Negative and TP: True Positive. TAR and FAR is calculated respectively as fraction of positive and negative scores exceeding my method.

TAR = TP/(TP+FN) and FAR = FP/(FP+TN)

VI. CONCLUSIONS

Lung cancer is one of the most dangerous diseases in the world. Proper diagnosis and early detection of lung cancer can increase the survival rate. CT scan images are collected from different hospitals and databases. A method is being proposed and developed for the detection of the stage of the cancer. Gabor Filters for image enhancement and Marker-Controlled Watershed segmentation gives the best results in the pre-processing phase. From the arising regions of interest (ROI), three properties are extracted from GLCM, such

TABLE I

DATASET TRUE/FALSE ACCEPTANCE RATE COMPARISON TEST

Patient	Type	Mokhled	Rashmita	Our-proposed
1	TAR	95.86	91.63	97.73
	FAR	4.14	8.37	2.27
2	TAR	93.47	90.02	96.23
	FAR	6.53	9.98	3.77
3	TAR	93.42	89.93	97.14
	FAR	6.58	10.07	2.86
4	TAR	91.93	90.67	95.66
	FAR	8.07	9.33	4.44
5	TAR	92.63	91.66	94.89
	FAR	7.37	8.34	5.11
6	TAR	93.12	93.41	96.27
	FAR	6.88	6.59	3.73
Total Success Rate		92.64%	89.97	95.32

as area, perimeter and eccentricity. These three properties measure the nodules to detect lung cancer levels. The size of nodule indicates the different levels of the tumor. Also for the purpose of classification, support vector machine modeling is an attractive approach. It combines generalization controls with a strategy to address dimension of nodules. In the classification problems, the generalization control is achieved by maximizing the margin, which is consistent with the weights of the weight vector in a canonical structure.

REFERENCES

[1] A. Staff, " *Cancer Facts & Figures 2018,*" Atlanta: American Cancer Society, Cancer, pp. 19-20, 2018.

[2] L. Sali, S. Delsanto, D. Sacchetto, L. Correale, M. Falchini, A. Ferraris, G. Gandini, G. Grazzini, F. Iafrate, G. Iussich, L. Morra, A. Laghi, M. Mascalchi and D. Regge, "*Computer-based self-training for CT colonography with and without CAD*", European Radiology, 2018.

[3] Q. Song, L. Zhao, X. Luo and X. Dou, "*Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images*", Journal of Healthcare Engineering, vol. 2017, pp. 1-7, 2017.

[4] S. Avinash, K. Manjunath and S. Kumar, "*An improved image processing analysis for the detection of lung cancer using Gabor filters and watershed segmentation technique*", 2016 International Conference on Inventive Computation Technologies (ICICT), 2016. Available: 10.1109/inventive.2016.7830084 [Accessed 24 January 2019].

[5] T. Jones, D. Duquette, M. Underhill, C. Ming, K. Mendelsohn-Victor, B. Anderson, K. Milliron, G. Copeland, N. Janz, L. Northouse, S. Duffy, S. Merajver and M. Katapodi, "*Surveillance for cancer recurrence in long-term young breast cancer survivors randomly selected from a statewide cancer registry*", Breast Cancer Research and Treatment, vol. 169, no. 1, pp. 141-152, 2018.

[6] P. Reid and S. Walter, "*Yield from a fast track referal system for radiologists suspecting lung cancer*", Lung Cancer, vol. 115, p. S16, 2018.

[7] F. Tuj Johora, M. Jony and P. Khatun, "*A New Strategy to Detect Lung Cancer on CT Images*", International Research Journal of Engineering and Technology (IRJET), vol. 05, no. 12, pp. 27-32, 2018.

[8] I. Valente, P. Cortez, E. Neto, J. Soares, V. de Albuquerque and J. Tavares, "*Automatic 3D pulmonary nodule detection in CT images: A survey*", Computer Methods and Programs in Biomedicine, vol. 124, pp. 91-107, 2016.

[9] S. Masrochah, "*Relationship between Slice Thickness to Artery Coronary Diagnostic Information on the Reconstruction of Maximum Intensity Protection (MIP)*",Journal of Medical Science And clinical Research, vol. 05, no. 06, pp. 23140-23145, 2017.

[10] R. Sehgal and S. Gupta,"*Lung Cancer Detection Using Neural Networks*", International Journal of Advanced Research in Computer Science and Software Engineering, vol. 6, no. 3, pp. 277-281, 2016.

[11] H. Chien and D. Mackay, "*How much complexity is needed to simulate watershed streamflow and water quality? A test combining time series and hydrological models*", Hydrological Processes, vol. 28, no. 22, pp. 5624-5636, 2013.

[12] P. P.Sundari, "*Analysis and Classification of Skin Lesions Using 3D Volume Reconstruction*", IOSR Journal of Computer Engineering, vol. 10, no. 2, pp. 35-40, 2013.

[13] M. S. AL-TARAWNEH, "*Lung Cancer Detection Using Image Processing Techniques*", Leonardo Electronic Journal of Practices and Technologies, no. 1583-1078, pp. 147-158, 2012.

[14] H. Rutika r, "*Automated detection and diagnosis from lungs CT scan images*", International Journal of Emerging Technologies and Innovative Research (www.jetir.org — UGC and issn Approved), ISSN:2349-5162,, vol. 2, no. 3, pp. pp785-787, 2015.

[15] S. Avinash, K. Manjunath and S. Kumar, "*An improved image processing analysis for the detection of lung cancer using Gabor filters and watershed segmentation technique*", 2016 International Conference on Inventive Computation Technologies (ICICT), 2016.

[16] M. Firmino, G. Angelo, H. Morais, M. Dantas and R. Valentim, "*Computer-aided detection (CADe) and diagnosis (CADx) system for lung cancer with likelihood of malignancy*", BioMedical Engineering OnLine, vol. 15, no. 1, 2016.

[17] H. Ren, Y. Zhou and M. Zhu, "*Tree Image Segmentation Based on an Improved Two-Dimensional Otsu Algorithm*", International Journal of Hybrid Information Technology, vol. 9, no. 9, pp. 199-210, 2016.

[18] Y. Zhang, H. Guo, F. Chen and H. Yang, "*Weighted kernel mapping model with spring simulation based watershed transformation for level set image segmentation*", Neurocomputing, vol. 249, pp. 1-18, 2017.

[19] Rana HK., Azam MS. and Akhtar MR. "*Iris Recognition System Using PCA Based on DWT*", SM Journal of Bio-metrics & Bio-statistics, vol. 2, no. 3, 2017.

[20] Rana HK, Azam MS, Akhtar MR, Quinn JMW, Moni MA., "*A fast iris recognition system through optimum feature extraction*", PeerJ Computer Science, vol. 5, pp. e184, https://doi.org/10.7717/peerj-cs.184, 2019.

[21] Y. Yuan, J. Wang, B. Li and M. Meng, "*Saliency Based Ulcer Detection for Wireless Capsule Endoscopy Diagnosis*", IEEE Transactions on Medical Imaging, vol. 34, no. 10, pp. 2046-2057, 2015.

[22] X. Jia and M. Meng, "*A deep convolutional neural network for bleeding detection in Wireless Capsule Endoscopy images*", 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2016.

[23] S. Suman, F. hussin, A. Malik, K. pogorelov, M. Riegler, S. Ho, I. Hilmi and K. Goh, "*Detection and Classification of Bleeding Region in WCE Images using Color Feature*", Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing - CBMI '17, 2017.

[24] Suzuki K., "*False-positive Reduction in Computer-aided Diagnostic Scheme for Detecting Nodules in Chest Radiographs*", Academic Radiology, Volume 13, Number 10, pp.10-15, February 2005.

[25] G. Aviram and M. Revel, "*Misclassification of Lymph Nodes in Lung Cancer Staging*", Chest, vol. 151, no. 4, pp. 733-734, 2017.

Single Image Haze Removal with Nonlinear Complex Diffusion Process

Sharmin Sultana, Nazifa Akter, Rahat Hossain Faisal and Md. Mostafijur Rahman
Dept. of Computer Science and Engineering,
University of Barishal,
Barishal, Bangladesh
sharmin.cse1.bu@gmail.com, nazifa.cse1.bu@gmail.com, rhfaisal@gmail.com and mostafij.csebu@gmail.com

***Abstract*—Hazy image enhancement is one of the challenging fields in the area of image processing because the presence of haze reduces the visibility of outdoor images. In order to remove haze, various algorithms have been proposed in recent years. Most of the algorithms remove haze but hardly enhance the image with better contrast. In this paper, we have proposed an effective method for removing haze from the image. The main benefit of our method is that it dehazes image as well as preserves ramp edges providing better contrast of the image. The qualitative and quantitative comparison among few states of the art methods have been performed which show that our method gives better output images compared to those methods.**

***Keywords*—Haze, Ramp Edge, Nonlinear Complex Diffusion Process, Contrast, Noise.**

I. Introduction

The appearance of the outdoor images is important in case of surveillance, intelligent vehicles and remote sensing system which may be affected by bad weather conditions such as rainy, foggy, haze and smoke etc. The image which is taken in bad weather condition offers low contrast and poor visibility of the scene. This condition is caused by small-sized water droplets which are present in the air [1]. Attenuation and airlight are two primary environmental extents cause scattering [2]. Attenuation is a light beam. It travels from a viewpoint via the atmosphere and gets attenuated because of the scattering by the atmospheric particle [1]. It reduces the contrast of the scene. Airlight guides to the change in color and this condition increases with the distance from the object [3].

For the better performance of the computer vision applications, haze removal is highly desired. For removing haze from an image we need to estimate the depth map or airlight map. Depth map holds information about the distance of the surfaces of scene objects from an angle of vision [4]. For estimating depth map various methods have been proposed based on two or multiple images. Two images are used for stereoscopic estimation of depth. Another approach uses multiple images of different times [2] or with different polarizing filters [5]. For existing images we can not use this kind of methods because they are constraining for the acquisition [3].

There are a good number of existing methods for removing haze from image. Robby T. Tan [6] proposed a method by observing two basic things. They first noticed that images with enhanced visibility have more contrast than images plugged by bad weather. The second thing is the airlight tends to be smooth. Based on these two observations, they developed a cost function in the framework of Markov random fields which can be efficiently optimized by various techniques. However, this method over enhances the image and the output image looks unnatural. Another method introduced by Raanan Fattal to remove the indeterminacy problem which is based on the independent component analysis (ICA). In this method optical transmission in hazy scenes is estimated and based on this estimation scattered light is removed to increase the visibility of the scene and recover haze from scene contrasts. Unfortunately, this method works only for color images not for gray images. As dense fog is colorless, this method also fails when there is dense fog. Another method proposed by Kopf et al. [7] which discards the need of various assumptions about the color of the scene. It is based on the use of a 3D model of the scene. But this method fails as it is an application dependant method and requires expert interaction. The methods mentioned above have one drawback which is computation time. So, to remove this problem Tarel et al. [8] proposed a method assuming airlight as a percentage between local standard deviation and local mean of the whiteness. Although this method is based on the linear assumption, it requires many parameters. He et al. [9] proposed another

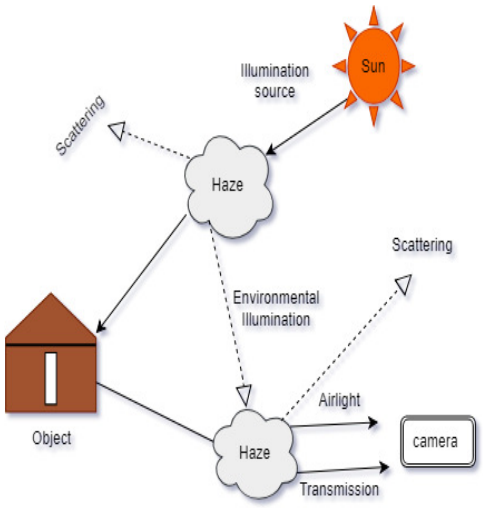


Fig. 1: Capturing image in hazy weather

effective method for outdoor RGB images based on the dark channel prior which also reduces computation time. Unfortunately, this method fails when the brightness of the scene objects are similar to atmospheric light. As most of the dehazing algorithm rely on user input, Sulami et al. [3] proposed a new method which does not rely on user input. Their proposed method based on recovering the atmospheric light vector in hazy scenes of a single input image. First, it estimates the color of the atmospheric light using few windows. Second, it tries to predict the magnitude of the atmospheric light by learning the magnitude of normalized color vector. However, this method is time consuming. Recently, another method proposed by Salazar et al. [10] uses a multilayer perceptron to compute the transmission map directly from the minimum channel and a contrast stretching technique to improve the dynamic range of the restored image.

From the aforementioned discussion it can be seen that though existing haze removal methods can remove haze, it hardly preserves ramp edges enhancing contrast of the images. Therefore, in this paper we proposed an effective haze removal method based on filtering approach. For filtering we used ramp edge preserving noise removal nonlinear complex diffusion process. Our proposed method is able to achieve equally and sometimes even better result to hazy images compared to existing methods. The novelties of this proposed method are as follows.

- A modified transmission map.
- We have refined transmission map using nonlinear complex diffusion process to preserve ramp edge.
- Post-processing using Multiscale Retinex Algorithm (MSR) to increase the contrast and brightness of the image.

II. Haze Model

The haze model which is widely used in computer vision and computer graphics to describe the formation of haze [2], [6], [11] is as follows:

$$I(x) = J(x)t(x) + A(1 - t(x)) \tag{1}$$

The goal of haze removal is to recover A, J and t from I where, I is the input image, J is the light reflected from its surface. x defines the pixel coordinates. A is the airlight vector representing single color in the image areas where t = 0. The scene transmission t(x) is distance dependant and is defined by Eq. (2).

$$t(x) = e^{-\gamma d(x)} \tag{2}$$

γ is the attenuation coefficient of the atmosphere and d(x) is the distance of the scene at pixel x from the sensor or camera. Here, the transmission acts as the matting between the scene and airlight. So, now from (1) and (2) haze model can be represented as,

$$I(x) = J(x)e^{-\gamma d(x)} + A(1 - e^{-\gamma d(x)}) \tag{3}$$

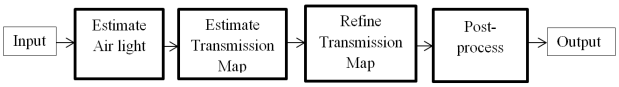


Fig. 2: Block diagram of proposed method

Here, in the right hand side first term is called direct attenuation [6] and second term is called airlight [6], [12]. For restoring haze free images information of airlight map A is needed. Airlight map depends upon the depth of the scene.

III. Proposed Method

In this paper, we have proposed a new haze removal method which first estimates the transmission map. After that, it refines the transmission map by nonlinear complex diffusion process. Nonlinear complex diffusion process removes haze from the image by preserving ramp edges. Hence, the haze-free image provides a clear concept about the objects. Afterward, we post-process the image using Multiscale Retinex Algorithm (MSR) because removal of haze from image gives a dark or low contrast view of the haze-free scene. So, post-process is needed because it will increase the contrast and brightness of the image. Our proposed method uses a modified transmission map which is different from other existing methods. Fig. 2 shows the block diagram of our proposed method which is described in details in the following subsections.

A. Estimate Airlight

In most of the existing methods, the airlight is estimated by the most haze-opaque pixels of the image. In Robby T. Tan's method, the pixel with the highest intensity is used as airlight and further refined [6]. However, this method does not always work because the brightest pixels could be any white object. Therefore, we have used dark channel for automatically estimating airlight. Dark channel is proposed by He et al. in [9] to find the minimum pixel in each color channel in hazy images. According to He et al. in [9] we took 0.1 percent brightest pixels in the dark channel. The pixel which has the highest intensity in the input image is chosen as airlight.

B. Estimate Transmission Map

In this section, We are going to estimate the transmission map. According to the dark channel prior method, the minimum intensity value can be described as,

$$J^{dark}(x) = \min_{y \epsilon \phi(x)} (\min_{c \epsilon (r,g,b)} (J^c, y)) \tag{4}$$

Here, J^c defines the color channel of RGB image J. $\phi(x)$ defines local patch around x. $\min_{c \epsilon (r,g,b)}(J^c, y)$ defines the minimum pixel of the RGB channel and $\min_{y \epsilon \phi(x)}$ is the minimum filter [13]. The dark channel value of the outdoor haze free images is close to zero when the image lacks haze [9]. It can be represented by $J^{dark} \rightarrow 0$. The region in which dark channel



(a) (b) (c)

Fig. 3: Transmission map of the hazy image (a) original (b) normalized (c) refined

value is large is likely to exhibit haze. According to He et al. in [9] the formula of estimating transmission map is

$$P = \min_{y \epsilon \phi(x)} (min_{c \epsilon (r,g,b)} (I^c, (y)/A^c)) \tag{5}$$

$$t(x) = 1 - \omega p \tag{6}$$

Where, ω is between 0 and 1. The transmission map which is used in our paper is different from the existing transmission map used by the method of He et al. Here, we have normalized the transmission map by the maximum value in p. This normalization technique reduces the value for further smoothing. This reduction of pixel value removes unnecessary information from the transmission map. The result of the normalized transmission map is shown in Fig. 3(b) where we can see that after normalization the values reduced to a certain level and the refining filter works well on this normalized transmission map.

C. Refine Transmission Map

For refining transmission map we have used nonlinear complex diffusion process [14]. After removing haze, the output image may not provide enough information about edges of the image. Specially ramp edge of image may not be preserved by many haze removal algorithms. So, we need to refine the transmission map in such a way that can smooths the image as well as preserves ramp edges. Nonlinear complex diffusion process does this very well.

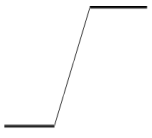


Fig. 4: Ramp Edge

The nonlinear complex diffusion process is performed using Eqs. 7 and 8:

$$I_t = \bigtriangledown (c(I_m(I)) \bigtriangledown I) \tag{7}$$

$$c(I_m(I)) = \frac{e^{i\theta}}{1 + \{\frac{I_m(I)}{k\theta}\}^2} \tag{8}$$

Here, I is the input image in case of denoising, noisy image. Where k is threshold parameter. $e^{i\theta}$ is called unit complex diffusion coefficient. The phase angle θ should be small ($\theta << 1$). In Eqs. 7 and 8, I_m is the imaginary part of the input image. Here, I_t consist of both real part and imaginary part of the image where imaginary part is the second derivative of the real part. The imaginary part is related to localized phase and zero crossing of the image where zero crossing forms thin line. This formula reduces noise preserving ramp features of the images.

As, here the imaginary part is normalized by θ. This process is not affected much by changing the value of θ as long as θ stays small [14]. The transmission map after implying nonlinear complex diffusion process shown in the Fig. 3 (c).

D. Restoration

When the value of transmission map $t(x)$ is close to zero, the direct attenuation $J(x)t(x)$ term can be close to zero. Here, transmission map $t(x)$ is restricted to a lower bound t_0, which means that a small amount of haze is present in every dense haze regions [9]. The output scene is recovered by [9]:

$$J(x) = \frac{I(x) - A}{max(t(x), t_0)} + A \tag{9}$$

E. Post-processing

The restored image may have low contrast after haze removal. So, for increasing contrast of this image, we applied Multiscale Retinex Algorithm (MSR) [15]. MSR is calculated by Eq. 10,

$$MSR_i = \sum_{n=1}^{N} W_n R_{ni} \tag{10}$$

Where, N is the number of scale, W_n is the weight of each scale and R_{ni} can be written as,

$$R_{ni} = [log(J_i(x,y)) - log(F_n(x,y) * J_i(x,y))] \tag{11}$$

Here, J_i is the input image which is the recovered scene from haze model. Here, i defines the color channel and R_{ni} is the retinex output image on the i-th channel. F is the

normalized surround function. This operation is performed on each channel. So, we can write from Eqs. 10 and 11

$$MSR_i = \sum_{n=1}^{N} W_n[log(J_i(x,y)) - log(F_n(x,y) * J_i(x,y))] \tag{12}$$

Here, $F_n(x,y) = C_n e^{-(x^2+y^2)/2(\sigma_n)^2}$ and $e^{-(x^2+y^2)/2(\sigma_n)^2}$ is the gaussian function where σ controls the amount of spatial detail which is retained.

According to Jobson et al. three scales are enough and the value of these three scales are experimentally fixed 15, 80, 250 by them. Here, weights can be identical. MSR ensures the overall contrast and details of local information of the images keeping the good effect.

IV. Experimental Analysis

In this section, we have made qualitative and quantitative comparison among our method and some existing methods. In our experiment, the transmission map is estimated using Eq. 6 described in Subsection III-B with a scaling parameter $\beta = 0.7$. For these images, the airlight is automatically estimated in Subsection III-A. The patch size is set to 3 here. We can notice from the following Figs. 5, 6, 7, 8, 9 and 10 that our method brings out hidden information of a hazy image robustly. The refined transmission map of our method is smoothed with preservation of ramp edges.

It is difficult to compare the performance of haze removal algorithms for general outdoor images. For this purpose, we have taken 6 challenging images for both qualitative and quantitative comparison. These images are also used by Salazar et al. in their proposed method [10].

For getting the experimental result of the images we have used MATLAB R2014a. Our system configuration is : Processor: Intel Core i5, RAM: 4 GB.

A. Qualitative Measurement

In this section, we have analyzed the visual quality of the six hazy images named "hazy wall","forest","hazed2", "mountain", ny17_input and y1_input. Haze covers important parts of these natural images. So, for clear view of all objects there needs to remove haze perfectly.

For "hazy_wall" in Fig. 5, He et al.'s method undercorrects the result and gives over saturated scene. Result produced by Sulami et al.'s method excessively corrects the red wall but fails to improve the scene of the grass. Salazar et al.'s method does not improve output image quality and there remains a little bit haze hidden under the leaf. In this case, our proposed method removes haze robustly and gives pleasing result than other methods.

For "forest" in Fig. 6, Sulami et al.'s method over corrects the haziest region leaving that region dark and undercorrects other regions. Salazar et al.'s method undercorrects the image while He et al.'s method removes haze robustly but gives a saturated scene. From Fig. 6, we see that our proposed method gives a well balanced result with better contrast enhancement than the existing methods.

For "hazed2" image in Fig. 7, He et al.'s method removes haze perfectly but gives over saturated view on the other side the result produced by Sulami et al.'s method contains artifacts. Salazar et al.'s method overcorrects the result as well as the grass are not visible here after haze removal. Fortunately, our proposed method removes haze very well and the edges of the grass can be clearly identified here.

For "mountain" image in Fig. 8, the method of He et al. and Sulami et al. give over corrected result. In the case of Salazar et al.'s method, the output is the result of limited ability to dehaze. Compared to these methods our proposed method gives a well balanced result.

For "ny17_input" image in Fig. 9, Sulami et al.'s method overly saturates the sky region which affects the whole image. Salazar et al.'s method gives an over corrected result. On the other side, He et al.'s method and our proposed method remove haze robustly and give the same quality result but the sky region in our method seems more satisfiable.

For "y1_input" image in Fig. 10, Sulami et al.'s method overcorrects the result while He et al., Salazar et al. and our proposed method remove haze efficiently and give well balanced result.

B. Quantitative Measurement

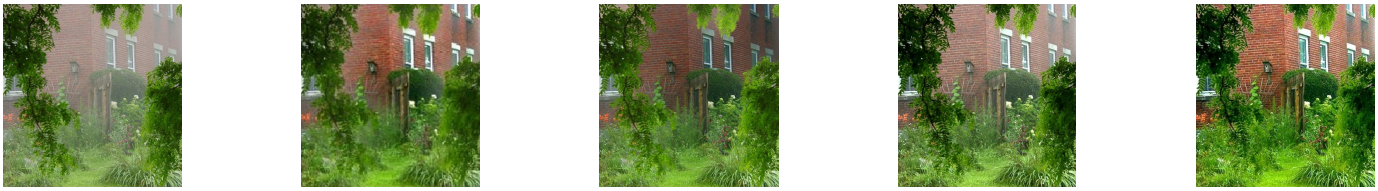
Only observing the visual quality of the image we can not define which method really works well on this type of image. So, we need a quantitative measurement. For this measurement, we consider here two quality metrics such as Contrast to Noise Ratio (CNR) and Root Mean Square (RMS). The detail description is provided in the followings.

1) **Contrast to Noise Ratio (CNR):**

Contrast to Noise Ratio (CNR) is used as a quantitative noise measurement to define image quality. CNR is important when there is indicative influence in an image, such as from haze [16]. We took CNR formula from mathwarks[1]. Here, the larger the CNR the better the visibility of the object in an image. From TABLE I, we see that in the case of "forest", "hazy_wall", "mountain" and "y1_input" images our proposed method gives higher value than other methods. So, we can say that our method gives better quality image than others. In case of "forest" image the CNR value is 89.52 which is greater than other methods. This means that in our method the quantity of noise is less. If we take a look in Fig. 6, we can see that our method provides sharp edges with better reduction of noise that is haze. This is true for other images too.

2) **Root-Mean-Square (RMS):** RMS contrast is the way to define the contrast of the images [17], [18], [19], [20]. It destines standard deviation of pixel luminance in an image which is divided by the average luminance of the whole image.

[1]https://www.mathworks.com/matlabcentral/fileexchange/33529-a-new-visibility-metric-for-haze-images



(a) (b) (c) (d) (e)

Fig. 5: ”hazy_wall” image enhancement (a) Original (b) He et al. (c) Sulami et al. (d) Salazar et al. (e) Proposed



(a) (b) (c) (d) (e)

Fig. 6: ”forest” image enhancement (a) Original (b) He et al. (c) Sulami et al. (d) Salazar et al. (e) Proposed



(a) (b) (c) (d) (e)

Fig. 7: ”hazed2” image enhancement (a) Original (b) He et al. (c) Sulami et al. (d) Salazar et al. (e) Proposed



(a) (b) (c) (d) (e)

Fig. 8: ”mountain” image enhancement (a) Original (b) He et al. (c) Sulami et al. (d) Salazar et al. (e) Proposed



(a) (b) (c) (d) (e)

Fig. 9: ”n17_input” image enhancement (a) Original (b) He et al. (c) Sulami et al. (d) Salazar et al. (e) Proposed



(a) (b) (c) (d) (e)

Fig. 10: ”y1_input” image enhancement (a) Original (b) He et al. (c) Sulami et al. (d) Salazar et al. (e) Proposed

TABLE I: Contrast to Noise Ratio (CNR) of all methods

Image Name	He et al.	Sulami et al.	Salazars et al.	Proposed
forest	56.07	89.25	82.35	89.52
hazy_wall	61.83	65.48	78.38	93.07
mountain	72.63	62.26	65.64	72.65
ny17_input	93.73	97.71	112.19	106.99
y1_input	72.34	57.08	49.72	72.87
hazed2	107.39	90.40	80.53	100.72

TABLE II: Root Mean Square (RMS) of all methods

Image Name	He et al.	Sulami et al.	Salazars et al.	Proposed
forest	0.18	0.15	0.20	0.22
hazy_wall	0.20	0.17	0.23	0.26
mountain	0.24	0.19	0.26	0.25
ny17_input	0.23	0.19	0.26	0.27
y1_input	0.29	0.26	0.29	0.30
hazed2	0.23	0.15	0.29	0.24

If L_i is the luminance of pixel number i, and there are N pixels then we can write, $L_m = \sum L_i/N$.

$$RMS = \sqrt{\frac{\sum (L_i - L_m)^2}{N}} \tag{13}$$

The value of RMS can vary from zero to one. Here, the largest RMS value defines the highest contrast of the image. From the Figs. 5, 6, 9, 10, 7 and 8, we see that our proposed method shows better contrast than other methods. For example, the RMS value of "forest", "hazy_ wall", "ny17_input" and "y1_input" are 0.22, 0.26, 0.27 and 0.30 which is larger than existing Fattal et al., Salazar et al. and He et al.'s methods. TABLE II shows that our method gives the highest contrast image than other methods in most of the cases.

V. Conclusion

This paper presents a new approach to remove haze based on filtering. The proposed method uses an improved transmission map and then refines it by nonlinear complex diffusion process. Computer vision applications such as remote sensing system, surveillance, automated car, etc. can be benefited from our proposed method as it gives enhanced image with clear view of all objects of an image. From the quantitative and qualitative measurement subsections it is clear that in most of the cases our method gives better or similar result compared to existing methods. In future, we will work on this approach for further improvement so that it can work for multiple images instead of single image.

References

[1] A. K. Tripathi and S. Mukhopadhyay, "Single image fog removal using bilateral filter," in *Signal Processing, Computing and Control (ISPCC), 2012 IEEE International Conference on*. IEEE, 2012, pp. 1–6.

[2] S. G. Narasimhan and S. K. Nayar, "Chromatic framework for vision in bad weather," in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, vol. 1. IEEE, 2000, pp. 598–605.

[3] M. Sulami, I. Geltzer, R. Fattal, and M. Werman, "Automatic recovery of the atmospheric light in hazy images," in *IEEE International Conference on Computational Photography (ICCP)*, 2014.

[4] A. Gotchev, "Computer technologies for 3d video delivery for home entertainment," in *Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing*. ACM, 2008, p. 1.

[5] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Instant dehazing of images using polarization," in *null*. IEEE, 2001, p. 325.

[6] R. T. Tan, "Visibility in bad weather from a single image," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.

[7] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski, *Deep photo: Model-based photograph enhancement and viewing*. ACM, 2008, vol. 27, no. 5.

[8] J.-P. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 2201–2208.

[9] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 12, pp. 2341–2353, 2011.

[10] S. Salazar-Colores, I. Cruz-Aceves, and J.-M. Ramos-Arreguin, "Single image dehazing using a multilayer perceptron," *Journal of Electronic Imaging*, vol. 27, no. 4, p. 043022, 2018.

[11] R. Fattal, "Single image dehazing," *ACM transactions on graphics (TOG)*, vol. 27, no. 3, p. 72, 2008.

[12] H. Koschmieder, "Theorie der horizontalen sichtweite," *Beitrage zur Physik der freien Atmosphare*, pp. 33–53, 1924.

[13] L. He, J. Zhao, N. Zheng, and D. Bi, "Haze removal using the difference-structure-preservation prior," *IEEE Transactions on Image Processing*, vol. 26, no. 3, pp. 1063–1075, 2017.

[14] G. Gilboa, Y. Y. Zeevi, and N. A. Sochen, "Complex difusion processes for image filtering," in *International Conference on Scale-Space Theories in Computer Vision*. Springer, 2001, pp. 299–307.

[15] D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image processing*, vol. 6, no. 7, pp. 965–976, 1997.

[16] H. Jiang, N. Lu, and L. Yao, "A high-fidelity haze removal method based on hot for visible remote sensing images," *Remote Sensing*, vol. 8, no. 10, p. 844, 2016.

[17] S. Rahman, M. M. Rahman, K. Hussain, S. M. Khaled, and M. Shoyaib, "Image enhancement in spatial domain: A comprehensive study," in *2014 17th International Conference on Computer and Information Technology (ICCIT)*. IEEE, 2014, pp. 368–373.

[18] S. Rahman, M. M. Rahman, M. Abdullah-Al-Wadud, G. D. Al-Quaderi, and M. Shoyaib, "An adaptive gamma correction for image enhancement," *EURASIP Journal on Image and Video Processing*, vol. 2016, no. 1, p. 35, 2016.

[19] K. Hussain, S. Rahman, M. M. Rahman, S. M. Khaled, M. Abdullah-Al Wadud, M. A. H. Khan, and M. Shoyaib, "A histogram specification technique for dark image enhancement using a local transformation method," *IPSJ Transactions on Computer Vision and Applications*, vol. 10, no. 1, p. 3, 2018.

[20] F. M. Amil, M. M. Rahman, S. Rahman, E. K. Dey, and M. Shoyaib, "Bilateral histogram equalization with pre-processing for contrast enhancement," in *2016 17th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*. IEEE, 2016, pp. 231–236.

Underwater Image Enhancement Based on Fusion Technique via Color Correction and Illumination Adjustment

Shahjadi Silma Subah
Department of Computer Science and Engineering
Bangladesh Army University of Engineering & Technology
Natore-6431,Bangladesh
nikitahuda1234@gmail.com

MD. Ashfaqul Islam
Department of Electrical and Electronic Engineering
Bangladesh Army University of Engineering & Technology
Natore-6431, Bangladesh
niloy064@gmail.com

MD. Mohyminul Islam
Department of Computer Science and Engineering
Bangladesh Army University of Engineering & Technology
Natore-6431,Bangladesh
shisierbauet01@gmail.com

Abstract— **Underwater image is affected by absorption and scattering when light propagates through water; resulting in poor visibility, reduced contrast, color attenuation. So before going into any restoration processing activity of an image, it is mandatory to enhance the image first. In this paper, a fusion based technique is introduced, where the method is to use wavelet decomposition. The proposed method has two inputs enhanced from the original underwater image. One is color corrected image and other is illumination adjusted image. Illumination adjustment is done by adaptive histogram equalization, gamma correction and filtering in the proposed system. Histogram equalization and adaptive histogram equalization techniques are used for color correction on the adjusted image. These modified images are then fused based on their maximum coefficient values and the resulted output shows promising result improving the haze of the underwater image. This is a very simple method and the resulted image is clearer than the hazy image.**

Keywords— Underwater image, Color illumination, Color enhancement, Wavelet decomposing, Fusion.

I. INTRODUCTION

In the last few decades, underwater image processing has received considerable attention in ocean engineering. Due to absorption and scattering [1] underwater images (UW) can face different types of difficulties resulting in degradations including color cast, noise, due to floating particles, low contrast, skewing and blurring [2]. In underwater it is hard to notice color, size of an object due to attenuation and mileage [3]. The objective of image enhancement is to recover the explicability of information in underwater images for human viewers, or to give corrected image for other automated image processing techniques [2]. Image enhancement techniques applied for in-air images cannot be applied to underwater images because of different optical properties of air and water. Non-uniform illumination [4] correction problem is ignored by many researchers, degradation by forward scattering is not attempted by most of the researchers [5]. Most of the dehazing techniques by researcher are complex and computation is heavy. So to enhance the underwater images it is necessary to develop algorithm that can correct all the four types of degradation, non-uniform illumination [4] and contrast degradation (caused by back scattering), blur (caused by forward scattering) and diminished colors (caused by scattering and absorption).
Rest of the article is indexed as follows. Section II presents some related article on Underwater Image Enhancement, proposed system architecture is described in section III. Section IV hold the conclusion and future work of this paper.

II. PREVIOUS RESEARCH WORK

In the field of underwater image enhancement, a lot of works have been done in the past to modify the haziness of images taken underwater. Authors in [6] estimated fusion based algorithm with two inputs, one is color corrected image and other is contrast enhanced image. Writers in [7] implemented wavelet based fusion technique to overcome color and contrast issues in underwater images. They employed CLAHE and histogram stretching techniques for contrast enhancement and color correction to address low contrast and color attenuation of the hazy images. Writers in [8] derived different images each having a significant feature by processing the original degraded image and then fuses all the images to get final output image which is a blend of all the features of processed images in it. Authors in [9] proposed a fusion based underwater image enhancement solution built on wavelet decomposition and multiscale fusion principle. Multiscale fusion principle is applied to the low frequency and high frequency components. Quality and reliability of the image processing is affected because of the floating particles of the water and that's why color correction is needed. Writers in [10] proposed a new method to enhance the underwater image depending on Dark Channel Prior. Authors of [11] enhanced the low quality images using unsupervised color correction method. The proposed approach is based on color balancing contrast correction of RGB color model and contrast correction of HIS color model. Writers in [1] proposed image enhancement method via color correction and illumination adjustment. For color correction they used simple normalization method and for illumination adjustment component, the Retinex theory is exploited.

III. PROPOSED SYSTEM ARCHITECHURE

Proposed system architecture mainly divided into three sections. Section A portrays illumination adjustment part, section B briefly describe color correction part and section C presents the fusion processing part.

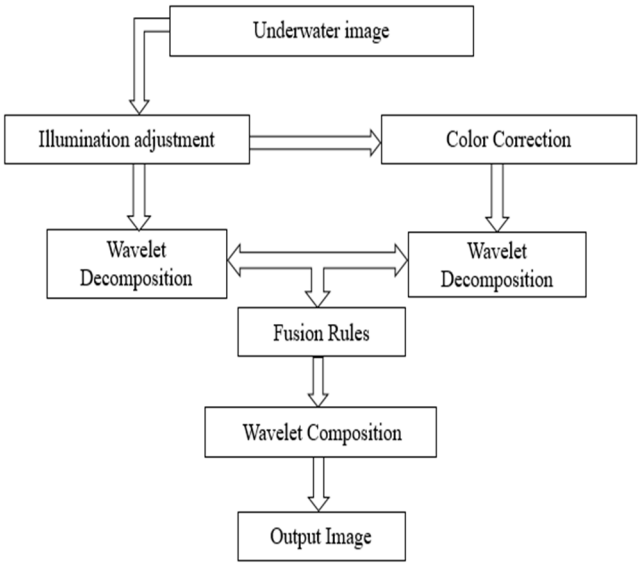


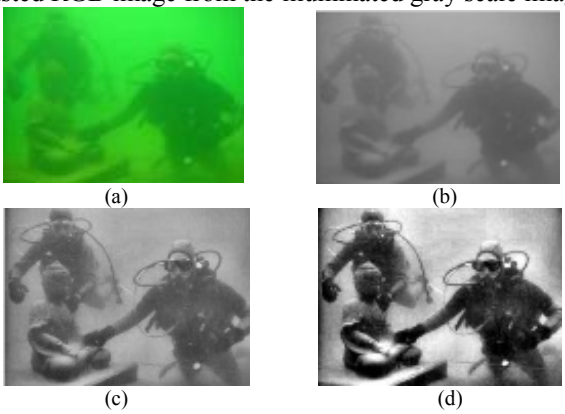
Fig. 1. Architecture of the proposed system.

A. Illumination Adjustment

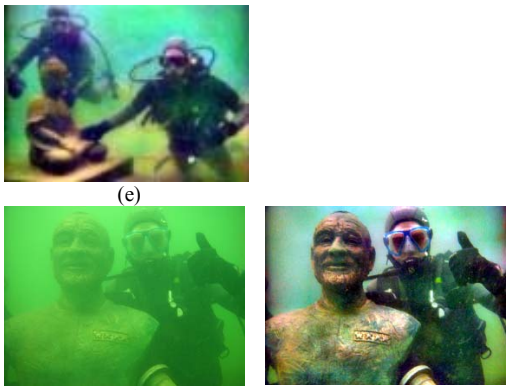
Illumination of an underwater image appears because of propagation of light in water. Illumination of an M×N dimension underwater image can be denoted by illuminant, t_c [3],

$$t_c = k_{|\sum_{j=k}^{L-1} h_c(j) > \frac{pMN}{100} AND \sum_{j=k+1}^{L-1} h_c(j) < \frac{pMN}{100}}$$

To remove illumination from underwater image in this proposed method the original image is first converted into gray scale image. Then Adaptive Histogram Equalization is applied on the gray scale image to manage the local contrast of the pixels from the neighboring areas [10]. The background from the AHE image is then subtracted and gamma correction is applied on the image to correct the images luminance. Median filter is then used on the image to remove any unwanted noise in the image. The process from background subtraction to filtering the image is done again two times (total three time) in order to convert back to the RGB color space of an image. The three resulted filtered images are then combined to produce the illumination adjusted RGB image from the illuminated gray scale image.



(a) (b) (c) (d)

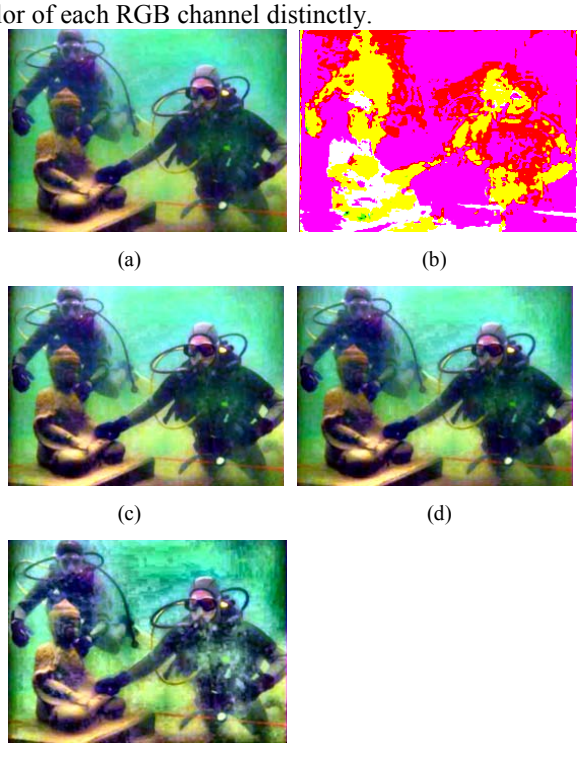


(e) (f) (g)

Fig. 2. (a) Original image 1, (b) Gray scale image for image 1, (c) Adaptive Histogram Equalized image for image 1, (d) Illumination adjusted gray scale image for image 1, (e) Illumination adjusted output image for image 1, (f) Original image 2, (g) Illumination adjusted output image for image 2

B. Color Correction

For color enhancement the input image is the output image of illumination adjustment. In underwater, images are affected by low contrast due to scattering. To enhance the color, illumination adjusted image is first converted to RGB space to HSV space. Hue, Saturation and Value are then processed with max luminance to highlight the higher visibility region in the output. Then histogram equalization is used on the processed image. Histogram equalization which is a useful technique for adjusting image intensities for to enhance contrast. But sometimes it spoils the unhazy parts and gives indiscriminate result. So, Adaptive histogram equalization [9] is then used on the histogram equalized image of the input image of color enhancement. AHE works by taking into account only small regions and based on their local cdf, enacts contrast enhancement of these regions. This technique executes histogram equalization on the intensity values of individual frames of the image and modifies the color of each RGB channel distinctly.



(a) (b) (c) (d) (e)



(f) (g)

Fig. 3. (a) Input image 1 for contrast enhancement (Illumination adjusted image), (b) HSV image for image 1, (c) Histogram equalized image for image 1, (d) AHE image for image 1, (e) Contrast enhanced image for image 1, (f) Input image 2 for contrast enhancement, (g) Contrast enhanced image for image 2

C. *Fusion Process*

Wavelet transforms are essentially extensions of the idea of high pass filtering. The process of applying the DWT can be represented as a bank of filters. At each level of decomposition, the signal is split into high frequency and low frequency components; the low frequency components can be further decomposed until the desired resolution is reached. The wavelet transform is designed to get good frequency resolution for low frequency components (average intensity values of the image) and high temporal resolution for high frequency components (edges of the image). Illumination adjusted image and contrast enhanced image are decomposed into low frequency and high frequency components. Since the wavelet transform is applied in both dimension, the low pass filter and high pass filter is applied on the rows first and then the columns. When low pass filter is applied on the rows it gives horizontal approximation value and applying high pass filter gives horizontal detail. The sub signal produced from low filter have a highest frequency equal to half of the original. Low pass filter and high pass filters is again applied on the columns of the horizontal approximation and horizontal detail resulting in four sub images [2]: approximate image, vertical detail and horizontal detail, diagonal detail respectively. This is one level decomposition of the input image as shown in figure 4 [8].

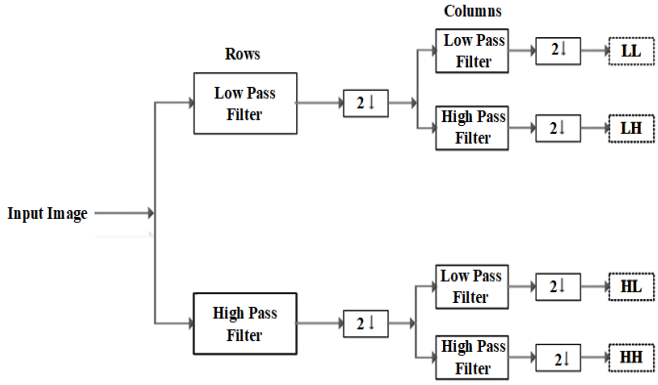


Fig. 4. One-level-2-dimension wavelet based image decomposition [8]

The approximate image which is the decomposed result of the first level becomes the input image and gets decomposed into the wavelet coefficient maintaining the same process as the first level in the second level decomposition. In the proposed method the decomposed wavelet coefficients of both illuminations adjusted image and color enhanced image is used and fused together using the maximum value of the approximation and detail coefficients.

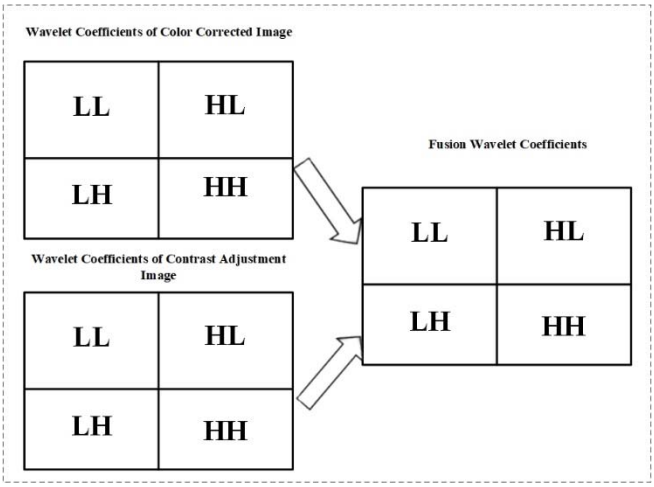


Fig. 5. Two-level-2D wavelet based image decomposition [8]

Every input image is decomposed into its coefficient and inversely composed respectively using wavelet transform to get the modified image as output. The decomposition in wavelet transform is done by down-sampling and up-sampling is used for the inverse composition.

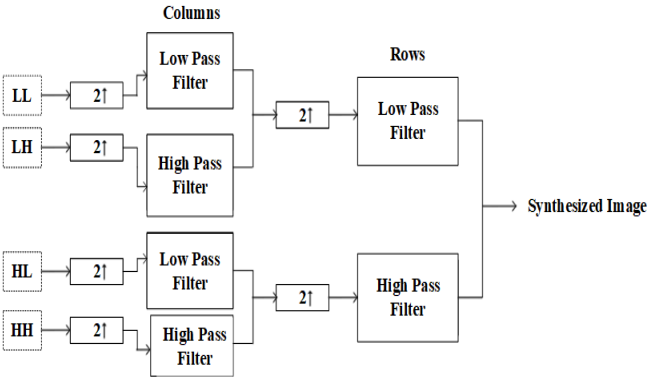


Fig. 6. One-level-2d wavelet based inverse composition [8]

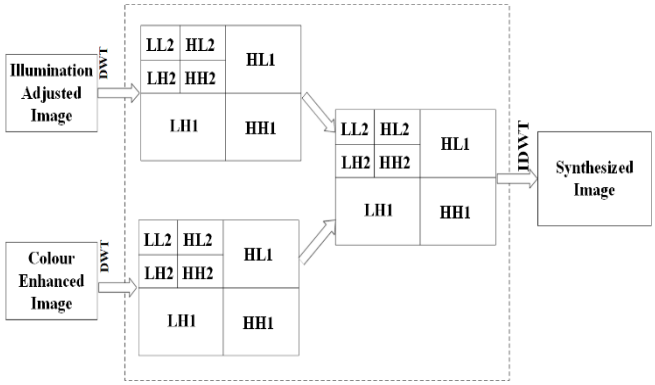


Fig. 7. Two-level-2D decomposition, fusion of coefficients and image synthesizing [8]

In figure 7, the output image of illumination adjusted image and color enhanced image is the input for DWT decomposition. Thereafter inverse transform is applied on the coefficients of its decomposition and fused together using the maximum values of the coefficients.

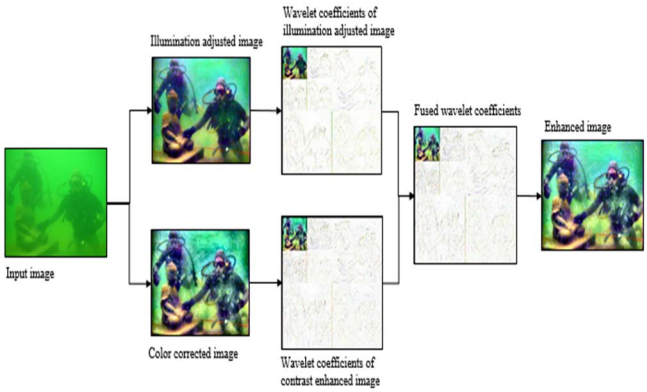
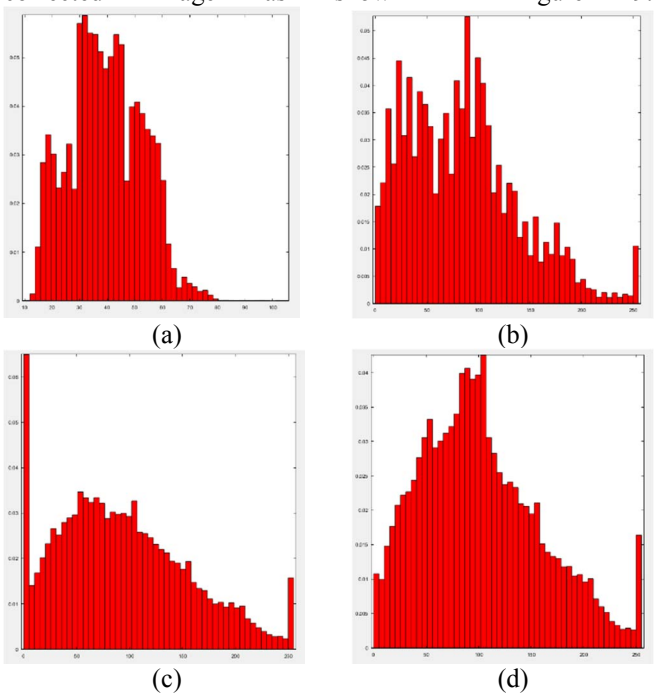


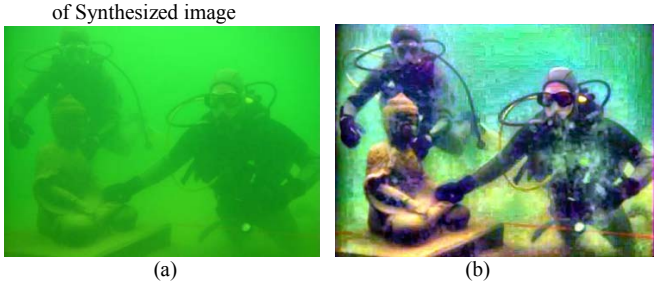
Fig. 8. Two-level-2D decomposition, fusion of coefficients and image synthesized of the proposed method

Histogram of an image illustrates a graph that can be used to appraise the quality of an image by comparing the values of an image before and after enhancement [6]. The histogram of underwater input image is darker and histogram of the enhanced output image has brighter pixel values which is the combination of illumination adjusted image and color corrected image as shown in figure 9.

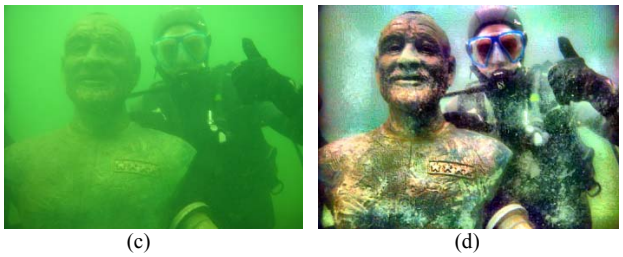


(a) (b) (c) (d)

Fig. 9. (a) Histogram of orginal image, (b) Histogam of Illumination Adjusted image (c) Histogram of Color Enhanced image (d)Histogram of Synthesized image



(a) (b)



(c) (d)

Fig. 10. (a) Input underwater image 1, (b) Synthesized image for image 1, (c) Input underwater image 2, (d) Synthesized image for image 2

TABLE I. Numerical Results of performance measurement parameters: Entropy for image (d) and comparison with Ancutti et al. [13] and Singh et al. [13] Method

Image	Entropy		
	Ancutti et al. [13]	Singh et al. [13]	Proposed
d	6.8683	7.5112	7.9053

IV. RESULT AND DISCUSSION

This paper proposed underwater image enhancement based on wavelet based fusion technique which will ensure the effective improvement of visibility of degraded underwater image. The turbid water images are victim of attenuation in water, blurring and scattering. The proposed method faces the problems head on correcting the color and illumination of the image to take two input image and fuses them together. This method generates admissible result and improves clarity for all the test images. The amount of information that stays in the output image is entropy. The amount of distortion is less in the final image if the value of entropy is less. Table 1 clearly shows that the resultant images obtained from our method has higher Entropy value as compared to Ancuti et al. and Singh et al. as desired for a good quality image. A main limitation of fusion based method is represented by the fact that the noise contribution may be amplified significantly with the depth causing undesired appearance of the distant regions. We can improve this in future. In future, the work will extend to underwater inspection and observation tasks, particularly in vision system for aquatic robot inspection and IOT.

ACKNOWLEDGMENT

We the authors are very obliged to Department of Computer Science and Engineering and Department of Electrical and Electronic Engineering, Bangladesh Army University of Engineering and Technology for the remarkable support and guideline during the process of this work.

REFERENCES

1. Wenhao Zhang 1, Ge Li2, Zhenqiang Ying 3, “A New Underwater Image Enhancing Method via Color Correction and Illumination Adjustment”, IEEE, Dec. 10 - 13, 2017
2. R. SwarnaLakshmi1, B. Loganathan2, “An Efficient Underwater Image Enhancement Using Color Constancy Deskewing Algorithm”, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 3, Issue 8, August 2015
3. Rashid Khan, “FUSION BASED UNDERWATER IMAGE RESTORATION SYSTEM”, NICOSIA, 2014

4. Huimin Lu1, Yujie Li2, Yudong Zhang3, Min Chen4, Seiichi Serikawa1, Hyoungseop Kim1, “Underwater Optical Image Processing: A Comprehensive Review”, 2017
5. S. S. Sankpal1 and S. S. Deshpande2, “A review on image enhancement and color correction techniques for underwater images”, Research India Publications, Volume 9, Number 1 (2016) pp. 11-23
6. Codruta O. Ancuti, Cosmin Ancuti, Christophe De Vleeschouwer, Philippe Bekaert, “Color Balance and Fusion for Underwater Image Enhacement”, IEEE, Vol. 27, No. 1, January 2018J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73
7. Amjad Khan, Syed Saad Azhar Ali, Aamir Saeed Malik, Atif Anwer, Fabrice Meriaudeau, “Underwater Image Enhancement by Wavelet Based Fusion”, IEEE,2016
8. Ritu Singh, Manitosh Biswas, “Contrast and Color Improvement based Haze Removal of Underwater Images using Fusion Technique”, IEEE, Sep. 21-23, 2017
9. Yafei Wang, Xueyan Ding, Ruoqian Wang, Jun Zhang, Xianping Fu, “Fusion-based underwater image enhancement by wavelet decomposition”, IEEE, 2017
10. Min Hao, Chao Chen, “Enhancing Underwater Image by Dark Chanel Prior and Color Correction”, IEEE, 2016
11. Kasif Iqbal, Micheal odetayo, Anne James, Rosalina Abdul Salam, “Enhancing the low quality images using Unsupervised Colour Correction Method”, Conference paper, October 2010
12. A. M. Aibinu, M. I. Iqbal, M. Nilsson and M. J. E. Salami “A New Method of Correcting Uneven Illumination Problem in Fundus Image”, Proceedings of the International Conference on Robotics, Vision, Information and Signal Processing ROVISP2007
13. Ritu Singh, Dr. Manitosh Biswas, “Adaptive histogram Equalization Based Fusion Technique for Hazy Underwater Image Enhancement”, IEEE, 2016
14. Amjad Khan, Fabrice Meriaudeau, Syed Saad Azhar Ali, Aamir Saeed Malik, “Underwater Image Enhancement and Dehazing Using Wavelet Based Fusion for Pipeline Corrosion Inspection”, IEEE, 2016

Automatic Examinee Validation System using Eigenfaces

Nuruzzaman Faruqui
Institute of Information Technology
Jahangirnagar University
Dhaka, Bangladesh
faruquizaman27@gmail.com

Mohammad Abu Yousuf
Institute of Information Technology
Jahangirnagar University
Dhaka, Bangladesh
yousuf@juniv.edu

Md. Fazlul Karim Patwary
Institute of Information Technology
Jahangirnagar University
Dhaka, Bangladesh
patwary@juniv.edu

***Abstract*—In most of the developing countries the seats of educational institutions are limited and barley 30% of the total eligible candidates get the opportunity to get admitted after facing competitive admission test. The scenario is worse in the job market. The growing graduation to unemployment ratio has become a major concern in developing and underdeveloped countries. As a result, job recruitment tests have become more competitive than ever. In order to survive in these types of competitive examinations, many candidates try to take unethical approaches. Because of large number of candidates, exam administrators cannot validate each and every candidate one by one. Moreover, modified image in registration paper makes it difficult to identify valid examinee. This is the weakest point of examinee validation system of competitive examination and significant number of candidates try to pass the examination by exploiting this weakness. The easiest way to exploit this weakness is to hire and send proxy examinee who has higher possibility of passing the exam instead of actual examinee. In this paper, we are proposing an automatic, fast, efficient and accurate yet simple to implement examinee validation system using eigenfaces. Our system will make it almost impossible for a proxy examinee to enter the exam-center.**

Keywords—examinee validation, face recognition, eigenface, segmentation, innovative registration.

I. INTRODUCTION

The rate of cheating in the competitive examination through proxy examinee is rising at an alarming rate [1]. It deprives the meritorious candidates from getting the positions they deserve. And if the cheaters do not get caught, they get employed in important positions. In the long run the productivity of the organizations they join, educational institutions they get admitted in end up in poor productivity. The purpose of competitive exams is to find out the appropriate candidates. If wrong person sits in important position through unethical means, they will not perform their responsibility as the proper candidate would have done. It is very dangerous for a nation, especially for developing countries. In developing countries excessive population creates extra pressure on both educational institutions and job market. Because of having limited seats, institutions can barely take 30% of total eligible candidates. The condition in job market is worse than this. The competition has become so tough that more candidates are being engaged in cheating in competitive examinations. At the same time, it is not possible for exam-administrators to verify every individual because of large number of candidates. This is where the weak point is and tendency of getting in exploiting this weak point is becoming acuter day by day.

Automatic examinee validation system is the most effective solution to proxy examinee problem. However, implementing such system is time consuming and challenging task. Using convolutional neural network, it is possible to validate examinee with 96.2% accuracy [2]. To gain such accuracy five images are needed of each individual. In the year 2017 against 2024 vacancies 343,532 candidates attended examination conducted by Bangladesh Public Service Commission (BPSC) [3]. In similar situation, taking five images of each candidate for training will yield 1,717,660 training images. Processing this large volume of image data and training the network using them is not a feasible solution. However, through linear segmentation of training data according to the number of exam-center, it is possible to distribute the computational task among multiple systems. In this approach the training time can be reduced significantly. Although it sounds feasible yet distributing the trained network, permanently setting up validation equipments in exam-centers, training new network before every examination make it too complex and expensive to implement.

We need a system that works in real-time and does not require training like deep neural network or convolutional neural network. The system has to be very simple, fast, accurate and instantly deployable. We are proposing an automatic examinee validation system using eigenfaces which is very simple, easy to implement and works very fast in real-time. Our system does not require any training. Theoretically, our system will work fast enough even for image database larger than 5×10^5 images. No matter how many images are there in the database, our system will always compare the test image with 10 training images and will yield over 97.88% accurate result.

The remaining part of the paper is as follows - relevant works are reviewed in section 2. The proposed method has been presented in section 3. In section 4, two ways have been shown which can be used to physically implement the validation system. The experimental result of the proposed method has been demonstrated in section 5. Finally, we concluded the paper in section 6.

II. BACKGROUND

The normalized distance between facial features and ratios among the features can be used for automatic face recognition. Bledsoe [4] and Kanade [5] have used eyes, nose, mouth and head outline as facial features. The existence of facial features, their relative positions and size are used to classify face and non-face classes. By comparing the features of images belong to face class, new faces can be recognized if they are already in existing face class. However, this approach cannot be extended to multi-view face recognition. Moreover, Diamond and Carey [6] have shown that individual facial features are inefficient for adult human face recognition. In addition, the

feature points are not prominent for all races. As a result, there are possibilities of having face images in non-face class.

Instead of classifying images based on individual facial features, face images can be considered as 2D matrix where 2D characteristics represent faces. Turk and Pentland [7] have developed eigenvector-based face detection and recognition system which can learn to recognize new faces in an unsupervised manner. Diverting from geometry or high-level visual characteristics, they have taken information theory-based approach. In this approach they have transformed images into characteristic feature images named eigenfaces. These features are the principle components of training set of face images. Then, projecting new images on the eigenfaces and comparing the position of new eigenface with existing eigenfaces, they have performed the recognition. This approach remains effective for multi-views such as profile face, view from 45° angle and straight view. Moreover, this recognition method is fast, simple and robust.

However, the size of the face space imposes challenges on eigenface construction. Sirovich and Kirby have represented their idea of reducing face space using principle component representation [8]. They have said face images can be reconstructed by storing small collection of weights of each face. However, a set of standards which will uniquely define how the weights will be distributed are required as well. Combining the weights and the standards, it is possible to reconstruct the actual face image from small image space. They have also suggested that if a face can be reconstructed from the weights and standards, a face can be recognized using the same principle as well. If a face image gives similar weights and standards as another face image, then both of the face images will be the same. In other words, weights and standards of one face image can be used to recognize the same face image. If the face image of the same person varies in dimension, the standards which are actually vectors adjust the weights proportionally. As a result, the size variation cannot affect the recognition accuracy.

Simple implementation, fast recognition and robustness of face recognition method of Turk and Pentland paved the way of using eigenface based system in application level. Wagh, Thakare, Chaudhari, and Patil have shown an application of detecting attendance fraudulent activities in class room using eigenfaces [9]. To meet the acceptance level at application level, only eigenfaces are not enough. Wagh et al. in their paper have combined luminance variant reduction technique to reduce impact of light variation among testing and training image. They have also used Viola-Jones algorithm to detect the face in order to get specific face space. After primary recognition a cross checking by students' image database have enhanced the accuracy of their system.

Eigenface-based face verification systems are enough secured for non-financial and non-confidential national defense related application. J. Galbally, J. Fierrez, J. Ortega-Garcia, C. McCool and S. Marcel have shown security vulnerabilities of Eigen-face based face verification system by hill-climbing attack [10]. The success rate of hill-climbing attack is more than 85%. It does not require a large number of training image to initiate the attack. It clearly suggests that precaution must have to be taken in eigenface-based verification system. However, hill-climb attack cannot impact effectively in real-time validation system. Which leaves a scope of implementing real-time automatic eigenface based validation system without the risk of having hill-climb or brute-force attack on the security system.

III. Proposed Method

Automatic examinee validation system using eigenface is a two-step process. The first step is related to registration system and the second step is the validation system. With the existing examination registration system, it is not possible to effectively validate examinees automatically. The training image required for validation must be acquired during registration process. In existing system, candidates register through online registration form for examinations and upload scanned copies of their photos. With a little modification of scanned images using photo editing software like Photoshop, it is very easy to alter the images in registration card to send proxy examinee instead of real examinee with very little possibility of getting caught. To leave no scope of sending proxy examinee we have to change the registration system first.

A. Proposed Registration System

With the existing online registration system, we propose an additional module which is a real-time image capturing unit. Instead of uploading the scanned copy of photo, the candidates will capture 10 photos according to given instructions. The device using which the candidates will register must have to have webcam for this registration process. Fig. 1 shows the interface of registration process. Candidates have to take 10 pictures imitating the poses recommended in the instruction. The image capturing module uses Viola-Jones face detention algorithm [12] to detect the faces and crop the face rectangle by 92×112 pixel. The images are labeled from 1 to 10. Once the candidate confirms registration, a directory is created based on choice of exam center (e.g. Dhaka) and a subdirectory is created under this newly created directory named according to the registration number of the candidate. In this way, candidates of different exam centers will have different directories. In those directories, the subdirectories will contain the images of the candidate. Each subdirectory represents individual registered candidate. As each registration number is unique, each candidate will have their images in specific subdirectory. These images are the training image of our system.

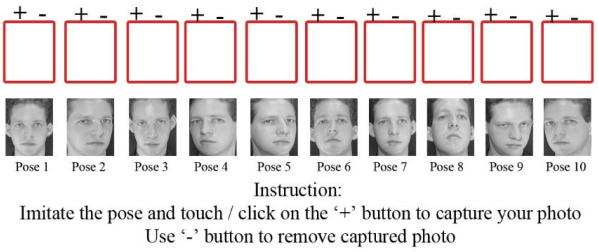


Fig. 1. Photo Capturing During Registration.

B. Examinee Validation

We have used eigenfaces for examinee validation. Even though convolutional neural network is much better at image recognition, we cannot use it in the validation system we are

proposing. Because the dataset our system has to handle is very large, which requires huge amount of time to train the network. And for every new examination, new network has to be trained. Which makes convolutional neural network unrealistic for the system we are building. Our system is based on the work of Turk and Pentland [7]. They have shown how eigenfaces can be used for face recognition. We calculated the eigenfaces in exactly the same way they did. However, we have modified the recognition phase to meet the requirement of our system. We have reduced the size of the face images by calculating and removing the mean. Instead of simply classifying as known and unknown, we have added an access method for validated candidate. The method proposed in [7] was near-real-time whereas our implementation works in real-time.

1) Prioritizing Eigenface for Recognition

Predefined measurement of facial features is not sufficient for face recognition. Instead of looking at face images from high level visual features it is much efficient in focusing from information theory perspective. Eyes, nose, lips, and hair are global feature of a face. From information theory perspective they carry very small amount of information as the probability of having these features is very high. Our purpose is to recognize faces. That is why, the variations among testing face and training faces are the major concern. The face image is a 2D matrix that carries information. From information theory perspective, our task is to extract the information from the 2D matrix and encode it efficiently. Then we will encode the training image in the same way and later we will compare among encoded pieces of information.

We have computed the variations among the face images stored in database. We are not considering those images as face images but 2D matrix those carry information. From these perspective, facial features described in [5] are simply information. There is no distinction among them from high level visual perspective (i.e. nose is not nose but some pixel variations). Once the information of the 2D matrices are collected we compare among them and identify them based on the variation of information they present. According to [8] the maximum varying regions are the principal component. They call it eigenvectors of the covariance matrix. When we represent these eigenvectors as images, it produces face images, indifferent to human eyes, but meaningful in computation in terms of variations among multiple faces.

All of the faces stored in the registration database can be represented in terms of linear combination of the eigenfaces. We have 10 images for each individual. That means there are 10 eigenfaces for each registered candidate. For our application, it is not mandatory to match all 10 of the eigenfaces with 100% accuracy with the testing image. Out of 10 eigenfaces, we compute the best eigenface based on the information variation of that face with other faces in the same directory (e.g. Dhaka). The number of training data is large even though we have classified them according to location of exam-center during registration process. As a result, reducing computational time is mandatory. Instead of taking all 10 of the eigenfaces, taking the eigenface that exhibit the highest variation when compared with other candidates' eigenface is maintained with highest priority for recognition.

Algorithm 1 Face recognition algorithm

Input: X: Exam Center; Y: Registration Number; Z = Input Image
1: **if** X = true **then** access X
  else terminate
2: **if** Y = true **then** access Y
    Image Array ← $Load\ training\ images$
  else terminate
3: **for** i = true
4: Processed Image[i] ← $Image\ Array[i] - mean(Image\ Array[i])$
5: Eigenface Array← $Eigenface\ Function\ (Processed\ Image)$
6: **end for**
7: Processed Input ← $Eigenface\ Function(Z - mean(Z))$
8: **for** i = true
9: Euclidean Distance ← $Processed\ Input - Eigenface\ Array\ [i]$
10: **if** Euclidean Distance > threshold **then** allow access to exam-center
  else deny access to exam-center

2) Algorithm for Face Recognition and Validation

The procedure or the face recognition and validation are shown in algorithm 1. As stated earlier, our system is built on [7]. We modified the approach taken in [7] and converted the method into an applicable validation system. The algorithm is initiated by user inputs which are actually the information of candidates. The candidates select the names of exam-centers, registration numbers and the system captures their frontal faces. If the inputs are correct, the algorithm executes the further instructions and return the result (line 10 of algorithm 1). It allows only valid candidates to enter the exam-center and prevent others from getting accesses.

3) Calculating Eigenfaces

Training images are 2D matrix consisted of 2-dimensional array having 92 and 112 values respectively. Each of the matrix has 10304 elements. The entire space does not represent the face. Moreover, we do not need the entire face. We are looking at the image from information theory perspective. To reduce the size of the image, we need to remove the information which will not be used in face recognition. By calculating and removing the mean from each image, we can reduce the size of the image. From matrix perspective, the value of common element will be replaced by zero. From visual perspective, the images will carry information that distinguishes them from one another.

Let the training set of face images of one candidate is be $F_1, F_2, F_3, \ldots F_{10}$,. To further reduce the size of face space, we calculate the average of the faces.

$$\text{Average face, } A = \frac{1}{10}\sum_{n=1}^{10} F_n \quad (1)$$

Now we have the average face. Using the average face, we can calculate the deviation matrix of other faces of the candidate by subtracting the A from each face.

$$\text{Deviation, } \phi_i = F_i - A \quad (2)$$

Then using principle component analysis (PCA) we find a set of vectors and values which form the eigenfaces. We followed the method of using PCA of [7] to calculate the eigenfaces. According to Turk and Pentland the eigenvalues (λ_k) and eigenvectors (μ_k) together form the covariance matrix,

$$C = \frac{1}{10}\sum_{n=1}^{10} \phi_n \phi_n^T = II^T \quad (3)$$

Here I is the array of deviation matrix. The covariance matrix is still large as it is consisted of 2 arrays containing 92 and 112 elements respectively. Computing over this large matrix is expensive and time consuming. However, [7] shows that eigenvectors can be determined my solving smaller M by M matrix and then combining them linearly. This makes the calculation much faster. At the end, the background is removed which gives us the eigenfaces only.

4) Validation using Euclidean Distance

Eigenfaces contain the pattern that distinguishes faces from one another. They are spanned in M' dimensional subspace where the original image is an 92×112 image space. The eigenvector M' which is associated with largest eigenvalues are chosen as the eigenface. Now for a new image, the face is detected using Viola-Jones algorithm [11] and cropped with a rectangle of 92×112 pixel. Then it is transformed into eigenface consisted of an eigenvector and eigenvalue. The weights of eigenfaces form vectors. These vectors represent the face images. The Euclidean distance between the weights of vectors of input image and training image describes if the new image belongs to the training image set or not. If the new face belongs to the training image located inside subdirectory named after the candidate's registration number, we consider the person as valid candidate. Otherwise the candidate is considered a proxy candidate.

IV. Implementation

The performance of the validation system depends implementation architecture. Here arises the issue of performance vs implementation cost trade-off, which leads to two different architecture – (i) Single Validation Unit (SVU) architecture and (ii) Parallel Validation Unit (PVU) architecture. The SVU architecture is preferred where implementation cost is the major concern. And the PVU architecture is preferred where performance is the major concern. One advantage of PVU architecture is here the implementation cost and performance can be adjusted by changing the number of units. On the other hand SVU architecture does not offer the flexibility of adjusting the performance vs cost parameter.

A. SVU Architecture

In SVU architecture there is one validation unit at the entrance of the exam perimeter (fig. 2). The examinees are validated one by one. Valid examinees get the access to the exam perimeter. The suspected examinees generate notifications in invigilators' office with their names and registration numbers. These examinees do not get the access to exam perimeters. Instead they are directed to invigilators' office for further investigation.

The SVU architecture is simple and easy to implement. However, it is suitable for small number of students. When the number of students increases, the total validation time increases linearly. In this case total time to validate S_t number of students V_t will be,

$$V_t = S_t \times V_i \tag{4}$$

Here V_i is average time to validate each examinee. Equation 4 shows that the total time required to validate all of the examinee is directly proportional to the number of total students. That is why for public exams where large number of candidates attend exams, SVU architecture fails to exhibits satisfactory performance. However, from network security perspective this approach is much secured. Because of having single unit, the validation server and invigilators' office can be established nearby. As a result, a wired communication is possible between server and the validation unit. It ensures security from attacks like eavesdropping or data manipulation by man-in-the-middle attack.

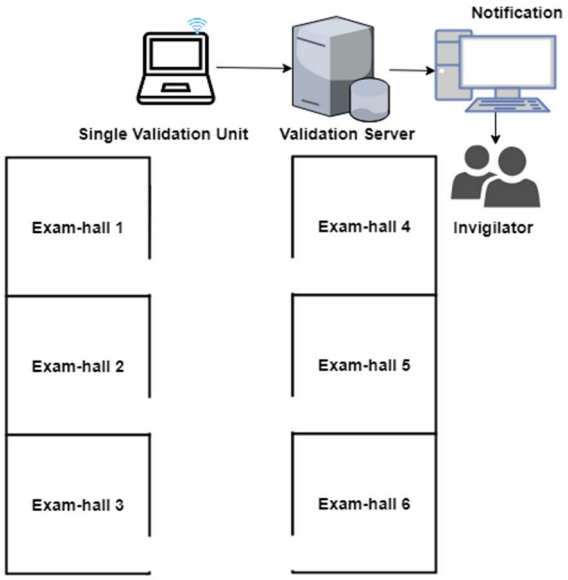


Fig. 2. SVU architecture

B. PVU Architecture

In PVU architecture (fig. 3) multiple SVU work simultaneously. In this model, the examinees can enter the exam perimeter without any validation. However, before entering the exam-hall they have to go through the validation process. Here, the valid examinees get access to respective exam-halls. And the suspected candidates are directed to the invigilators' office. In this case with username and registration number, the exam-hall number are sent in the notification.

The implementation of PVU system is not as simple as SVU system. The complexity and cost increases with the increase of validation units. However, the total time required to validate all of the examinees reduces linearly with the increases of number of validation units. Here, the total time required to validate S_t number of students V_t will be,

$$V_t = \frac{1}{U} \times (S_t \times V_i) \tag{5}$$

Here U represents the total number of validation unit. The more the validation unit the less the time required to validate in PVU method. Equation 5 shows that it is possible to reduce the total validation time by increasing the validation unit. That means the performance of the system can be enhanced. However, because of having multiple units in different spatial locations, wireless communication among units and server is the most convenient communication method. In this case, the

network security becomes a vital issue. Intruders can eavesdrop in the network to steal information. At the same time, they can alter the validating data by man-in-the-middle attack. Proper network security needs to be implemented to use PVU method.

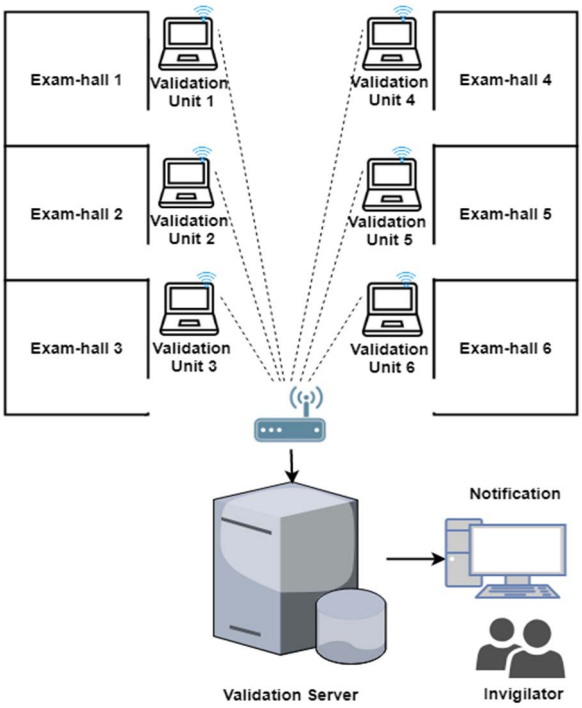


Fig. 3. PVU architecture

We prefer implementing the PVU architecture because of its performance and flexibility.

C. Validation System

The validation process starts with the user inputs. The examinee enters his username, registration number and password. The server validates the access credentials and if the credentials are valid, it serves the image capturing web application. At the same time, the server loads the image of the examinee those were taken during the registration process. In the next step, the examinee captures his image and send it to the server for validation. The server checks if recent image and previously saved images are matching using our proposed method. If the images match with acceptable rate, the candidate is declared as valid examinee. If the images do not match, then the server marks the candidate and generates an alarm at the invigilators' office. Concurrently, the candidate receives a message containing the instruction to visit the invigilators' office. The invigilators will further investigate the issue and will decide if the candidate is a valid candidate or proxy examinee.

V. Performance Analysis

We used Olivetti Research Laboratory (ORL) dataset of faces for performance analysis. Each of the images are in 8-bit grayscale. The images are in PGM format. This dataset contains 400 images of 40 person. There are 40 directories representing 40 individuals and each of those directories has 10 images of same person labeled from 1 to 10. To conduct the experiment, we used Windows 10, 64 bit operating system, with 8192 MB primary memory, powered by Intel(R) Core(TM) i3-6100 CPU, 3.70GHz (4 CPUs). All of the algorithms have been implemented using MATLAB R2017a.

We have evaluated the performance by the dataset loading time, recognition time and error rate.

A. Loading Time

The loading time maintains linear relationship with the number of images. We computed the loading time for 5, 10, 15, 20, 25, 30, 35 and 40 persons. Normally 35-40 candidates attend exam in a single room. That is why we considered it unnecessary to calculate the loading time for more than 40 candidates. To load dataset of 5 people, which is 50 images, our system took 0.098 seconds. The time was extended to 0.876 seconds for 40 candidates, which are 400 images.

B. Recongition Time

We have experimented with all 400 images. In none of the cases the recognition time exceeded 1.98 seconds. To calculate the average recognition time, we have calculated the average for every five individuals at a time. Finally, we calculated the average of these average times. And the average recognition time is 1.37 seconds. The recognition time includes the time to generate eigenfaces of training images and testing image, mean subtraction time, indexing time and Euclidean distance calculation time. Altogether, our algorithm took 1.37 seconds to successfully validate a face. However, we did not consider the time when the

Table 1 Performance summary

Serial	Person	Loading Time (seconds)	Recognition Time (seconds)
1	5	0.098	1.12
2	10	0.211	1.36
3	15	0.309	1.6
4	20	0.429	1.98
5	25	0.487	1.18
6	30	0.547	1.22
7	35	0.656	1.21
8	40	0.876	1.31

algorithm gave wrong output. The recognition time has been calculated for correct output only. In Fig. 4 the comparison

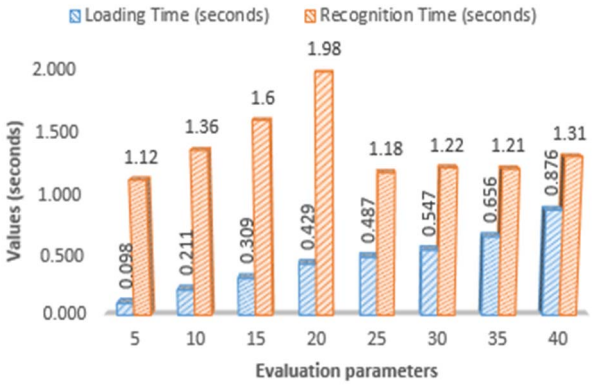


Fig. 4. Loading and recognition time.

between loading time and recognition time has been shown. In our system, the loading time is pre-validation process. It will be loaded before starting the validation process. The impact of recognition time is the major concern. Because

recognition will be conducted in real-time and it determines the speed of the validation process.

Table 2 Accuracy and Error Calculation

Trial	Correct Recognition	Incorrect Recognition	%Accuracy	%Error
1	392	8	98	2
2	388	12	97	3
3	388	12	97	3
4	396	4	99	1
5	392	8	98	2
6	392	8	98	2
7	388	12	97	3
8	396	4	99	1

C. *Accuracy and Error Calculation*

We define the error as incorrect recognition. Our system recognizes faces from information theory perspective using eigenfaces. Although the probability of having same pattern for two different eigenfaces of two different persons is close to zero, sometimes the patterns demonstrate little difference which falls below our threshold level. As a result, there is possibility of having incorrect recognition. However, it is not feasible to employ additional algorithm to detect the error rate. That is why, we have taken human inspection approach to calculate the error of our validation system.

We have 400 images of 40 person. Each of them has 10 images with different poses. In order to find out the error rate, we loaded each of the images of every individual separately and ran the program to check if it can recognize the face correctly. We repeated the process for 8 times. In the first trial, out of 400 images 392 faces were recognized correctly, whereas 8 recognitions were incorrect. In the second trial 388 faces were successfully recognized. But 12 faces gave mismatched result. The result has been enlisted in table 2.

To calculate the percentage error, we used (6), where TI means 'total image' and CR means 'correct recognition.

$$\%Error = \frac{TI-CR}{TI} \times 100 \qquad (6)$$

Equation (6) gives us the error. From the error we calculated the percentage accuracy. Fig. 5 shows the accuracy in percentage scale.

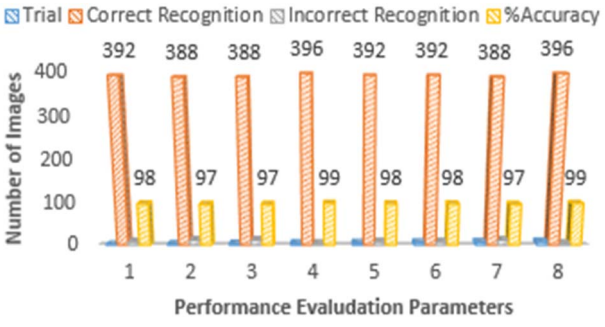


Fig. 5. Accuracy and error

The highest accuracy of our system was 99% where only 4 images were incorrectly recognized. The red bars represent the total number of correctly recognized image and the yellow bars represent the percentage accuracy.

CONCLUSION AND FUTURE WORK

An effective, efficient, simple and inexpensive method has become essential to prevent proxy examinees from taking examinations. Manual process is no longer effective because of excessive and growing number of candidates in competitive tests. The automatic examinee validation system we have proposed has all of the features required to conduct a fair environment for competition. Our proposed method is fast and can recognize faces with above 97.88% accuracy. Moreover, this system does not require any training. The registration method itself processes the images for recognition system, which validates examinees before entering the exam-center. Although it adds few more steps in registration process, considering the performance, efficiency and accuracy the system we are proposing is an excellent solution to prevent proxy examinees and ensure that only the deserving candidates will survive in competitive examinations. In our future work, we will implement the validation system physically to conduct a dummy examination. If the system performs as good as the lab experiment, we will launch the validation system as an open source examinee validation system.

REFERENCES

[1] "One-stop cheating service in exams", The Daily Star, 2018, April 8, Retrievd from https://www.thedailystar.net/backpage/one-stop-cheating-service-exams-1559536

[2] S. Lawrence, C.L. Giles, A. C. Tsoi, A.D. Back, "Face recognition: a convolutional neural-network approach", IEEE Transactions on Neural Networks, vol. 8 , pp. 89-113 , January, 1997.

[3] "38th BCS preliminary exam held", Dhaka Tribune, December 29, 2017, Retrived from https://www.dhakatribune.com/bangladesh/2017/12/29/38th-bcs-preliminary-exam

[4] W. W. Bledsoe, "The model method in facial reccognition," Panoramic Research Inc., Palo Alto, CA, Ref. PRI:15, August 1966.

[5] T. Kanade, "Picture processing system by computer complex and recognition of human faces," Department of Information Science, Kyoto University, Novermber 1973.

[6] S. Carey, R. Diamond, "From piecemeal to configurational representation of faces," Science, Vol. 195, January 21, 1977, pp.312-313.

[7] M. A. Turk, A. P. Pentland, "Face Recognition Using Eigenfaces," in Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Maui, HI, USA, USA, 3-6 June, 1991.

[8] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," J. Opt. Soc. Am. A, Vol. 4, No. 3, March 1987, pp. 519-524.

[9] P. Wagh, R. Thakare, J. Chaudhari, S. Patil, "Attendance system based on face recognition using eigen face and PCA algorithms," in International Conference on Green Computing and Internet of Things (ICGCIoT), Noida, India, 8-10 October, 2015.

[10] J. Galbally, J. Fierrez, J. Ortega-Garcia, C. McCool, S. Marcel, "Hill-climbing attack to an Eigenface-based face verification system," First IEEE International Conference on Biometrics, Identity and Security (BIdS), Tampa, FL, USA, 22-23 September, 2009.

[11] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, USA, 8-14 Dec. 2001.

[12] M. Turk, A. Pentland, "Eigenfaces for Recognition", Journal of Congnitive Neuro-science, March 1991.

[13] K. Cheng, L. Xiang, T. Hirota and K. Ushijima, "Effective Teaching for Large Classes with Rental PCs by Web System WTS," in Proc. Data Engineering Workshop 2005 (DEWS2005), 2005, 1D-d3.

[14] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," ACM Computing Surveys, 2003, vol. 35, no. 4, pp. 399-458.

[15] S. Nishiguchi, K. Higashi, Y. Kameda and M. Minoh, "A Sensor-fusion Method of Detecting A Speaking Student," IEEE International Conference on Multimedia and Expo (ICME2003), 2003, vol. 2, pp. 677-680.

[16] R.E. Burkard and E. C¸ ela, "Linear Assignment Problems and Extensions", In Handbook of Combinatorial Optimization, Du Z, Pardalos P (eds). Kluwer Academic Publishers: Dordreck, 1999, pp. 75-149.

[17] A. K. Jain, A. Ross, and S. Pankanti, "Biometrics: a tool for information security," IEEE Trans. on Information Forensics and Security, vol. 1, no. 2, pp. 125–143, 2006.

[18] P. Mohanty, S. sarkar, and R. Kasturi, "From scores to face templates: a model-based approach," Pattern Analysis and Machine Intelligence, vol. 29, pp. 2065–2078, 2007.

[19] C. Soutar, "Biometric system security. http://www.bioscrypt.com/assets/security soutar.pdf."

[20] A. Adler, "Sample images can be independently restored from face recognition templates," in Proc. CCECE, vol. 2, 2003, pp. 1163–1166.

[21] U. Uludag and A. K. Jain, "Attacks on biometric systems: a case study in fingerprints," in Proc. SPIE-IE, vol. 5306, no. 4, 2004, pp. 622–633.

[22] F. S. Samaria and A. C. Harter, "Parameterization of a stochastic model for human face identification," in Proc. 2nd IEEE Wkshp. Applicat. Comput. Vision, Sarasota, FL, 1994.

[23] K.-K. Sung and T. Poggio, "Learning human face detection in cluttered scenes," in Computer Analysis of Images and Patterns, G. Goos, J. Hartmonis, and J. van Leeuwen, Eds. New York: Springer–Verlag, 1995, pp. 432–439.

[24] K. Sutherland, D. Renshaw, and P. B. Denyer, "Automatic face recognition," in Proc. 1st Int. Conf. Intell. Syst. Eng. Piscataway, NJ: IEEE Press, 1992, pp. 29–34.

[25] D. L. Swets and J. J. Weng, "Using discriminant eigenfeatures for image retrieval," IEEE Trans. Pattern Anal. Machine Intell., to appear, 1996.

[26] M. Turk and A. Pentland, "Eigenfaces for recognition," J. Cognitive Neurosci., vol. 3, pp. 71–86, 1991.

[27] J. Weng, N. Ahuja, and T. S. Huang, "Learning recognition and segmentation of 3-D objects from 2-D images," in Proc. Int. Conf. Comput. Vision, ICCV 93, 1993, pp. 121–128.

[28] L. Wiskott, J.-M. Fellous, N. Kr¨uger, and C. von der Malsburg, "Face recognition and gender determination," in Proc. Int. Wkshp. Automat. Face Gesture Recognition, Z¨urich, Switzerland, 1995.

Prediction of Potato Disease from Leaves using Deep Convolution Neural Network towards a Digital Agricultural System

Md. Al-Amin
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
alamin15-376@diu.edu.bd

Tasfia Anika Bushra
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
anika.cse@diu.edu.bd

Md Nazmul Hoq
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
salim.cse@diu.edu.bd

***Abstract*—Potato is one of the most used crops in the world and 2nd most important crop in Bangladesh. Our economy is largely affected by the production of potato. But its production is hampered due to different diseases of potato leaves. These diseases decrease production and increase the price of potatoes. Our objective is to develop an automated system which will predict the potato disease and helps farmers to take necessary steps. In this paper, we have implemented a deep Convolutional Neural Network (CNN) which provides 98.33% accurate result in predicting different diseases of potatoes. To best of our knowledge, this is the highest accuracy gained for only potato disease prediction. The system is cost effective, less time consuming and provides an efficient way of predicting potato diseases from leaves. This will help the farmers and lead our country towards a digital agricultural system.**

Keywords— Potato Disease Prediction, CNN, Deep Learning.

I. INTRODUCTION

The potato is grown as a major crop in different countries (e.g. China, India, Russian Federation) of the world [1]. Potato (*Solanum tuberosum*) is the 2nd most important crop in Bangladesh, after rice. The total world potato production is estimated at 388,191,000 tonnes in 2017 [2]. Currently, the potato crop is grown on a significant scale in about 130 countries. Potato production is threatened by several diseases resulting in considerable production losses, and deteriorates the quality and increases the price of potatoes. The two main diseases of potato plants are *Phytophthora infestans* (Late Blight) and *Alternaria solani* (Early Blight). The farmers are facing many problems for these diseases which can be solved by early prediction of the diseases. Manual prediction of potato disease is difficult, costly, and time-consuming, while the computerized system is more efficient and cost-effective. Recently, Machine Learning (ML) system are extensively used to automate the processes. Machine learning can be an effective method for continuous monitoring of plant health status and early detection of plant diseases. Every disease has some visible symptoms on the plants, particularly on leaves. These symptoms, for example, a visible pattern of the affected leaves help to predict the disease. In this way, machine learning offers a solution to the issue of agricultural productivity and ensures food security.

Convolutional neural network (CNN) is one of the most widely used techniques for image recognition, image classification, medical imaging etc. CNN takes an input image, processes it and classify it to a certain category. CNN takes an input image as an array of pixels. Each input image will pass through a series of convolution layers with filters like pooling, fully connected layers(FC), etc. Different activation functions are applied to classify an image with probabilistic values between 0 and 1.

The main goal of this work is to develop machine learning based effective and error-free disease prediction system for the potato plants. We have used Plant Village dataset for developing our deep convolutional neural network model.

This paper is organized as follows: The related works are presented in Section II. Section III explains the dataset of the proposed work. Section IV shows the methodology and the architecture of the deep convolutional neural network. Section V represents our result and discussion. Finally, conclusion and future work are presented in Section VI.

II. RELATED WORK

Researchers explore plant disease detection using deep learning, SVM, RGB imaging, X-ray, ultrasound, and multispectral and hyperspectral technologies. In paper [3], a multiclass support vector machine model was developed for classifying potato disease detection. The model was applied to 300 different images. Different performance parameters such as accuracy, sensitivity, recall, and F1-score were calculated for measuring the performance of the model. Another work has been done for apple disease identification using Multi-class SVM Classifier which was used to train and test images [4].

In paper [5], a deep CNN model was developed for classifying 13 leaves' disease for Apple and Tomato plants. The model was able to detect leaf presence and distinguish between healthy leaves and 13 different diseases, which can also be visually diagnosed. The complete procedure was described, respectively, from collecting the images used for training and validation to image preprocessing, image augmentation and finally training the deep CNN and fine-tuning of the model. The experimental results on the developed model achieved precision between 91% and 98%, for separate class tests, and provide an average precision of 96.3%. Ferentinos and Konstantinos P. [6] used AlexNetOWTBn and VGG deep learning architecture to classify 25 different diseases. Faster R-CNN, R-FCN, and SDD has been used for classifying real-time tomato plant diseases and pest recognition.

A few works [7-11] has been done for several diseases classification problem using deep learning. In paper [12] authors used K-Means clustering. They converted an image from RGB color space to L*a*b* color space and then used color coherence vector (CCV), global color histogram (GCH), complete local binary pattern (CLBP), Gabor features, local binary pattern (LBP) and local ternary pattern (LTP). Then authors performed feature level fusion by fusing for more than two texture and color features and also applied random forest classifier on segmented image. Fruit disease was shown after classification by tagging every pixel using this method.

III. DATASET DESCRIPTION

Plant Village (www.plantvillage.psu.edu), a publicly available image database, contains 54,306 images of the different plant leaves. We have analyzed 2250 images of potato leaves. Table I shows the detailed statistics of the image dataset.

TABLE I. DETAILS STATISTICS OF IMAGE DATASET

Disease	Number of images
Early Blight	1000
Late Blight	1000
Healthy	250

IV. METHODOLOGY

We have used CNN for predicting potato disease using potato leaves. Fig. 1 shows the block diagram representing the methodology of predicting potato diseases. Our methodology is divided into two major steps. The steps are:

- Preprocessing of images.
- Deep convolutional neural network.

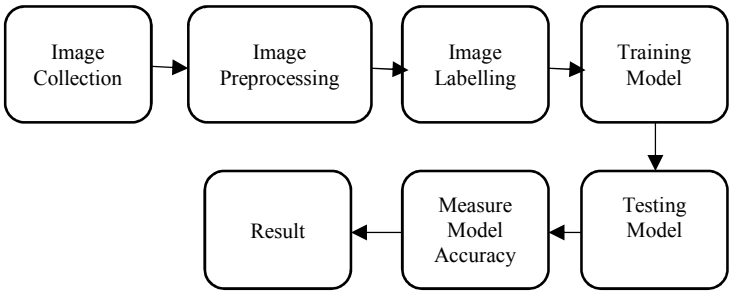


Fig. 1. The methodology of predicting potato diseases

The details of the two major steps are described below:

A. Preprocessing of images

Preprocessing refers to the transformations on the raw data before it is fed to the machine learning or deep learning algorithm. For instance, training a convolutional neural network on raw images will probably lead to bad classification performances[10], whereas preprocessing increase efficiency. The preprocessing is also important to speed up training, for instance, centering and scaling techniques.

- Gray Scaling:

The original size of Plant Village images is 256×256 pixels. We have converted all the RGB images to grayscale images. The color channel is converted to 1 channel from 3 channel. Fig 2. shows the input image and the result of applied grayscale techniques.



(a) (b)

Fig. 2. (a) Input image (b) Grayscale image

- *Removing Blur from Images*:

We have used Gaussian blur technique for removing blur from the image. First, we have added blur and then subtracted the blurred image from the original image. Then we have added a weighted portion of the mask to get the de-blurred image.

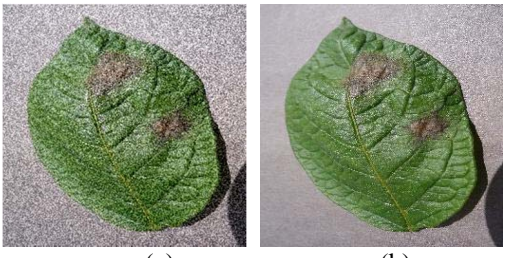
$$g_{mask}(x,y) = f(x,y - f'(x,y)) \quad \ldots\ldots\ldots(1)$$

$$g(x,y) = f(x,y + k * g_{mask}(x,y)) \quad \ldots\ldots(2)$$

Here $f'(x,y)$ is the blurred image and k is a weight for generality

- Removing Noise from Images:

We have removed salt and pepper noise from the images. We have used the median filter to remove noise. After removing the noise, the images become more clear, sharp and noiseless. After removing noise, we have converted the image to an array of pixels for feeding into our model. Fig 3. shows the input image and noise reduced image.



(a) (b)

Fig. 3. (a) Input image (b) Noise-free image

- Labeling images:

After preprocessing all the images we have used label binarizer to label all images into corresponding class. Then we have split our dataset into train, test and validation set. Table II shows the detailed

statistics of the train, test, and validation set with the split ratio.

TABLE II. DETAILS STATISTICS OF TRAIN, TEST AND VALIDATION SET

Class	Train	Test	Validation
Late Blight	800	100	100
Early blight	800	100	100
Healthy	200	25	25
Total	1800	225	225

B. Architecture

Our proposed architecture has 5 convolution layers and 2 fully connected dense layer. The first convolution layer has 32 filters where the filters are of size 3×3, the second and third layer has 64 filters with filter size 3×3 and last two layers have 128 filters and the filter size is 3×3. Rectified Linear unit activation function (*ReLu*) is used for all layers. Total three MaxPooling layers are used and six Batch normalization are used in our model. The pool size of MaxPooling layer is 2×2. Batch normalization is used to speed up the learning. Dropout 25% was added to the first, third and fifth layer to reduce overfitting of the data. Two fully connected dense layers have been used. The first one has 1024 filters and the last one has only 3 filters for three output. We used Adam optimizer to update the weights. Table III. shows the model of our proposed deep Convolutional Neural Network.

TABLE III. SUMMARY OF THE PROPOSED DCNN MODEL

Layer	Output Shape
Conv2D_1	(None,256,256,32)
Activation_1	(None,256,256,32)
BatchNormalization_1	(None,256,256,32)
MaxPooling2D_1	(None,85,85,32)
Dropout_1	(None,85,85,32)
Conv2D_2	(None,85,85,64)
Activation_2	(None,85,85,64)
Batch Normalization_2	(None,85,85,64)
Conv2D_3	(None,85,85,64)
Activation_3	(None,85,85,64)
Batch Normalization_3	(None,85,85,64)
MaxPooling2D_2	(None,42,42,64)
Dropout_2	(None,42,42,64)
Conv2D_4	(None,42,42,128)
Activation_4	(None,42,42,128)
Batch Normalization_4	(None,42,42,128)
Conv2D_5	(None,42,42,128)
Activation_5	(None,42,42,128)
Batch Normalization_5	(None,42,42,128)
MaxPooling2D_3	(None,21,21,128)
Dropout_3	(None,21,21,128)
Flatten_1	(None,56448)
Dense_1	(None,1024)
Activation_6	(None,1024)
Batch Normalization_6	(None,1024)
Dropout_4	(None,1024)
Dense_2	(None,3)
Activation_7	(None,3)

C. Parameters of Training model

The training model of CNN architecture is dependent on some parameters(e.g. learning rate, batch size, epoch, etc.). Learning rate is a parameter that controls how often we need to adjust the weights of the CNN model with respect to the gradient loss. Batch size defines the total number of training examples present in a single batch. We used batch size 32 so each time 32 sample images are present in a single batch. Epoch means the number of time a full dataset is passed forward and backward through the CNN. And shuffle parameter is used for shuffling the entire dataset after completing one epoch. Used parameters and the values of our training model are given in Table IV

TABLE IV. PARAMETERS OF TRAINING MODEL

Parameter Name	Value
Learning Rate	10^{-3}
Batch Size	32
Epoch	150
Shuffle	True

D. Analysis and evaluation tools

All the preprocessing and classification algorithm implementations, as well as evaluation of the classifiers, were performed using python version 3.X.

V. RESULT AND DISCUSSION

We have split our total data into train, test and validation dataset. We have applied our model to the training dataset. We have achieved an overall testing accuracy of 98.33%. We have also observed precision, recall and F1 score for better understanding and analyzing the result. Precision is the ratio of correctly predicted true observations to the total predicted true observations. Recall is the ratio of correctly predicted true observations to all observations of the actual class. F1 Score is the weighted average of Precision and Recall. Table V shows precision, recall and F1 score of the testing dataset.

TABLE V. PRECISION, RECALL AND F1 SCORE

Precision	Recall	F1 Score
0.9851	0.9809	0.9826

We have used confusion matrix to represent the visualization of the performance of our algorithm on the testing dataset. From confusion matrix, one can easily find out which class of data provides good accuracy and which provide the worst accuracy. Fig. 4 shows the graphical representation of the confusion matrix containing test data for healthy(class 0), early blight (class 1) and late blight (class 2). From the figure, we can observe that 2 data items are wrongly predicted for late blight (class 2) data. For the rest of the classes i.e. healthy (class 0) and early blight (class 1), we have successfully predicted all the diseases.

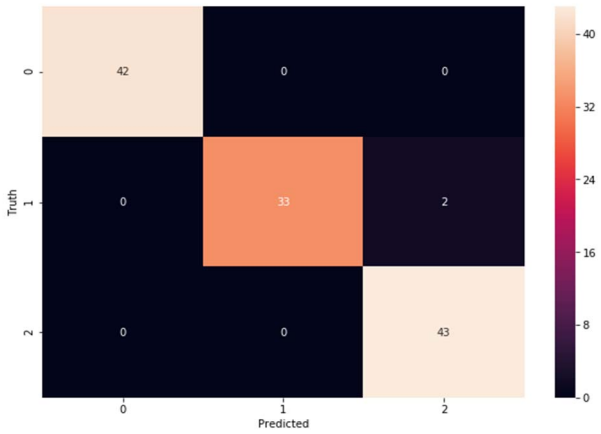


Fig. 4. Confusion Matrix represnting the visual of the performance

Finally, we have tested our trained model with test data. We took individual image of potato leaf as input to predict the disease. Fig. 5 shows the predicted result when we applied our method to some affected images. The value on the top left corner represents the accuray of the predicted result. We can see that for a sample healthy plant we got 97% accuracy, for a sample early blight plant we got 96% accuracy and for a sample late blight plant we got 99% accurate result.

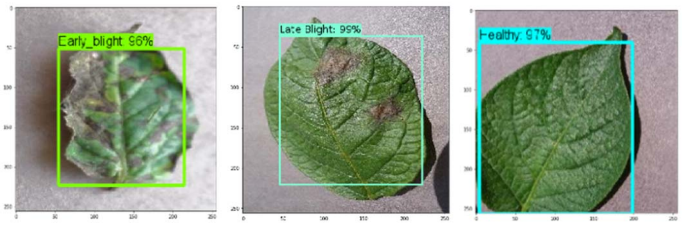


Fig. 5. Result of sample predicted disease images (a) early blight (class 1) (b) late blight (class 2) (c) healthy (class 0)

We have found two recent works on potato disease prediction. Islam et al. [3] worked with the same dataset we have used and got 95% accurate result. They used RGB imaging technique. Hu et al [13] used hyperspectral imaging technique and got 95% accuracy. Few other works have also been found regarding disease prediction of different other plants. For example Rice and Tomato. There are also several works on mixed data of different plants. But for simplicity, we have shown a comparison of only single plant disease prediction performance. Comparing with other works, we achieved the highest accuracy of 98.33%. Table VI. shows the comparison between various detection algorithm for leaves diseases prediction.

TABLE VI. COMPARISON OF VARIOUS DETECTION TECHNIQUES/ALGORITHM OF LEAVES DISEASE PREDICTION

.

Reference	Techniques	Plant	Accuracy (%)
This Work	Deep CNN	Potato	98.33
Islam et al. *2017* [3]	RGB Imaging	Potato	95
Hu et al. 2016 [13]	Hyperspectral Imaging	Potato	95
Asfarian et al. 2013 [14]	Texture Analysis and PNN	Rice	83
Xu et al. 2011 [15]	K-Nearest Neighbor	Tomato	82

VI. CONCLUSION AND FUTURE WORK

We presented a deep CNN based approach to predict the two most common diseases named late blight, early blight along with the healthy potato. We have achieved 98.33% accuracy which is the best result comparing others work on the same dataset according to our knowledge. This work can be very effective to increase the production of potato as well as decrease the price of the potatoes. Our approach will offer farmers a feasible, efficient and timesaving way of potato disease identification. This will lead to a digital agricultural system. In the future, we want to increase the accuracy by augmenting the images of the dataset. This work can also be extended for other plants to predict the diseases. As a future work, a smartphone-assisted system might be developed to automate the detection process.

REFERENCES

[1] B. Leff, N. Ramankutty, and J. A. Foley, "Geographic distribution of major crops across the world, ' *Earth's Future'*, 2004. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2003GB002108. [Accessed: 19-Feb-2019].

[2] *www.potatopro.com, ' The Potato Sector'*, 2019. [Online]. Available: http://www.potatopro.com/world/potato-statistics. [Accessed: 19-Feb-2019].

[3] Islam, Monzurul, Anh Dinh, Khan Wahid, and Pankaj Bhowmik. "Detection of potato diseases using image segmentation and multiclass support vector machine." In *2017 IEEE 30th Canadian conference on electrical and computer engineering (CCECE)*, pp. 1-4. IEEE, 2017.

[4] Dubey, Shiv Ram, and Anand Singh Jalal. "Detection and classification of apple fruit diseases using complete local binary patterns." In *2012*

Third International Conference on Computer and Communication Technology, pp. 346-351. IEEE, 2012.

[5] Sladojevic, Srdjan, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. "Deep neural networks based recognition of plant diseases by leaf image classification." *Computational intelligence and neuroscience*2016 (2016).

[6] Ferentinos, Konstantinos P. "Deep learning models for plant disease detection and diagnosis." *Computers and Electronics in Agriculture* 145 (2018): 311-318.

[7] Fuentes, Alvaro, Sook Yoon, Sang Kim, and Dong Park. "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition." *Sensors* 17, no. 9 (2017): 2022.

[8] Oppenheim, Dor, Guy Shani, Orly Erlich, and Leah Tsror. "Using Deep Learning for Image-Based Potato Tuber Disease Detection." *Phytopathology* ja (2018).

[9] Hanson, A. M. G. J., M. G. Joel, A. Joy, and J. Francis. "Plant leaf disease detection using deep learning and convolutional neural network." *International Journal of Engineering Science*5324 (2017).

[10] Barbedo, Jayme GA. "Factors influencing the use of deep learning for plant disease recognition." *Biosystems engineering* 172 (2018): 84-91.

[11] Pal, Kuntal Kumar, and K. S. Sudeep. "Preprocessing for image classification by convolutional neural networks." In *2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, pp. 1778-1781. IEEE, 2016.

[12] Leemans, Vincent, Hugo Magein, and M-F. Destain. "Defects segmentation on 'Golden Delicious' apples by using colour machine vision." *Computers and Electronics in Agriculture* 20, no. 2 (1998): 117-130.

[13] Hu, Y. H., X. W. Ping, M. Z. Xu, W. X. Shan, and Y. He. "Detection of Late Blight Disease on Potato Leaves Using Hyperspectral Imaging Technique." *Guang pu xue yu guang pu fen xi= Guang pu* 36, no. 2 (2016): 515-519.

[14] Asfarian, Auzi, Yeni Herdiyeni, Aunu Rauf, and Kikin Hamzah Mutaqin. "Paddy diseases identification with texture analysis using fractal descriptors based on Fourier spectrum." In *2013 International Conference on Computer, Control, Informatics and Its Applications (IC3INA)*, pp. 77-81. IEEE, 2013.

[15] Xu, Guili, Fengling Zhang, Syed Ghafoor Shah, Yongqiang Ye, and Hanping Mao. "Use of leaf color images to identify nitrogen and potassium deficient tomatoes." *Pattern Recognition Letters* 32, no. 11 (2011): 1584-1590.

Human Identification Based on Color Stimuli

Md. Mahmudul Hasan[1], Nafiul Hasan[2], Md. Mohaiminul Islam[3], Mahozabin Suhi[4]
[1,2,3,4]*Department of Electrical and Electronic Engineering*
[1,2,3,4]*Bangladesh Army University of Engineering & Technology,*
Qadirabad, Bangladesh
[1]mubin.eee.kuet@gmail.com,[2]nafiulhasan71@gmail.com,[3]mohaiminulbauet@gmail.com, [4]mahojabinshuhi@gmail.com

***Abstract*—Human identification by using Electroencephalogram is becoming promising field and reliable to improve security systems. It is difficult to acquire EEG at a certain mental condition always such as concentration or relaxation. This paper represents a simple model to identify individuals and finding most effective primary color by using features of EEG by means of color stimuli. A comparison between primary and secondary colors for identification has also been made. Standard additive primary colors blue, green, red and one secondary color yellow were selected for experiment. Four neural networks were built by extracting various features of EEG in the domain of time and frequency. All artificial neural networks showed satisfactory performance with minimum mean square error for identification. Among the four selected colors blue color based ANN showed minimum mean square error of 6.238×10^{-08}.**

Keywords—Electroencephalogram (EEG), Human Identification, Color Stimuli, Artificial Neural Network (ANN), Biometrics.

I. INTRODUCTION

Human identification system uses identifier's features which makes it unique. Existing identification system are mostly either knowledge or possession based [1]. Password, PIN, fingerprint, voice etc. alike based database are very vulnerable in today. Facial recognition system is better but it also has been failed to complete reliability such as in the incident of Boston marathon bombing where the system fails to identify. Fingerprint can be stolen and all the data sets can be easily hacked. So human identification by using features of individual's brain wave becomes first choice for reliability [2-4].Biometrics uses special physiological or behavioral feature of individuals. As EEG consist of many features of brain, it is becoming a prominent research field [5].

Features of EEG due to different color stimuli is studied. Individual color stimuli has an effect on human emotion [6]. As EEG isn't constant and to reproduce subject need a specific environment such as concentration or relaxation, to make this system independent of hard experimental condition; the alternative method is to use color stimuli. Previous works shows that statistical and dynamic features gives best performance for human identification while using features of EEG [7]. So both domain's features is used in this proposed model.

The primary color red, green and blue are known as RGB colors. Primary colors are mixed together to generate secondary color. Researches shows significant effect of RGB on visual impact. A color theory namely opponent process describes the interpretation of color information with human visual system[8].

Practically most of the time the dataset have nonlinear relation. The dataset can be generalized by artificial neural network easily. Moreover ANN don't have input pre requisite. Besides theoretically ANN belongs to deep learning. To deal with the data complexity, volatility artificial neural network is chosen in this proposed work.

This paper presents a prominent methodology for human identification by using color stimuli. BIOPAC MP36, Acqknowledge 4.1 was used for acquiring EEG [9] and feature detection, MATLAB 2014a used for artificial neural network design.

The following sections of this article is described accordingly: section II represents the chosen methods, section III describes EEG processing and analyzing, section IV shows selection of features and comparison of colors. Section V and VI briefly discuss classification, experiment result and conclusion with future work respectively.

II. PROPOSED MODEL

A. Projected Model

The proposed model for identification of individual using color stimuli is visualized in Fig. 1 below. Having acquired EEG signal with BIOPAC with the assistance of MP36, features were taken out by Acqknowledge 4.1 and MATLAB 2014a which used to design artificial neural network. Deep learning with MATLAB leads to classification approach. By analyzing mean square error, regression, gradient, accuracy a decision is taken about whether the proposed model will work satisfactorily or not, which primary color is best and what type of color is better between primary and secondary.

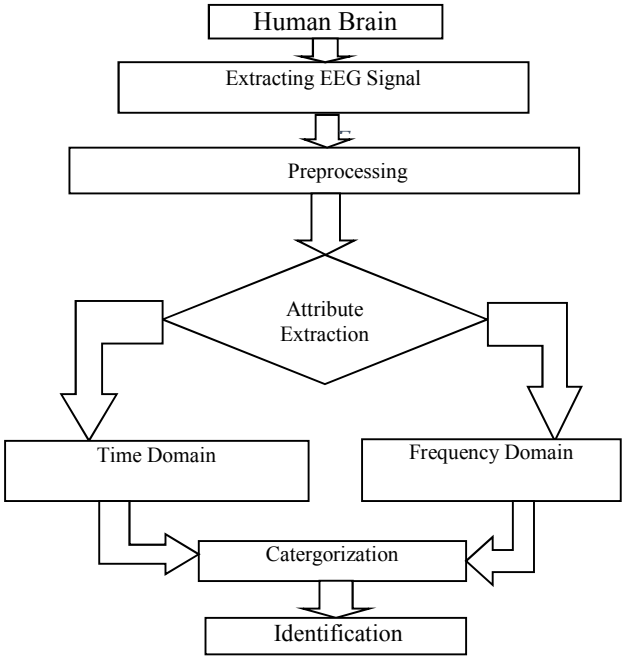


Fig 1. Flow chart of the planned approach for Biometrics using EEG

B. *Electroencephalography Acquiring and Exploratory Condition*

The test was conducted in Biomedical Engineering Lab of KUET (former BIT, Khulna), Bangladesh. Subjects were three male persons having no psychological and major visionary problem such as color blindness. Subjects were instructed to concentrate on colors. Total twenty individual trials were taken for each subject and color. Used two trials for testing purposes. Colors were shown in computer desktop screen.

C. *Experimental Setup*

1) Hardware description: BIOPAC data acquisition unit MP36 was used. Fig. 2 shows EEG signal acquisition system during the experiment.

2)Required software: Different softwares namely Acqknowledge 4.1 and EEGlab were used for signal processing approach. Relative waveforms is given by BIOPAC student Lab software after it's built in signal acquisition and processing system. Different features of signal were measured and analyzed in this platform.



Fig. 2. Pictorial view during experiment at BME, KUET.

III. Signal Processing And Analysis

A. *Filtering and Artifact Removing*

After acquiring the raw EEG signal it showed artifact and noise which due to EOG and EMG noise because of hand motion, eye blinking and line frequency (50 Hz). The raw EEG was filtered between .5 Hz to 44 Hz by band pass FIR.

B. *Smoothing*

Smoothing makes average and replace it with the real value. In words smoothing is averaging. Moderate smoothing (factor 5) was done. In (1) n, m are no. of points and samples respectively.

$$x_{output} = \sum_{r=m-\frac{n}{2}}^{r=m+\frac{n-1}{2}} \frac{x_{input}(r)}{n} \text{.....(1)}$$

C. *Domain analysis*

Standard deviation (STDDEV), Maximum value (MAX), kurtosis, skewness these statistical methods were measured in the study as time domain features. For measuring the overall power of EEG, Power Spectral Density (PSD) measures were used in frequency spectra and as a part of DFT, Fast Fourier Transform (FFT) was used as a parameter. Hamming window was used during FFT.

IV. Extracted Data

A. *Contrast of the EEG attributes of the individuals*

In time and frequency domain we used maximum, SD, skewness, kurtosis, PSD mean, PSD max, FFT mean, FFT max where the first four attributes are time domain and remaining fours are frequency domain features. The features of subject 1 in time domain while looking at blue color are illustrated below in Table I.

TABLE I. Response for blue color of subject 1

Trial	Maxm. (Micro-V)	Std. Dev. (Micro-V)	Skewness (Micro-V)	Kurtosis (Micro-V)
1	13.30831	5.79718	-0.0591	3.01581
2	8.776440	4.83765	-0.36379	2.19582
3	11.77697	3.59855	0.36564	3.68987
4	15.70188	5.25637	0.07720	3.26976
5	7.191940	3.88736	-0.1355	2.41993
6	9.168280	3.50900	0.10184	2.61476
7	7.122060	3.33194	0.10835	2.41100
8	10.48134	4.81770	-0.05164	2.35189
9	17.64717	4.89702	0.80992	4.74267
10	10.48335	4.36073	-0.23235	3.21103
11	11.19927	4.68437	0.46547	2.39521
12	9.036970	4.06186	-0.21058	2.64188
13	13.93632	4.75774	0.59110	3.15644
14	8.946480	4.73119	-0.41294	2.77898

15	9.362860	3.97022	-0.08332	2.75909
16	12.61301	4.51157	0.12616	3.52430
17	8.308320	3.74044	0.01996	2.66426
18	11.40869	3.42292	0.10140	3.87897
19	8.13347	3.42092	-0.06007	2.67656
20	8.43347	3.43864	-0.06569	2.85472

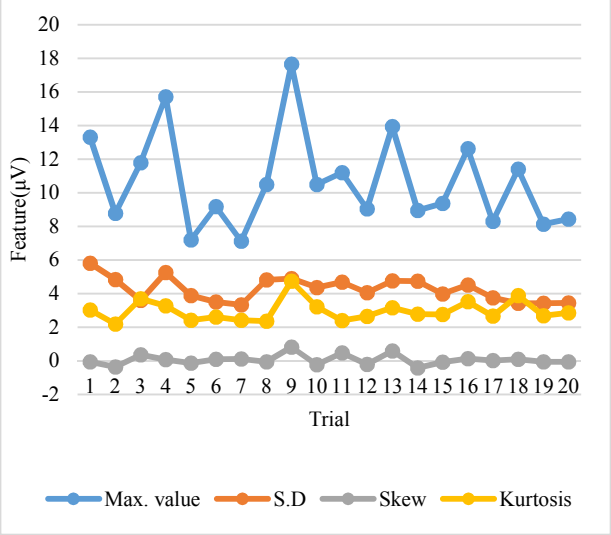


Fig. 3. Time domain features of subject 1 while looking at blue color.

Fig. 3 shows that the fluctuation of features of time series varies a little for subject 1 for blue stimuli in twenty different trials.

B. *Comparisons of features for different individuals*

After plotting definite feature of all subjects taken in twenty trials, it shows small variation of feature for same subject and comparatively more variation for different individuals. Fig. 4 shows the mean PSD for three subjects at twenty trial. Similar graphs are obtained for max PSD, mean FFT and max FFT.

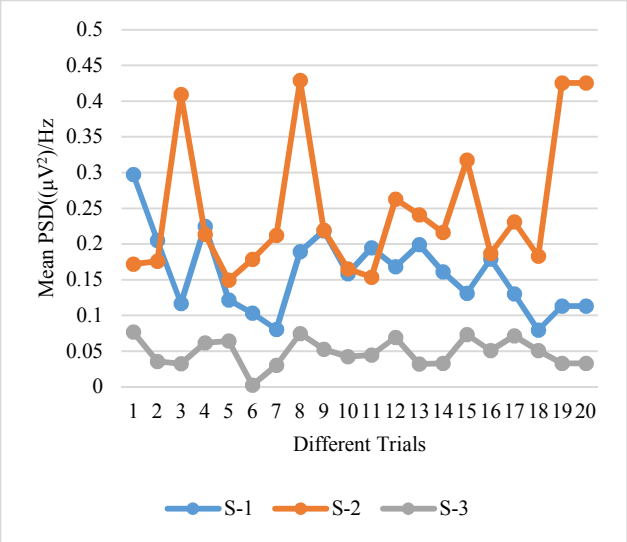


Fig. 4. Comparison of Mean PSD of different subjects.

V. Classification Results & Discussion

In this experiment we have used four color stimuli. So four neural network were built. MATLAB built in nntool GUI was used. Two layers feed forward back propagation ANN was used utilizing TANSIG transfer function and Levenberg Marquardt learning algorithm were used to design neural network.

A. *Scenario 1: Yellow color*

The ANN designed for yellow color showed mean square error of 0.22371 at epoch 1 with regression of 0.977 and successfully identified three different individuals with accuracy of 92.82%, 93.85%, 95.99% respectively. It gave the worst performance comparatively.

B. *Scenario 2: Green color*

The ANN designed for green color showed mean square of 7.7336×10^{-07} at epoch 24, overall regression 1 and successfully identified subjects with accuracy of 94.82%, 95.66%, 96% respectively. It showed good performance.

C. *Scenario 3: Red color*

The ANN designed for red color showed mean square error of 2.5798×10^{-07} at epoch 1, overall regression 1 and successfully identified all subjects with accuracy of 98.32%, 99.2%, 99.92% respectively.

D. *Scenario 4: Blue color*

The ANN designed for blue color showed mean square of 6.238×10^{-08} at epoch 5, overall regression 1 and successfully identified all subjects with accuracy of 99.98%,100%,100% respectively. Fig. 5 shows best validation performance of the neural network designed for blue color.

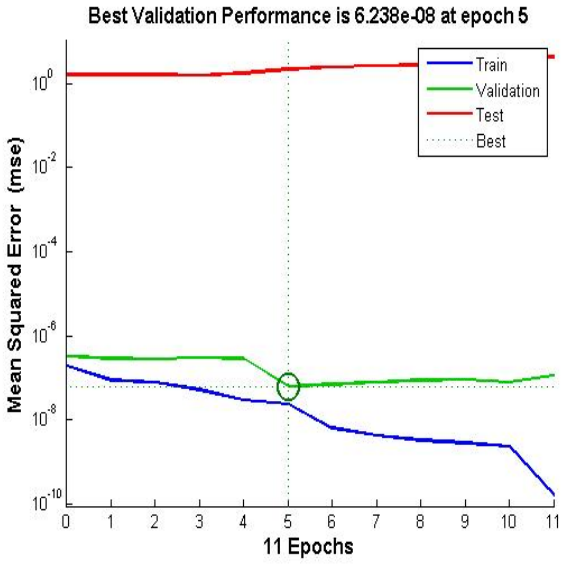


Fig. 5. Performance of neural network designed for blue color.

TABLE I I. COMPARISON DESIGNED NEURAL NETWORKS

Attributes	Yellow	Red	Green	Blue
Mean Square Error	0.223	2.579×10^{-07}	7.733×10^{-07}	6.238×10^{-08}
Overall Regression	0.977	1	1	1
Accuracy (Average)	94.22%	95.49%	99.14%	99.99%

The above table shows a comparison among ANN designed for different colors. So now the results clearly signifies that the ANN built for blue color gives best performance with 99.99% identification accuracy. Mean square error is an effective parameter for evaluating the performance of any ANN model. Lower mean square error signifies better efficiency of the designed ANN and vice versa. The value of regression varies from zero to one. Having unity or near to unity regression represents a greater efficiency and vice versa. Accuracy is measured by comparing the output and target value in terms of percentage.

VI. CONCLUSION

After the subsequent steps of EEG signal processing, it is to confirm that human identification can be done easily by means of color stimuli and EEG analysis. ANN for blue color gives lowest error of 6.238×10^{-08} whereas ANN designed for yellow color gives the highest error of 0.22371. Being primary color, blue provides best performance as it shows mean square error illustrating better performance of the network.

Development of an efficient online EEG analysis system for real application is under progress. SVM analysis of proposed model is under consideration.

REFERENCES

[1] C. Ashby, A. Bhatia, F. Tenoreb, and J. Vogelstein , "Low-Cost Electroencephalogram (EEG) based Authentication", Proceedings of the 5th International, IEEE EMBS Conference on Neural Engineering, Cancun, Mexico, April 27 - May 1, 2011.

[2] M. M. Hasan, M. H. A. Sohag, M. E. Ali, and M. Ahmad, "Estimation of the Most Effective Rhythm for Human Identification using EEG Signal", 9th International Conference on Electrical and Computer Engineering (ICECE),2016.

[3] A. Riera, A. Soria-Frisch, M. Caparrini, C. Grau, and G. Ruffini,"Unobtrusive biometric system based on electroencephalogram analysis," EURASIP J. Adv. Sig. Proc., 2008.

[4] C. R. Hema, M. P. Paulraj, and H. Kaur, "Brain signatures: A modality for biometric authentication," International Conference on Electronic Design (ICED), pp. 1–4, December 2008.

[5] Q. Gui, Z. Jin, W. Xu, "Exploring EEG-based Biometrics for User Identification and Authentication," IEEE conference on Signal Processing in Medicine and Biology Symposium (SPMB), pp. 1-6, Dec. 2014.

[6] D. Zakzewski, I. Jouny, and Y. C. Yu, "Statistical Features of EEG Responses to Color Stimuli", 40th Annual Northeast Bioengineering Conference (NEBEC), 2014.

[7] M.M. Hasan, M. M. A. Sohag, and M. Ahmad, "EEG Biometrics Based on Small Intra-individual and Large Inter-individual Difference of Extracted Features", International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), 8-10 December 2016, Rajshahi-6204, Bangladesh.

[8] D.Q. Ultama, T.L.R. Mengko, R.Mengko,M.N. Aulia, " Color blind test quantification using RGB primary color cluster", International Conference on information Technology Systems and Innovation(ICITSI),2016.

[9] Reference manuals of AcqKnowledge software guide and MP system hardware guide. (2008). BIOPAC Systems Inc., 42 Aero Camino,Goleta, CA – 0067, USA.

Design of a Convolutional Neural Network Based Smart Waste Disposal System

Md. Samiul Haque Sunny
Department of Electrical & Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
mdshs31@gmail.com

Debopriya Roy Dipta
Department of Electrical & Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
roydebopriya54@gmail.com

Shifat Hossain
Department of Electrical & Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
shifathosn@gmail.com

Hossain Mansur Resalat Faruque
Department of Electrical & Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
resalat_f@yahoo.com

Eklas Hossain
Department of Electrical Engineering and Renewable Energy
Oregon Institute of Technology
Oregon, United States
eklas.hossain@oit.edu

Abstract— **In recent times, waste management problem has become a crucial challenge for Bangladesh, which is having a detrimental impact on the environment. This paper presents the proposition of designing a smart dustbin similar to an Automated Teller Machine (ATM) along with an intelligent embedded system, which has been dubbed as Automated Teller Dustbin (ATD). An efficient convolutional neural network (CNN) based image classifier is developed, which is able to detect and recognize any object regarded as garbage by analyzing training features. Additionally, it can also count the number of labeled objects and assign a price value to each object. The waste brought by any individual to the ATD will readily be recognized by the image classifier and the recycle value, which has been assigned for that object can be withdrawn by that individual. Therefore, a direct exchange of waste and its equivalent price is possible, which will incentivize people to use our proposed smart dustbin. After the installation cost, the operation and maintenance cost can be gained by recycling the garbage in it. A pre-trained CNN-based model ALexNet has been utilized to train and test the model with a dataset of 20 images for each of the 10 categorized objects collected from different waste management shops in Dhaka, Bangladesh. The model that has been trained for object recognition has attained an accuracy of 96%, which bears testimony to the feasibility of our proposal.**

Keywords— Waste management, Convolutional Neural Network, Automated Teller Dustbin, Object detection, Object recognition

I. INTRODUCTION

Environmental pollution in Bangladesh is rising drastically due to uncontrolled discharge of wastes in our environment [1]. Moreover, the rapid growth of population has made it even harder to handle waste management, particularly in urban areas such as Dhaka, which happens to have the highest density of population in the world (44,000 per square kilometers) [2]. According to Bangladesh Waste Database 2014, the approximate amount of waste generated in its urban areas is about 23,688 tons per day, which has been estimated to increase to 47,000 tons per day by 2025 [3]. If this growth continues, Bangladesh will soon surpass the maximum limit of waste generation that has been envisaged [3]. This disastrous outcome can only be mitigated by implicating smart waste management technologies to collect wastes properly for recycling purposes. In Bangladesh, recycling process is done manually and people involved in this process are mostly women and children. Fig. 1 represents the present conditions of recycling plastic waste collection and sorting management involving women and children. Predominantly, waste is categorized into two classes- organic waste and non-organic waste. Food waste covers most of the organic waste.



(a)



(b)

Fig. 1 Present condition of recycling plastic waste in Bangladesh.

On the contrary, non-organic wastes are comprised of metal, plastic, paper or different electric and electronic wastes, which are commonly referred to as e-waste. Due to unplanned location of dustbin as well as reluctance of people to use dustbin heavily contribute to the increasing rate of waste disposal at roadside in Bangladesh. Many researches, as well as experiments, have already been carried out about the management of wastes [4-6]. However, most of them are focused on waste management after they are readily available in dustbin [7-8]. Nevertheless, in this paper, we are more concerned about the before scenario i.e. how to collect and categorize waste in the first place; how to incentivize city residents to use the bins properly instead of littering the roadsides. For coming up with a solution, this paper concentrated on designing a smart dustbin, which will act like an Automated Teller Machine (ATM). In our proposed model, in exchange for waste in the smart dustbin, anyone can withdraw the currency equivalent of the recycle value of that waste. This kind of trading will entice people to use this smart dustbin and will have a promising impact on reducing the waste disposal on roadsides. The proposed smart dustbin has an intelligent embedded system interfaced with it, which is actually a Convolutional Neural Network-based (CNN) object recognizer. For recognizing any throw-out waste, different object images categorized in 10 classes have been collected as data from multiple waste management shops located in Dhaka. Therefore, any solid non-organic recyclable waste will readily be recognized and labeled by the CNN object recognition model. According to the labeled name, the equivalent price of the recycle value will be assigned for that object, which can be withdrawn instantly from our smart dustbin. Organic waste or other solid waste, which are not recognizable will be sorted out differently through a different channel and their equivalent price will be assigned based on their weight. Hence, in our original design, a weight sensor is also suggested for handling non-recyclable or organic wastes. The proposed design has been named **Automated Teller Dustbin** (ATD). From this point onwards, this shorthand will be used for simplicity.

The rest of the paper has been organized as follows: in section II, the proposed methodology along with the algorithm and mathematical explanation has been carried out. The performance result of the model has been evaluated in section III, whereas in section IV the key features of this research have been enlisted as a conclusion.

II. Proposed Methodology

The ATD has mainly two input pathways for waste collection-one for organic waste and the other for non-organic waste. The input for organic waste is integrated with a weight sensor. Based on the weight, automatic calculation of the price of that particular waste material is performed. On the other hand, for non-organic recyclable objects, a webcam is placed in the opening, which will take live footage of the object. A pre-trained network model named AlexNet has been utilized for detection as well as for recognition of the object. As fodder for the AlexNet, 20 images for each of the 10 categorical objects have been collected as our primary dataset. When new objects are introduced, ATD can recognize that object and put a label according to which, price is determined. If multiple objects are present in one single snapshot, it locates the objects at first, then recognition for each individual object is carried out and finally, the labeling and pricing procedures are undertaken. In Fig. 2, the key concepts have been illustrated using a block diagram.

In this research our sole focus is about handling non-organic wastes, as it required much consideration due to the variation on their recycled prices. The non-organic waste management channel of our ATD comprises of a webcam and an object recognizer. For recognizing any object, the very first step is to detect the object, then separate it and finally build a model to successfully recognize that object. The flowchart that has been followed in this paper has been shown in Fig. 3.

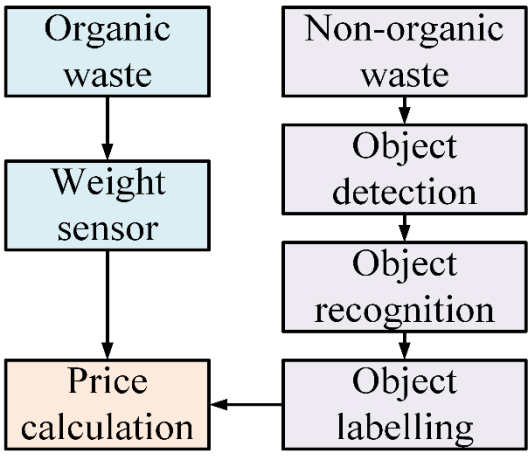


Fig. 2 Block diagram describing the work-flow of the dual input ATD.

III. System Modeling

A. R-CNN:

Region-based CNN (R-CNN) has been performed for detecting objects. This method is adopted to capture the region of an object. The detection process is basically based on hierarchical grouping algorithm. For image segmentation, the bottom-up approach has been adapted for the selective search. The total algorithm has been described in Algorithm 1 [9].

Algorithm 1: Hierarchical Grouping Algorithm
Input: I (Image)
Output: L, a set representing object location
Initial Region: K ={k_1, k_2,…, k_n}
Similarity Set: S = phi

for each neighbor region pair (k_i, k_j)
    Determine similarity s(k_i,k_j)
S=S U s(k_i,k_j)

While B ≠ phi
    Highest similarity, s(k_i,k_j)=max(S)
    Merge two regions k_t=k_i U k_j
    Cancel similarities for k_i : S=S\s(k_i, k_{new})
    Cancel similarities for k_j : S=S\s(k_{new},k_j)
    Determine similarity set between k_t and its neighbors S_t
    S= S U S_t
    K= K U K_t
Extraction of L from regions R

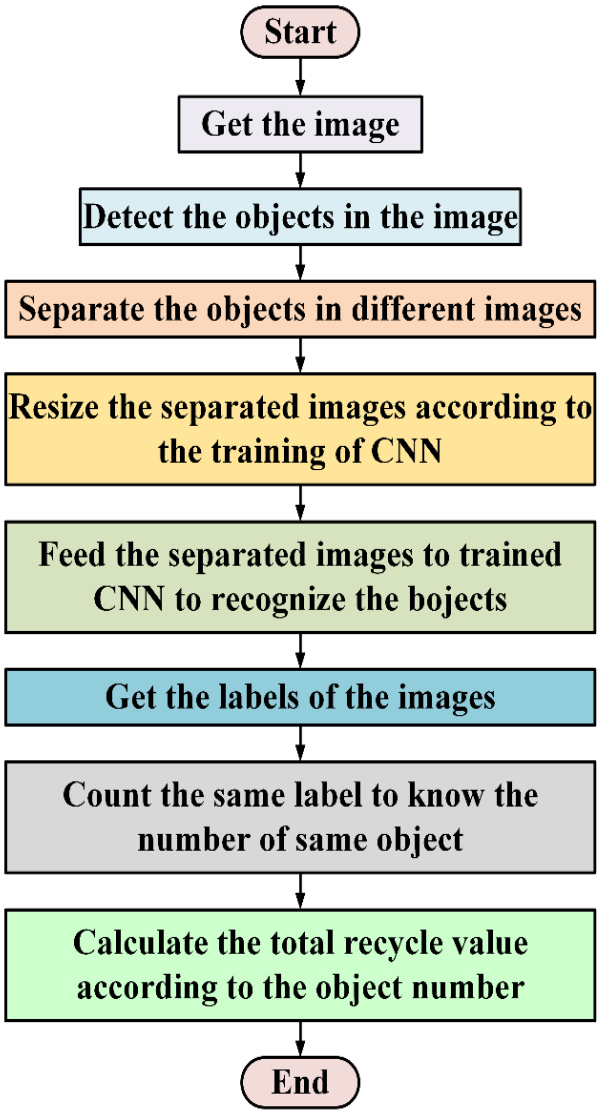


Fig. 3 Flow chart of the non-organic waste management procedure.

B. Modified AlexNet:

Object recognition can be made possible by employing machine learning methods. However, for improving the precision of the output, a large dataset is required through which, the model will be trained. AlexNet is one such model that has been trained by using over 15 million labeled images that fall roughly to 22000 categories [10]. By taking advantage of this large dataset, a well-trained AlexNet model has been developed which has 25 layers in its architecture. This pre-trained model has been modified and tuned for classifying 10 different objects. Therefore, our modified AlexNet model is basically a CNN model, which contains 5 convolutional layers and 2 fully-connected layers. The architecture of our proposed CNN model is shown in Fig. 4. The size of the input image is 224×224×3 which is fed into the first convolutional layer and is filtered with 96 kernels of size 11×11×3 with a stride of 4 pixels. The 2nd convolutional layer filters the input with 256 kernels, each of

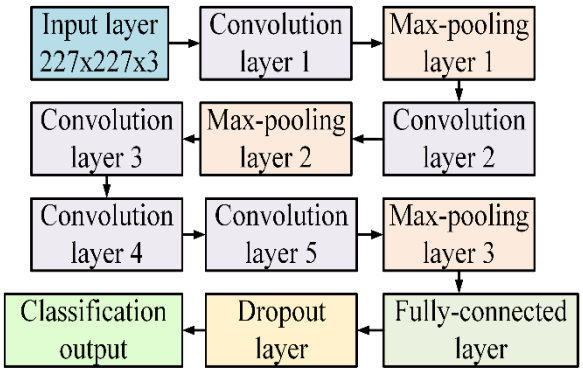


Fig. 4 Architectural block diagram of the CNN-based AlexNet model. It basically contains 5 convolutional layers and 2 fully-connected layers.

whose size is 5x5x48. After the first two convolutional layers, a max-pooling layer is entered for dimensionality reduction. The 3rd and 4th convolutional layers contain 384 and 256 kernels having size of 3x3x256 and 3x3x192, respectively. For reducing over-fitting, a drop-out layer is also added. In this layer, a neuron is made to drop out with a probability of 0.5% which obviously increases the convergence rate of the model. Nevertheless, this layer plays a vital rule to retain the model from being considerably over-fitted. For each convolutional layer, 'ReLU' activation function is associated with each neuron, as it offers much faster approach to train the whole model rather than other activation function based on saturation property e.g. 'tanh'.Only for the final fully-connected layer, 'softmax' activation function is used for determining the probabilities of each class based on which the classification is done

Let us consider the input feature map as F_i (i=1, 2,…., I), where I represents the total number of feature maps. Each of these input feature maps is connected to other feature maps in the convolutional layer based on the number of weight metrics. If the other feature map is defined by Oj (j=1, 2, …..,J), where J is the total number of other feature maps, then the total number of weight metrics will be (I×J). Let us consider the weight matrices as $w_{i,j}$=(i=1, 2, …, I; j=1, 2, …..,J).

By applying convolution operation in signal processing, the mapping can be explained. Each unit of single feature map in the convolution layer can be determined by using the following equation [11]:

$$q_{j,m} = \sigma\left(\sum_{i=1}^{I}\sum_{n=1}^{f} F_{i,n+m1} w_{i,j,n} + w_{F,j}\right) \qquad (1)$$

Where, j=1, 2,…,J. Here, $q_{j,m}$ is the mth unit of jth feature map. $w_{i,j,n}$ represents the nth element of the $w_{i,j}$ weight vector and f is the filter size.

In the maximum pooling layer, max-pooling function is employed for dimensionality reduction, which is shown in the following equation [11]:

$$p_{i,m} = \max_{n=1}^{H} q_{i,(m-1)\times x+n} \qquad (2)$$

Here, x is the shift size, H is the pooling size. Similarly, by using average function, the output can be determined as follows:

$$p_{i,m} = c \sum_{n=1}^{H} q_{i,(m-1)\times x+n} \quad (3)$$

Where, c is the scaling factor. Fully-connected layer is constructed based on the neural network algorithm. This is the most important layer of CNN. For improving performance, dropout layer has been added which helps reduce over-fitting.

IV. Result and Discussion

1) Data Collection:

In our datasets, non-organic recyclable waste has been categorized in 5 groups. These are- plastic, metal, glass, paper, and e-waste. For plastic, we have collected 20 images of each plastic bottles and boxes. In a similar fashion, for each group, we have chosen two objects and for each of them, we have collected 20 images. Hence, our total datasets have 200 images of 10 different objects along with their recycle values.

2) Object detection:

The process initiates by introducing a small region in the whole image and gradually merging them with hierarchical grouping. By analyzing the similarities between all neighboring regions, two most similar regions are selected and then grouped together. After that, the resulting group region is compared with all other neighboring regions and a new group is formed. The whole process continues until it becomes a single region. A demo of object detection and segmentation is shown in Fig. 5.



(a)



(b)

Fig. 5 (a) Detected objects in a figure (b) segmented object figures.

3) Object recognition:

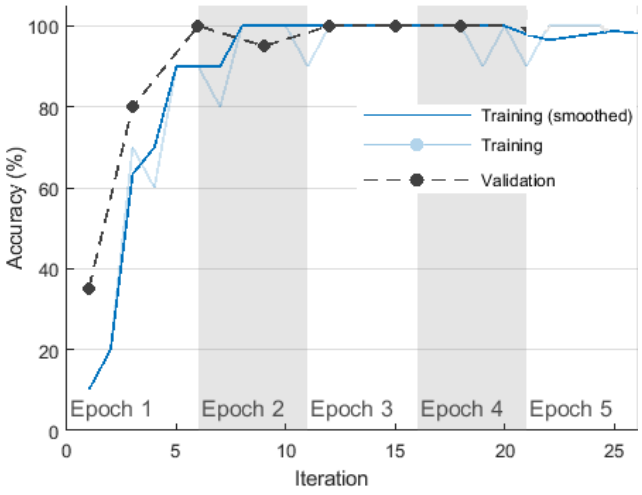
After detection and segmentation, the images of the objects are fed into the trained CNN for recognition and labeling purposes. The output of the CNN-based model along with a few labeled objects are shown in Fig. 6.



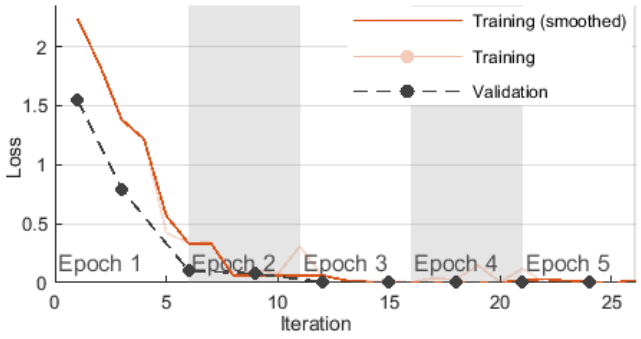
Fig. 6 Different objects that have been successfully recognized by CNN-based modified AlexNet model.

4) Performance evaluation:

In this paper, the pre-trained AlexNet model has been adapted for the purpose at hand. The training progress achieved by the modified AlexNet is shown in Fig. 7(a). From the figure, we can see that the accuracy level is progressing with each epoch and finally the validation accuracy has attained around 96 % accuracy, which is quite impressive. The loss profile of our model has also been evaluated in Fig. 7(b), which shows a gradual decrease of loss with each iteration.



(a)



(b)

Fig. 7 (a) Training progress in terms of accuracy rate. As is evident from the graph, accuracy crosses the 80% threshold after a few iterations and then saturates to 96%, (b) Loss profile with each iteration for the designed CNN model.

5) Pricing:

The price of the recycled waste according to the present market price is listed below in Table I.

TABLE I. WASTE RECYCLING PRICE LIST

Waste Category			**Price (BDT/kg)**
Non-Organic Waste	***Plastic Bottles***	Clear and Light Blue PET	10
		Colored PET	3
		HDPE natural	27
		HDPE mixed color	10
		Mixed	3
	Plastic Film	Polyethylene (PE) printed	18
		Polypropylene (PP) printed	2
		Clear natural	28
	Glass containers (jars and bottles)	Brown	1
		Clear	2
		Green	0.5
		Mixed	1
	Metal	Steel can	11
		Non-Ferrous metal	214
		Ferrous metal	10
		Aluminum can	10
	Paper	News	9
		White	25
	E-waste	Battery	50
		Display units (LCDs, CRTs)	150 per piece

V. CONCLUSION

The objective of this research was to present an idea, which can help manage roadside waste in a significant manner, especially in Bangladesh. As for the origin of road side waste, most of the time, the citizenry is responsible for not using the conventional dustbins properly. In this paper, we have presented the design of a smart dustbin with the aid of a CNN-based object recognizer. The very idea behind this design is to incentivize people to use this smart dustbin as it involves monetary compensation in exchange for a few specific waste materials. Therefore, it will not only reduce the amount of waste on the roadside but also save the long effort of waste collection. For implementing this concept, this paper has proposed the design of an ATD integrated with a CNN based object recognizer, which can readily recognize any object and then determine the recycle value of that object. With our collected datasets, the developed model has gained an accuracy of 96%, which is well-enough to move ahead for implementation in the real world.

REFERENCES

[1] Hasan S, "Problems of municipal waste management in Bangladesh: An inquiry into its nature," Habitat International, vol. 22, no. 2, pp. 191-202, 1998.

[2] I. Enayetullah, A. H. Md. Maqsood, I. Lehtenon, " Bangladesh Waste Database 2014", Waste Concern, Dhaka, Bangladesh.[Online]. Available:http://wasteconcern.org/wpcontent/uploads/2016/05/Waste-Data-Base_2014_Draft-Final.pdf. [Accessed Feb. 10, 2019].

[3] M. Shamsuzzoha (2017) Newage article page. [Online]. Available: http://www.newagebd.net/article/17041/waste-management-and-green-dhaka. [Accessed Feb. 12, 2019].

[4] A. S. Wijaya, Z. Zainuddin, M. Niswar, "Design a Smart Waste Bin for Smart Waste Management", *2017 5th International Conference on Instrumentation Control and Automation(ICA) Yogyakarta*, August 9–11, 2017.

[5] B Keerthana, Sonali M Raghavendran, S Kalyani, P Suja, V.K.G. Kalaiselvi, "Internet of Bins-Trash Management in India", *Second International Conference on Computing and Communications Technologies(ICCCT '17)*, 2017.

[6] D. D. Vu, G. Kaddoum, "A waste city management system for smart cities applications", *2017 Advances in Wireless and Optical Communications (RTUWO)*, pp. 225-229, 2017.

[7] F. Folianto, Y. S. Low, W. L. Yeow, "Smartbin: Smart waste management system", 2015 IEEE Tenth InternationalConference on Intelligent Sensors Sensor Networks and Information Processing (ISSNIP), pp. 1-2, April 2015.

[8] S. Mirchandani, S. Wadhwa, P. Wadhwa, R. Joseph, "IoT enabled dustbins", *Big Data IoT and Data Science 2017 International Conference on*, pp. 73-76, 2017.

[9] [9] J.R.R. Uijlings, K.E.A. van de Sande, T. Gevers, A. W. M. Smeulders, "Selective Search for Object Recognition", *International Journal of Computer Vision,* vol. 104, no. 2, pp. 154-171, Sep. 2013.

[10] A. Krizhevsky , I. Sutskever , G. E. Hinton, "ImageNet classification with deep convolutional neural networks", in *Proc. International Conference on Neural Information Processing Systems*, Dec. 2012, pp. 1097-1105.

[11] O. Abdel-Hamid, A.-R. Mohamed, H. Jiang et al., "Convolutional neural networks for speech recognition", in *Proc. IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 22, no. 10, pp. 1533-1545, 2014.

Brain Tumor Classification Using Convolutional Neural Network

Sunanda Das, O.F.M. Riaz Rahman Aranya, Nishat Nayla Labiba
Department of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
sunanda@cse.kuet.ac.bd, aranya.riaz@gmail.com, nishatnaylawork@gmail.com

***Abstract*— Medical image classification has gained tremendous attention in recent years, and Convolutional Neural Network (CNN) is the most widespread neural network model for image classification problem. CNN is designed to determine features adaptively through backpropagation by applying numerous building blocks, such as convolution layers, pooling layers, and fully connected layers. In this paper, we mainly focused on developing a CNN model for classifying brain tumors in T1-weighted contrast-enhanced MRI images. The proposed system consists of two significant steps. First, preprocess the images using different image processing techniques and then classify the preprocessed image using CNN. The experiment is conducted on a dataset of 3064 images which contain three types of brain tumor (glioma, meningioma, pituitary). We achieved a high testing accuracy of 94.39%, average precision of 93.33% and an average recall of 93% using our CNN model. The proposed system exhibited satisfying accuracy on the dataset and outperformed many of the prominent existing methods.**

Keywords*— Brain Tumor; Convolutional Neural Network; Kernel; Histogram Equalization; Feature Maps; Adam Optimization;

I. INTRODUCTION

The brain tumor is one of the most feared diseases in medical science. According to the American Cancer Society's [1] estimation of the brain tumor that will be diagnosed in the United States in 2019 includes 23,820 malignant tumors of the brain or spinal cord (13,410 in males and 10,410 in females). These estimation does not include benign (non-cancer) tumors. According to their evaluation, there is a possibility that 17,760 people (9,910 males and 7,850 females) could die from brain and spinal cord tumors in 2019. Survival rates of a brain tumor vary according to the type of the tumor and age of the patient. The Central Brain Tumor Registry of the United States (CBTRUS) [2] shows that the survival rate can extensively vary according to age. These data were collected on the patients who were diagnosed for the brain tumor during the year 2000-2013. The report shows that the 5-year relative survival rates for Meningioma tumor are 87.0%, 77.4%, 70.5%, 53.7%, 47.3% for age 20-44, 45-54, 55-64, 65-74, 75+ respectively.

A brain tumor is the abnormal growth of cells in the brain. A brain tumor can be primary or secondary. A primary brain tumor originates in the brain itself or tissues adjacent to it, i.e. the brain-covering membranes (meninges), cranial nerves, pituitary gland or pineal gland, while a secondary brain tumor occurs when cancer cells from other organs like lung, kidney, breast, etc. spread to the brain [3]. Primary brain tumor initially arises because of the mutations of their DNA. These mutations allow the abnormal cell to grow and the normal cell dies. It can cause brain damage and sometimes it can be life-threatening.

In this work, we proposed a CNN model that is able to accurately classify the brain tumor. Hence, the treatment for the tumor can be started at an early stage.

The remaining part of the paper is organized as follows: Section II demonstrated the related works that have been done in this field. The proposed methodology with some theoretical explanations is investigated in Section III. The result is illustrated in details with some performance measures in Section IV. The conclusion of the paper is drawn in Section V.

II. RELATED WORK

Numerous works have been done on brain tumor classification in recent years. These works are briefly outlined as follows:

Yang et al. [4] proposed the CBIR (Content-Based Image Retrieval) method to evaluate a large image dataset. This method considers the tumor region as a query and tries to find a tumor that has the same pathological properties. The system used MID (Margin Information Descriptor) as the feature. It illustrates the image content by using the surrounding tissue of a tumor. Their proposed system was able to achieve a precision of 89.3%. But the distance metric determined by their methods is globally linear, and it disappoints to produce many local projections for separate regions. Furthermore, their system is developed using manual segmentation of the brain tumor.

Huang et al. [5] suggested a CBIR system with region-specific BoVW (Bag of Visual Words) model. It uses intensity information of the lesion region and the intensity change of the surrounding of lesion region. The proposed system also summarized that normalization and whitening gave better performance in retrieving the image. Their proposed model obtained an average precision value of 91.0%. However, the brain tumors in MRI images are outlined manually which is not proficient and advantageous. When the contour of the segmentation is far from the tumor boundary, the method strives to catch the intensity variation encompassing the tumor.

Furthermore, Huang et al. [6] described a CBIR based system. It uses BoVW model to extract information features and REML (Rank Error-based Metric Learning) algorithm to decrease the gap between low-level features and high-level connotative concepts. This system outperformed many methods with an average precision of 93.1%. However, the partition learning that was used in the system to obtain discriminative features, many of the times

gives local optimal solutions. Moreover, the distance matric REML is globally linear and strives to determine multiple regional projections for various regions in the feature space.

III. Proposed Methodology

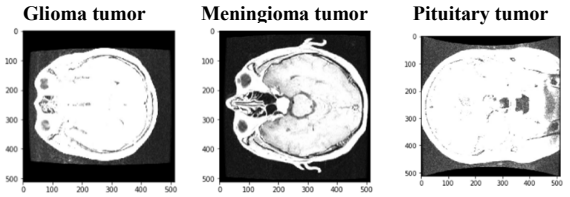
This section illustrates the detailed methodology that is used to classify brain tumors. It has the following steps: dataset collection and description of the dataset, preprocessing the images, classification of brain tumors using CNN and performance measures for the system. Each step of the proposed system is outlined as follows:

A. Dataset Collection and Description

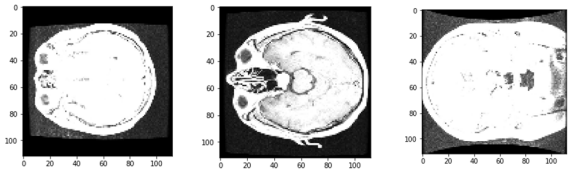
The brain tumor dataset retrieved from [7] comprises 3064 T1-weighted contrast-enhanced images. The dataset is collected from 233 patients with three kinds of brain tumor: meningioma (708 images), glioma (1426 images), and pituitary tumor (930 images). The dataset is arranged in MATLAB data format (.mat file). Each file reserves a struct comprising various fields of an image. These fields include the label (1 for meningioma, 2 for glioma, 3 for pituitary tumor), PID (patient ID), image data, tumor border, and tumor mask.

B. Preprocessing

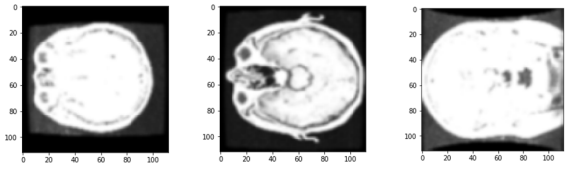
We used two fields, i.e. label and image data from each of the mat files. Fig. 1 illustrates the stepwise preprocessing of the work. The original images are 512×512 in size.



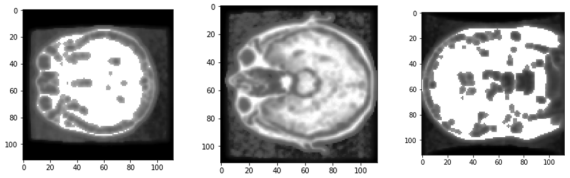
(a) Original image (512×512)



(b) Resized image (112×112)

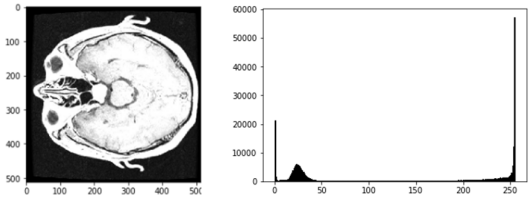


(c) Filtered image (112×112)

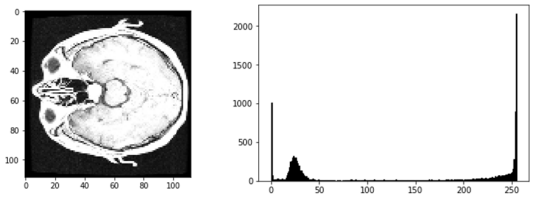


(d) Equalized image (112×112)

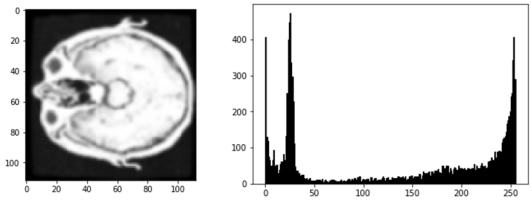
Fig. 1. Stepwise preprocessing outcome for tumor classification



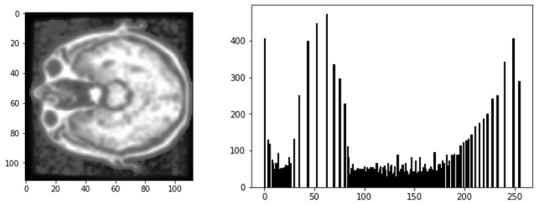
(a). Histogram of original image (512×512)



(b). Histogram of resized image (112×112)



(c). Histogram of filtered image (112×112)



(d). Histogram of equalized image (112×112)

Fig. 2. Histogram of different preprocessing step image

First, we resize the images in 112×112 and then apply Gaussian filter with kernel size 5×5. Gaussian filter smooths the image. In 2-D, Gaussian function has the following form:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\pi\sigma^2}} \quad (1)$$

where x denotes the distance from the origin in the horizontal axis, y denotes the distance from the origin in the vertical axis, and σ denotes the standard deviation of the Gaussian distribution.

Then, histogram equalization is used on each of the images. Histogram graphically illustrates the intensity distribution of an image. Histogram Equalization is an image processing technique, and it enhances the contrast of images. It manages to achieve this by spreading out the most frequent intensity values of the image. Fig. 2 shows the histogram of different preprocessing step image.

The total dataset is divided into three sets. These sets include train set (2051 images), validation set (513 images) and test set (500 images). Hence, the train, validation and test set contain 66.9%, 16.7%, 16.3% of the dataset respectively. The train set is used to train the CNN model.

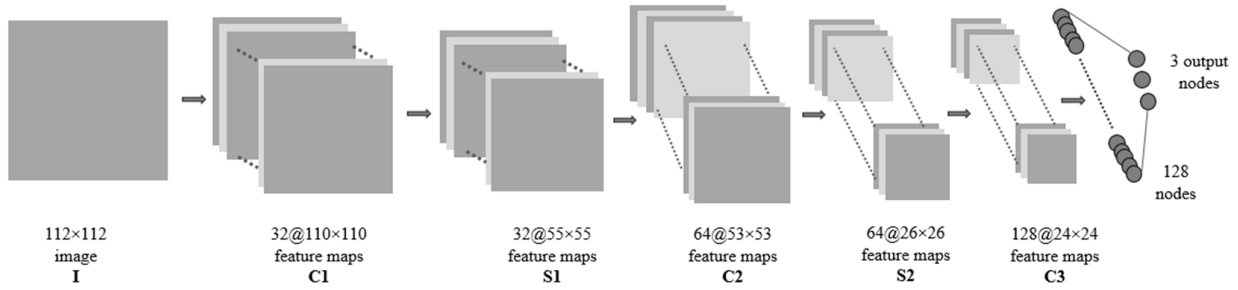


Fig. 3. CNN architecture of the model (**I, C, S** represent input image, convolution and subsampling respectively)

The validation set gives an unbiased assessment of the model while tuning the model's hyperparameters. The test set determines how well the model has been trained by evaluating the accuracy of the picked approach.

C. Tumor Classification using CNN

CNNs are unusually multi-layer neural networks. It is an effective recognition algorithm applied in pattern recognition and image processing [8]. It simulates the biological neural network through shared weight network structure. CNN is broadly practiced in the field of computer vision problem. It comprises the property of parameter sharing which decreases the number of parameters needed for the model compared to ANN (Artificial Neural Network). Moreover, the quality of the features extracted by CNN is very standard. In the CNN model we apply a series of convolution + pooling operations, followed by a fully connected layer. As we are implementing multiclass classification, the final output layer is softmax.

In our work, we used convolution on the input image using a convolution filter to generate a feature map. A 3×3 kernel is used as convolution filter. Strides = (1, 1), padding = 'valid' is used for this operation. Stride defines how much the convolution filter is moved at each step. Padding is 'valid' represents zero padding. Fig. 3 illustrates the CNN architecture that is used for this work. It includes 3 convolution layers (C1, C2, C3) with a kernel size of 5×5, 2 subsampling layers (S1, S2) with a pooling window of 2×2.

In first convolution (C1), the input image I with size 112 ×112 convolved with 32 filters with zero padding and stride = 1 which produced 110×110 sized 32 convolved features. Hence, total parameters for this operation are (3×3+1) ×32 = 320. These parameters will be learned by the algorithms in the training phase. In every convolution layer ReLU (Rectified Linear Unit) is used for activation function. It increases the non-linearity in the images. The features that are generated by the C1 layer are fed into first subsampling layer (S1). In subsampling layer max pooling is used with a window size of 2×2. Hence, 32 pooled feature maps each with size of 55×55 are produced. Pooling decreases the number of parameters which reduces the training time. Then a dropout layer is used with a dropout rate of 25%. Dropout is a regularization technique that is used to reduce overfitting [9]. Dropout is only used during training phase of the model. In dropout, some randomly selected neurons are ignored during training phase according to the dropout rate, i.e. their contribution in activating downstream neurons is shut down on the forward pass and during the backward pass, any weight updates are not applied on the neurons.

TABLE I. TRAINABLE PARAMETERS FOR CNN MODEL

Layers	Parameters
Conv2D (C1)	320
Conv2D (C2)	18,496
Conv2D (C2)	73,856
Dense (1)	9,437,312
Dense (2)	387
Total	**9,530,371**

In second convolution layer (C2), 64 convolution filters are used on the previous layer's features map that result in 64 convolved features map each with a size of 53×53. Hence, total parameters for this operation are (32×3×3+1) ×64 = 18,496. These features are fed to second subsampling layer (S2) that also uses max pooling with window size of 2×2. Therefore, 64 pooled features map each with 26×26 are generated. Then, a dropout layer with 25% dropout rate is used.

In the third and final convolution layer (C3) 128 filters are used which results in 128 features map each with size 24×24. Here, total parameters are (64×3×3+1) ×128 = 73,856. Then we used another dropout layer with a dropout rate of 40% and flatten the output. Then the generated features are fed to a dense layer with 128 nodes. Here, total parameters are (24×24×128+1) ×128 = 9437312. Then, again a dropout layer with 30% dropout rate is used. Finally, we used another dense layer with softmax activation and 3 nodes for the classification. Hence, total parameters will be (128+1) ×3 = 387. The total trainable parameters for the model are shown in Table I.

D. Performance Evaluation

The performance of the system is analyzed using confusion matrix, precision, recall, f1-score, support and ROC (Receiver Operating Characteristic) curve. For calculating precision, recall, f1-score the following formulas are used.

$$Precision = \frac{True\ Positive}{True\ positive + False\ Positive} \quad (2)$$

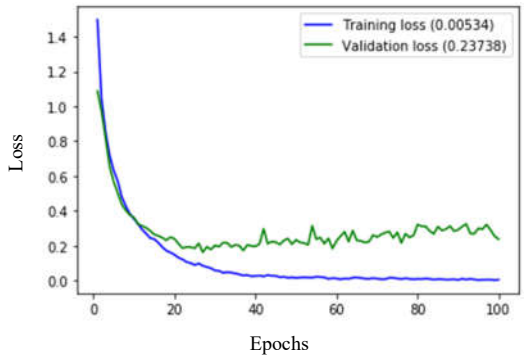
$$Recall = \frac{True\ Positive}{True\ positive + False\ Negative} \quad (3)$$

$$F1\text{-}score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

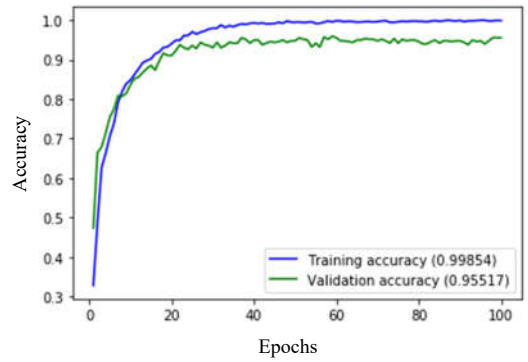
ROC curve represents a plot of the true positive rate (Sensitivity) versus the false positive rate (100-Specificity) for different cut-off points. In micro-average ROC scheme, the necessary individual value for different classes is summed up to compute the average while a macro-average ROC curve will calculate the required values for each class independently and then take the average. The area under the ROC curve (AUC) is a measure of how well the model is distinguishing between different classes. An area of 1 is considered to be the best in evaluating test cases.

IV. Result Analysis

After developing the CNN model, the model is compiled with Adam optimizer. It is a very powerful optimization algorithm used for training deep neural networks. It combines the advantages of two optimization methods i.e. AdaGrad (Adaptive Gradient Algorithm) and RMSProp (Root Mean Square Propagation) [10]. For fitting the training data with the model, a batch size of 256 and 100 epochs are used. Fig. 4 illustrates the loss and accuracy of the training set and validation set for 100 epochs. After evaluating the model on test data, we achieved a loss of 28.16% and an accuracy of 94.39%. Fig. 5 depicts the ROC curve. The closer the ROC curve is to the upper left corner, the higher the overall accuracy [11]. It also includes micro-average and macro-average ROC curve. Furthermore, it shows that the area under the curve ranges from 0.99 to 1 for the different classes which are desirable for the test cases.Table II represents the confusion matrix for the CNN model. For testing the model, 500



(a). Loss vs Epochs



(b). Accuracy vs Epochs

Fig. 4. Loss and Accuracy curve for 100 epochs

TABLE II. CONFUSION MATRIX FOR THE MODEL

	Class 1 (Glioma)	Class 2 (Meningioma)	Class 3 (Pituitary)
Class 1 (Glioma)	88	12	3
Class 2 (Meningioma)	13	213	0
Class 3 (Pituitary)	0	1	170

images are used. The model identifies class 3 (pituitary tumor) more accurately than other classes. From 171 pituitary tumor images, it can successfully recognize 170 as pituitary while locating 1 as meningioma and 0 as glioma. The ability of the model to predict the other tumor class is also satisfactory.

Table III illustrates some performance measure indices for the model. It shows that precision, recall, f1-score and support are used to evaluate the model. The model achieved the precision of 88%, 94%, 98% for Class 1(Glioma), Class 2 (Meningioma) and Class 3 (Pituitary) respectively which gives the average precision of 93.33%.

A comparison is exhibited in Table IV between several methods from different papers which used the same dataset. It shows that our model gives the highest average precision of 93.33% among the different methods. The CBIR method with BoVW and REML [6] also performs well with the average precision of 93.1%.

TABLE III. PERFORMANCE MEASURE INDICES

Class	Precision	Recall	F1-score	Support
Class 1 (Glioma)	0.88	0.85	0.87	103
Class 2 (Meningioma)	0.94	0.95	0.94	226
Class 3 (Pituitary)	0.98	0.99	0.99	171

TABLE IV. COMPARISON WITH EXISTING METHODS IN TESTING PHASE

Methods	**Average Precision (%)**
CBIR [4]	89.3
CBIR with BoVW [5]	91
CBIR with BoVW and REML [6]	93.1
Proposed model using CNN	**93.33**

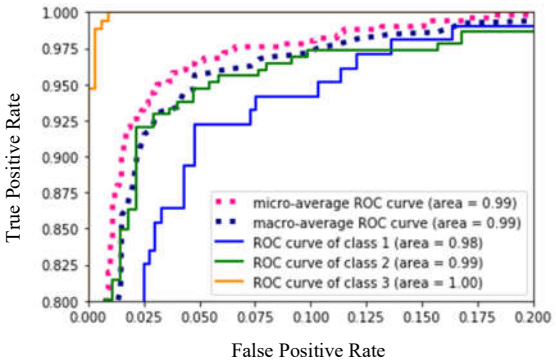


Fig. 5. ROC (Receiver Operating Characteristic) curve of the system

V. CONCLUSIONS

Brain tumor classification is very crucial in the domain of medical science. In this paper, we concentrated on developing a CNN classifier which classifies among three important tumor classes (glioma, meningioma, pituitary). Initially, the proposed system preprocesses the image data. The preprocessing includes filtering images using Gaussian filter and applying histogram equalization technique on the filtered images. Then the system classifies the images using the CNN model. The number of parameters of the model is too high, and the model is trained on a significantly small amount of data. Hence, there is a possibility of overfitting. To prevent the overfitting, a regularization technique, i.e. dropout regularization is used on the model. It helps the model to concentrate on the most prominent patterns during the training phase. Therefore, there is a better chance of generalization which keeps the model stable. The model ends up with the accuracy of 94.39% and an average precision of 93.33%. Hence, the CNN classifier will be very significant in the medical field and in saving precious lives.

REFERENCES

[1] cancer.org, 'Key Statistics for Brain and Spinal Cord Tumors', January, 2019. [Online]. Available: https://www.cancer.org/cancer/brain-spinal-cord-tumors-adults/about/key-statistics.html, [Accessed: Jan. 9, 2019].

[2] Q. T. Ostrom, H. Gittleman, J. Xu, C. Kromer, Y. Wolinsky, C. Kruchko, and J. S. Barnholtz-Sloan, "Cbtrus statistical report: primary brain and other central nervous system tumors diagnosed in the united states in 2009–2013," *Neuro-oncology*, vol. 18, no. suppl5, pp. v1–v75, 2016.

[3] mayoclinic.org, 'Brain tumor'. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/brain-tumor/symptoms-causes/syc-20350084, [Accessed: Dec. 15, 2018].

[4] W. Yang, Q. Feng, M. Yu, Z. Lu, Y. Gao, Y. Xu, and W. Chen, "Content-based retrieval of brain tumor in contrast-enhanced mri images using tumor margin information and learned distance metric," *Medical physics*, vol. 39,no. 11, pp. 6929–6942, 2012.

[5] M. Huang, W. Yang, M. Yu, Z. Lu, Q. Feng, and W. Chen, "Retrieval of brain tumors with region-specific bag-of-visual-words representations in contrast-enhanced mri images," *Computational and mathematical methods in medicine*, vol. 2012, 2012.

[6] M. Huang, W. Yang, Y. Wu, J. Jiang, Y. Gao, Y. Chen, Q. Feng, W. Chen, and Z. Lu, "Content-based image retrieval using spatial layout information in brain tumor t1-weighted contrast-enhanced mr images," *PloS one*, vol. 9, no. 7, p. e102754, 2014.

[7] J. Cheng, 'brain tumor dataset', 2017. [Online]. Available: https://figshare.com/articles/brain_tumor_dataset/1512427, [Accessed: Oct. 5, 2018].

[8] T. Liu, S. Fang, Y. Zhao, P. Wang, and J. Zhang, "Implementation of training convolutional neural networks," *arXiv preprint arXiv:1506.01195*, 2015.

[9] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov,"Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[10] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[11] M. H. Zweig and G. Campbell, "Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine." *Clinical chemistry*, vol. 39, no. 4, pp. 561–577, 1993.

Improved Segmentation Approach for Plant Disease Detection

Md. Arifur Rahman
Department of Computer Science & Engineering
Ahsanullah University of Science & Technology
Dhaka, Bangladesh
rafiddhakacollege@gmail.com

Md. Mukitul Islam
Department of Computer Science & Engineering
Ahsanullah University of Science & Technology
Dhaka, Bangladesh
mukitulislamratul@gmail.com

G M Shahir Mahdee
Department of Computer Science & Engineering
Ahsanullah University of Science & Technology
Dhaka, Bangladesh
shahirmahdee1@gmail.com

Md. Wasi Ul Kabir
Department of Computer Science & Engineering
Ahsanullah University of Science & Technology
Dhaka, Bangladesh
wasi.cse@gmail.com

***Abstract*—The Agricultural sector plays a vital role in sustainable economic growth and food security. However, crop diseases often cause a great threat in achieving this goal. As such, a successful outcome depends entirely on proper detection and classification of plant diseases. This has created many opportunities of new possibilities for researchers. Nowadays, a lot of work is being done to recognize and classify plant diseases more precisely using computer vision. The objective of this research is to create a methodology that will provide a better solution to classify plant diseases. This work mainly focuses on implementing an improved segmentation technique using a combination of thresholding and morphological operations. For classification, we have used the deep neural network. Our proposed method has achieved 99.25% accuracy in Plant Village database.**

***Index Terms*—segmentation, image processing, disease detection, plant leaf disease, deep neural network.**

I. Introduction

Bangladesh is a developing country and it is very important to increase food production in order to support its rapidly growing economy. Bangladesh is trying to become self-sufficient in producing food. However, crop loss due to various diseases poses a great threat in achieving that goal. There are several plant leaf diseases in Bangladesh. Leaf disease detection using one's bare eye is quite difficult for both general people as well as farmers. Furthermore, consulting an expert in this field is not always possible for the rural farmers of Bangladesh. As a result, farmers face several problems in controlling the diseases on crops and often have to face a massive loss in terms of production. Sometimes, it is also very difficult to identify a disease as symptoms of many different leaf diseases are quite similar. Most of the times, farmers fail to take the necessary steps even if they may have been able to identify the problem. Moreover, in some rural areas of Bangladesh, farmers cannot get rid of these problems because they do not have proper knowledge or education on how to do so. Most of the cases involve leaf diseases which are not recognized properly, and farmers end up using insecticides which may not be suitable for that particular disease. If an artificial intelligent (AI) system can detect disease from an image of any diseased leaf, it will be incredibly helpful for this industry. From this thought, we are going to provide new techniques of image pre-processing and a new combination of feature set from the processed images to make a trained model that will show how accurate our methodology is to detect diseases accurately.

In this work, the proposed new methodology can detect leaf diseases accurately. Among various plant leaf diseases, tomato leaf disease is one of them. This work is based on three tomato leaf diseases, they are - Bacterial Spot, Late Blight and Septorial Spot. These diseases are common tomato leaf diseases in Bangladesh. The method we proposed consists of four main phases which are: image enhancement, segmentation, feature extraction and classification. This work focuses mainly on image segmentation and feature extraction techniques.

II. Related Works

Tm et al. [1], have used convolution neural network to detect tomato leaf diseases. Their proposed methodology has three stages- data acquisition, data pre-processing and classification. The plant village dataset is used as dataset. For image pre-processing, the image was normalized using the mean and the standard deviation. This is also known as Z-score method. For classification, Le-Net architecture is used for convolution neural network. Though the proposed method gave a high accuracy, it is also computationally heavy which may cause problems for larger datasets.

C. G. Dhaware et al. [2], represent techniques for detecting infected parts of plant leaves. Images are pre-processed with the help of RGB to HSV conversion technique. Cluster based method background subtraction is used for segmentation. Correlation, energy, homogeneity etc. features are also used. For image classification, support vector machine (SVM) technique

is used. In this approach, the dataset was very small and thus showed high accuracy. Additionally, there was no exact diseases classification of infected leaf.

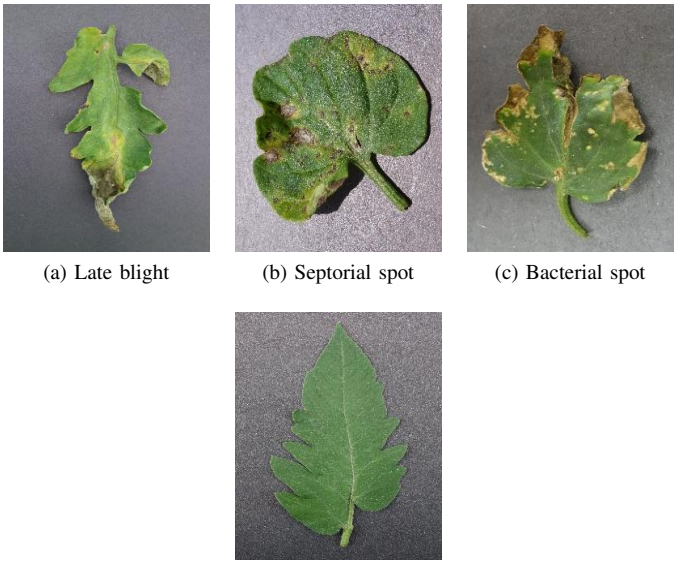
U. Mokhtar et al. [3], describe a method of detecting tomato leaf diseases detection based on support vector machines. In image pre-processing phase they had used three techniques such as leaf image isolation and extraction, image resizing and background subtraction technique with some morphological operations. Color and texture have been considered as features. For extracting texture Gabor filter is used. For the classification phase, support vector machine technique is used.

T. N. Tete et al. [4], describe a method of detecting plant disease based on a clustering method. This process is divided into four parts- pre-processing, segmentation, feature extraction and classification. Infected parts are segmented with the help of thresholding and k-mean clustering method. For classification, artificial neural network algorithm is used.

III. Dataset Description

Images are collected from publicly available Plant Village [5] database. It contains about 54306 images of diseased as well as healthy leaves of 14 different types of crops. Fig 1 shows some sample images of tomato leaf from the dataset. Here 2000 images of tomato leaves are analyzed which have following class labels:

- Late blight affected tomato leaf.
- Septorial spot affected tomato leaf.
- Bacterial spot affected tomato leaf.
- Healthy or Non-diseased tomato leaf.



(a) Late blight (b) Septorial spot (c) Bacterial spot

(d) Healthy leaf

Fig. 1: Images of infected and healthy tomato leaves

IV. Proposed Methodology

The proposed method includes four important stages namely: enhancement, segmentation, feature extraction and classification. Steps of proposed methodology are shown in Fig. 2

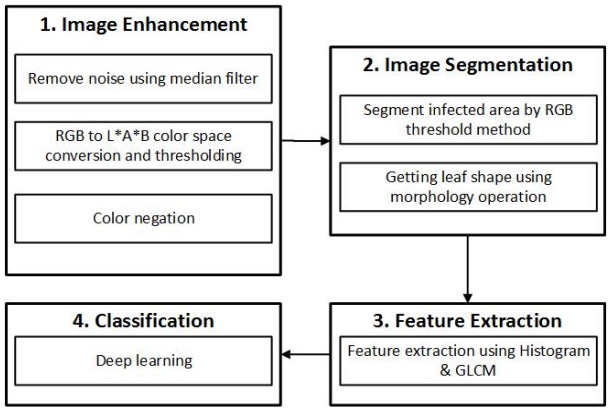


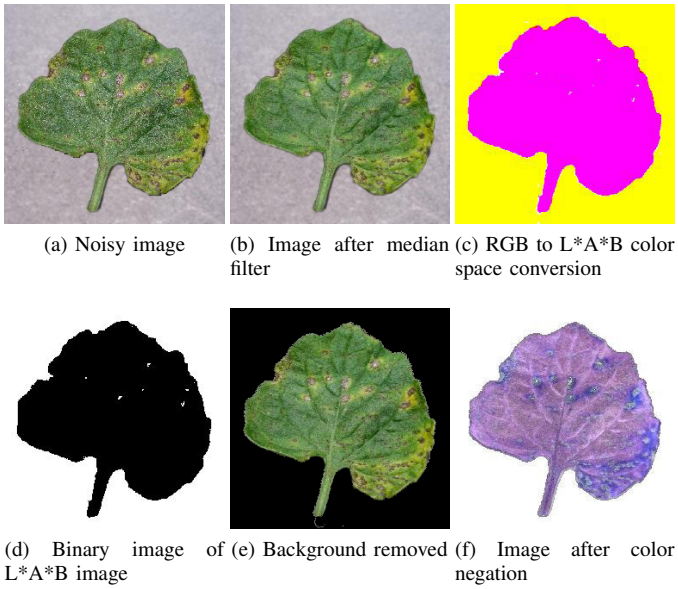
Fig. 2: Flowchart of our proposed methodology

A. Image Enhancement

Image enhancement technique is used so that input image will be more detailed and become more suitable for further experiment and analysis. Here, there are three image enhancement techniques-

- Median filter
- RGB to L*A*B color space conversion and threshold
- Color negation

First we have used median filter technique, which provides almost noise free image. On this less noisy image RGB to L*A*B color space conversion and threshold techniques are used [7], which removes background of the leaf image and provides an image having a black background. Finally, color negation technique is applied on that image which will be helpful in segmentation process. Image enhancement process are presented in Fig. 3.



(a) Noisy image (b) Image after median filter (c) RGB to L*A*B color space conversion

(d) Binary image of L*A*B image (e) Background removed (f) Image after color negation

Fig. 3: Image enhancement

B. Image Segmentation

Here two image segmentation techniques are applied. These techniques are -

- RGB threshold
- Morphological operation.

After color negation, the healthy green part of the leaf becomes magenta and the infected area become a different color. Then RGB threshold technique is used and threshold value based on blue and red level intensity as magenta color is located midway between red and blue. After RGB threshold we would be able to get infected part of the leaf. After that, morphological operation is used to achieve the leaf shape. Here we used 3*3 cube shaped structural element. If A is our color negated image and B is the structural element then mathematical formulation of output image are presented in equation 1 and output of segmentation process are shown in Fig. 4.

$$Y = (A \oplus B - A)' \tag{1}$$

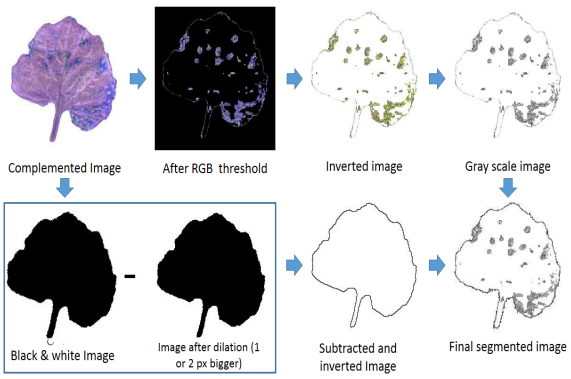


Fig. 4: Output of segmentation process

C. Feature Extraction

Some unique properties of each image from the dataset are needed to be extracted for machine learning purposes. Those unique properties are known as features of that data and the process of extracting the features is called feature extraction. In this phase, we have extracted features from GLCM and graylevel histogram and the size of feature vector is 260. Fig. 5 shows a sample image which is used for feature extraction.

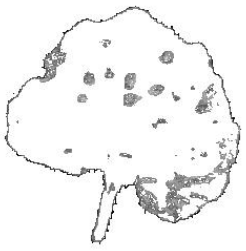


Fig. 5: Segmented grayscale image for feature extraction

1) Gray-Level Co-Occurrence Matrix (GLCM): A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM). The GLCM characterizes the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image. By creating a GLCM and then we can extract statistical measures from this GLCM. We can retrieve following statistical information from a GLCM of an image - contrast, correlation, energy, homogeneity [8] [9] [10]. Mathematical formulation of GLCM features are given below TABLE I-

TABLE I: Mathematical Formulation of GLCM Features

Feature	Mathematical Formulation
Contrast	$\sum_{i,j} \lvert i-j \rvert^2 p(i,j)$
Correlation	$\sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j}$
Energy	$\sum_{i,j} p(i,j)^2$
Homogeneity	$\sum_{i,j} \frac{p(i,j)}{1+\lvert i-j \rvert}$

Each of the GLCM statistical features signify some unique property of the image. Contrast is a measurement which represents local variations present in the image. Value of correlation indicates a linear relationship between gray levels of pixel pairs. Energy measures the repetition of pixel pairs of an image which is also known as textural uniformity. Homogeneity has a relationship with contrast and energy feature of GLCM. At constant energy, homogeneity is inversely proportional to contrast.

2) Gray level Histogram: An image is a set of pixels or points of different intensity values [11]. A histogram for any image is such representation where it shows the number of pixels in an image at each different intensity value found in that image. There are 256 different possible intensities for an 8-bit grayscale image. We have extracted 256 different 8-bit values from gray level histogram for each segmented image like Fig. 5.

D. Classification

Neural networks are a set of algorithms that are designed to recognize patterns. A deep learning neural network is an artificial neural network with multiple layers between the input and output layers. The deep learning neural network finds the correct mathematical manipulation to turn the input into the output, whether it be a linear relationship or a non-linear relationship. In this work, deep neural network is used for classification.

We have used Python as programming language for setting up our deep learning: neural network model. Deep learning library Keras has been used in this purpose. Keras is a high-level neural networks API, written in Python and capable of running on top of TensorFlow.

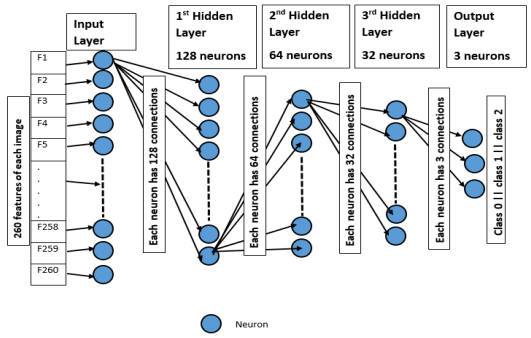


Fig. 6: Deep learning neural network model

Fig. 6 shows our deep learning neural network model. Our neural network model is a fully connected network, i.e., each layer of neuron is connected with all other neurons of the next layer. To simplify the image shown in Fig. 6, we are showing how one neuron of each layer is connected with all other neurons of next layer.

In our deep neural network model, there is one input layer, one output layer and three hidden layers. The input layer has 260 neurons as our dataset has 260 features of each data and the output layer has three neurons as we are classifying among three disease classes. In hidden layer neurons, we have used Rectified Linear Unit (ReLU) as the activation function [12] and in output layer neurons we have used Softmax as the activation function.

In this model, we have used Stochastic Gradient Descent (SGD) for weight update. SGD does sequential weight update or one update at a time. SGD slowly decrease the learning rate and SGD almost certainly converging to a local or the global minimum. We have used Minimum Squared Error (MSE) loss function. MSE is the sum of squared distances between our target variable and predicted values.

$$MSE = \frac{\sum_{i=1}^{n} (y_i - y_i^p)^2}{n} \tag{2}$$

In the equation 2, y_i and y_i^p represents actual and predicted value respectively.

V. Experiment & Result Analysis

For experiment, we have selected three diseases from Plant Village database to evaluate our proposed method. The disease name and corresponding number of images are shown in TABLE II. We have used 80% of the total image for training and the remaining 20% image data for testing. All of these images have been chosen randomly.

TABLE II: Disease Name and Corresponding Number of Images

Disease Name	Number of Images
Late Blight	500
Septorial Spot	500
Bacterial Spot	500
Healthy leaf	500
Total	2000

TABLE III: Performance Metric

	Precision	Recall	F1 Score	Support
0: Late Blight	0.97	1.00	0.99	100
1: Septorial Spot	1.00	0.97	0.98	100
2: Bacterial Spot	1.00	1.00	1.00	100
3: Healthy leaf	1.00	1.00	1.00	100
Avg/Total	**0.99**	**0.99**	**0.99**	**400**

TABLE IV: Confusion Matrix

```
[[100   0   0   0]
 [3    97   0   0]
 [0     0 100   0]]
 [0     0   0 100]]
```

The performance metric for histogram & GLCM feature combination is presented in TABLE III. The average precision, recall and f1-score hold a good score for our proposed feature combination. TABLE IV shows the confusion matrix, where 100 of 100 Late Blight disease, 97 of 100 Septorial spot disease, 100 of 100 Bacterial Leaf Spot disease and 100 of 100 healthy leaves are detected correctly. In this experiment, our proposed method achieves 99.25% accuracy.

We also compared our segmentation method with widely used k-mean clustering segmentation method [13] [14]. Both methods are tested under the same environment and the testing environment setups are shown in TABLE V.

TABLE V: Environment Setup

Segmentation Method	Features	Classifier
Our Proposed Method	Histogram & GLCM	Deep learning
K-mean Clustering	Histogram & GLCM	Deep learning

The precision and recall curve is a widely used metric to understand which model performs better. Fig. 7 and Fig. 8 shows precision and recall curve of k-mean method and our proposed method respectively. From these figures, we can see that our proposed method outperforms the K-mean clustering method.

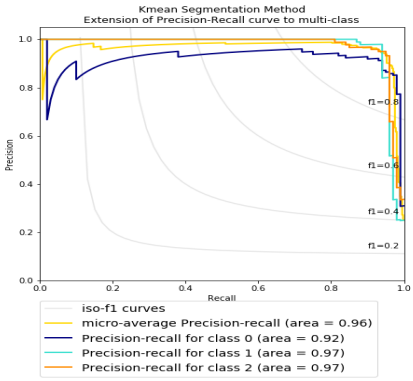


Fig. 7: Precision and recall curve of k-mean model

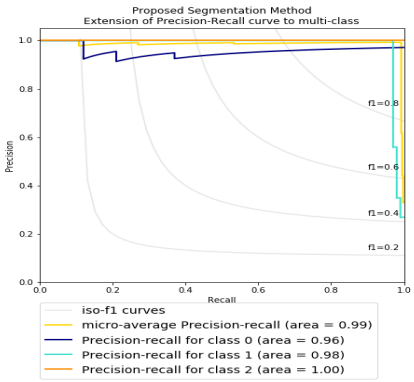


Fig. 8: Precision and recall curve of our proposed model

TABLE VI: Accuracy Comparison

	Our Proposed Method	**K-mean Clustering**
Accuracy	99.25%	93.25%

TABLE VI clearly shows that our proposed method is more effective and accurate than k-mean clustering method.

VI. Conclusion

The proposed method of this research focuses on improved segmentation approach for plant leaf disease identification. It is an effective segmentation technique and doesn't require any kind of heavy computational power. However, there is a scope to improve our work in the future by reducing feature vector size as well as keeping the classification accuracy higher. This research work can be extended by integrating a larger number of diseases of different species of plants.

References

[1] P. Tm, A. Pranathi, K. SaiAshritha, N. B. Chittaragi, and S. G. Koolagudi, "Tomato leaf disease detection using convolutional neural networks". In *Proc. of the Eleventh International Conference on Contemporary Computing* (IC3), pages 15, Aug 2018.

[2] C. G. Dhaware and K. H. Wanjale, "A modern approach for plant leaf disease classification which depends on leaf image processing". In *Proc. of the International Conference on Computer Communication and Informatics (ICCCI)*, pages 14, Jan 2017.

[3] U. Mokhtar, M. A. S. Ali, A. E. Hassenian, and H. Hefny, "Tomato leaves diseases detection approach based on support vector machines". In *Proc. of the 11th International Computer Engineering Conference (ICENCO)*, pages 246250, Dec 2015.

[4] T. N. Tete and S. Kamlu, "Detection of plant disease using threshold, k-mean cluster and ann algorithm". In *Proc. of the 2nd International Conference for Convergence in Technology (I2CT)*, pages 523526, April 2017.

[5] spMohanty, Pennsylvania State University, Plantvillage. [Online]. Available: https://github.com/spMohanty/PlantVillage-Dataset. [Accessed: 12-November-2018].

[6] Monzurul Islam, Anh Dinh, Khan Wahid and Pankaj Bhowmik, "Detection of potato diseases using image segmentation and multiclass support vector machine". In *Proc. of the IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 14, April 2017.

[7] Dhanashree Gadkari, "Image quality analysis using glcm", Thesis, University of Central Florida, Florida, 2000.

[8] P. Mohanaiah, P. Sathyanarayana, and L. GuruKumar, "Image texture feature extraction using glcm approach". In *International Journal of Scientific and Research Publications*, vol. 3(5), May 2013.

[9] M. Dhakate, "Diagnosis of pomegranate plant diseases using neural network". In *Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*, pages 14, Dec 2015.

[10] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift". In *Proc. of the 32nd International Conference on Machine Learning*, pages 448-456, 2015.

[11] K. R. Aravind, P. Raja, K. V. Mukesh, R. Aniirudh, R. Ashiwin, and C. Szczepanski. "Disease classification in maize crop using bag of features and multiclass support vector machine". In *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, pages 11911196, Jan 2018.

[12] H. Q. Cap, K. Suwa, E. Fujita, S. Kagiwada, H. Uga, and H. Iyatomi. "A deep learning approach for on-site plant leaf detection". In *2018 IEEE 14th International Colloquium on Signal Processing Its Applications (CSPA)*, pages 118122, March 2018.

[13] S. Maity, S. Sarkar, A. v. Tapadar, A. Dutta, S. Biswas, S. Nayek, and P. Saha, "Fault area detection in leaf diseases using k-means clustering". In *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)*, pages 15381542, May 2018.

[14] R. Anand, S. Veni, and J. Aravinth. "An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method". In *2016 International Conference on Recent Trends in Information Technology (ICRTIT)*, pages 16, April 2016.

An Automated Approach to Detect Breast Cancer Tissue Using Ultrasound Images

T. M. Shahriar Sazzad
Dept.: CS
American International University Bangladesh
Dhaka, Bangladesh
tss5standrews@gmail.com

Misbah Ul Hoque
Dept.: CSE
Green University of Bangladesh
Dhaka, Bangladesh
misbah.hoque.cse@gmail.com

A.K.M. Ashek Farabi
Dept.: CSE
Green University of Bangladesh
Dhaka, Bangladesh
akmashekfarabi@gmail.com

Ismail Hosen
Dept.: CSE
Green University of Bangladesh
Dhaka, Bangladesh
rubel7854@gmail.com

Mursheda Akter Somi
Dept.: CSE
Green University of Bangladesh
Dhaka, Bangladesh
somiislam8@gmail.com

Abstract— **Cancer is a kind of cell that develops out of control of the regular drives that normalizes growth. Breast cancer is one of the deadly forms of cancers for the human. Initial detection of cancerous tissues is likely to monitor the patient's strength to deal with the disease and live more. Electronic modalities are applied to diagnose breast cancers. Due to low cost and safety, ultrasound scanning is considered as one of the most frequently used electronic modalities for breast cancer detection. This research study proposed an automated approach for breast cancer detection using ultrasound gray-scale images. To this end, enhancement operation was used at the beginning to minimize the gray-scale color shifts. Median filter was used to facilitate better segmentation removing maximum unwanted noises. Threshold-based segmentation (Threshold based OTSU) was used as the test images used in this study are gray-scale images. Finally, necessary feature information that is used in the pathology laboratory by pathology experts is used in this study to identify the ROI (Region of interests) to detect breast cancer tissues. This study proposed approach achieved an average accuracy rate of 93.11%. The depicted experimental outcomes of the proposed approach outperform other contemporary existing available approaches in terms of accuracy while maintaining the medical experts' acceptable accuracy rate.**

Keywords— Breast Cancer, Ultrasound Gray-scale Images, Threshold based OTSU.

I. INTRODUCTION

CANCER is the uncontrolled growth of abnormal cells anywhere in a body [1]. There are more than two hundred kinds of cancer including breast cancer, lung cancer, ovarian cancer, leukemia, prostate cancer, cervical cancer, pancreatic cancer, brain tumor etc. [1]. The most widely recognized cancer is breast cancer that found more in ladies both within the developed and less developed countries. Breast cancer is believed to be an ailment of the developed world, right around half of the breast cancer cases and 58% of all deaths happen in less developed nations (GLOBOCAN 2008) [2]. Breast cancer creates breast cancer tissues [3]. Signs of breast cancer may incorporate an irregularity in the breast, a trade in breast shape, dimpling of the skin, and liquid originating from the nipple, a newly inverted nipple, or a purple or scaly patch of skin [4]. Lack of early detection of breast cancer is high proportion in women. It becomes severe in the late stage and the lack of adequate diagnosis and treatment facilities [2]. Early detection of breast cancer can be possible by using breast ultrasound screening.

Breast ultrasound screening (BUS) is one of the most widely used diagnostic tools to utilize and classify abnormalities of the breast. It is most appropriate for substantial scale breast cancer screening and determination in low-asset nations and areas. Automatic breast ultrasound image segmentation is vital for most cancer diagnosis and arranging treatments [5]. Ultrasound is an easy and painless system that uses sound waves to generate images of the inner of the body [5].

This research study focused on medical image processing approaches for breast cancer tissue identification and proposed an automatic breast cancer detection approach where ultrasounds images are processed. To detect the cancer region accurately this research study incorporated ultrasounds images and were processed in two different stages which include pre-processing (Enhancement, Filtering, Segmentation), and post-processing (Feature selection, Identification of ROI, Classification).

II. RELATED WORKS

Many research works have been carried out on breast cancer analysis. For breast cancer segmentation research work in [6] incorporated ultrasound images and allowed faster and more accurate radiographic contouring. The aim of the paper was to present a new method for ultrasound image segmentation of breast lesions. The Dice Index was greater than 80%. However; this existing work depends on the radiologist and not fully automated (semi-automated).

Another research work [7] proposed an automated approach to achieve error-free detection for breast cancer. The study conducted on three stages including Pre-processed Image, Pectoral Muscle Removal, and Segmented Breast Profile. Decision trees, K-nearest Neighbor (KNN), Support Vector Machine (SVM) and Random Forest classifier were used and their performances were compared. The Random Forest Classifier performed better compared to other methods in breast cancer detection.

Research work [8] reported that deep learning approach performs well in general object detection. Additionally,

convolutional neural network-based method, Single Shot Multi-Box Detector (SSD) was also used. A new dataset was collected consisting of 579 benign and 464 malignant lesion cases and the result showed that the convolutional neural network-based method achieved the best performance in terms of both precision and recall. But the resultant bounding box carries extra regions which are not cancer tissue at all. Their future work indicates that they need some automated segmentation techniques to detect tumor areas.

Proposed approach by [9] mentioned that the nodule of the breast can be automatically segmented by neutrosophic set and watershed method with some extra unwanted regions.

Research work of [10] applied watershed algorithm in breast cancer ultrasound images to extract precise contours of the tumors. During pre-processing median filter used to preserve the edges of the tumor lesion. Threshold was calculated for background and foreground areas. As a result, the method defined the contours of breast tumors in ultrasound images, and the results were close to ground truth data. However; their work failed to detect breast tumor region/regions where shadows are present inside the images.

III. Proposed Method

The working steps of the proposed method are shown in Fig. 1. The first step begins with image acquisition. Further steps are subcategorized as pre-processing and post-processing. Pre-processing includes image enhancement, filter operation and segmentation. Post-processing includes feature extraction, identification of ROI and finally classification approach to improve the accuracy rate.

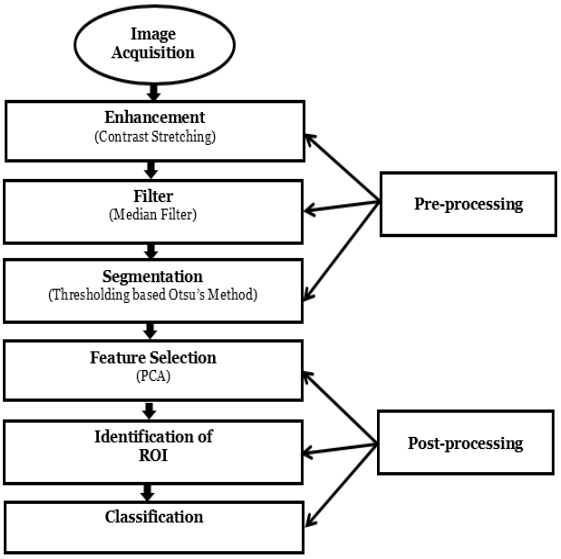


Fig. 1: Proposed method to detect breast cancer tissues

A. Image Acquisition

Image acquisition is the first step for image processing approach. Ultrasound images were collected from the pathology laboratory for this study test data set. Collected test data set was .PNG and each image size was 350 × 350.

B. Pre-processing

Pre-processing has three stages which include

a. Enhancement
b. Filter
c. Segmentation

a. Enhancement

Ultrasound RGB grayscale images were converted into HSV format as HSV color model performs better [18]. Contrast enhancement was performed on V channel. The resultant V has subtracted from initial V and the resultant images were converted back to RGB. An enhanced image of this study is shown in Fig. 2.

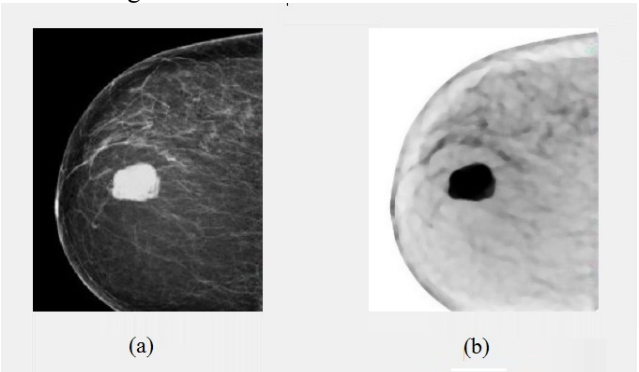


Fig 2: (a) Original test image, (b) Applied enhancement approach (contrast stretching)

b. Filter Operation

It is important to remove unwanted regions as much as possible before any segmentation approach is applied as a fact that a better segmentation leads to a better identification of regions of interests. In this study median filter was applied and compared with the results of mean and Gaussian filters. Median filter showed better results in compare to other approaches as a fact that median filter is capable to preserve image edges [15-17]. After applying a number of sampling window (starting from [3x3] to [21x21]) this study found that [7×7] sampling window performs better results. Fig. 3 indicates the filtered result of this study test image.

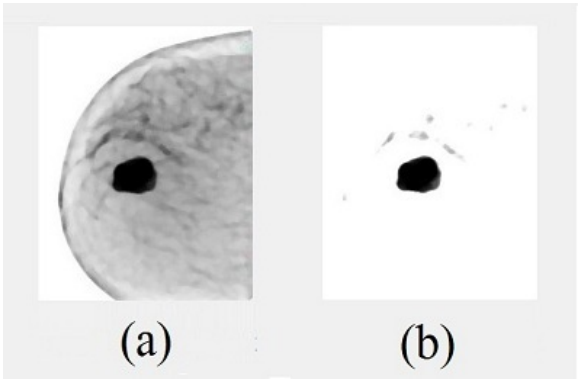


Fig 3: (a) Test image from step one (image enhancement), (b) resultant image using median filter

c. Segmentation

To extricate significant features from an image it is necessary to use segmentation approach. Threshold based gray-scale segmentation approach is considered as one of the most popular approaches. Among them Otsu's method used widely as it computes the threshold value a few times than other techniques [11].

In this study Otsu's threshold-based segmentation approach was incorporated. However; instead of considering all 3 color channels of RGB gray-scale image (R, G, B channel) this study investigated and discovered that only green and

blue channels are necessary to segment the breast cancer tissues more precisely. Therefore; for this study test images only green and blue channels were considered for segmentation using Otsu's method. Fig. 4 shows the result of the image after applying Otsu's threshold segmentation approach.

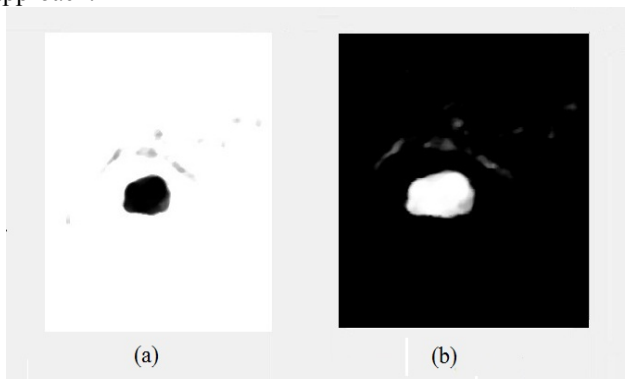


Fig 4: (a) Test image from step two (Median Filter), (b) Resultant image after segmentation

C. Post-processing

There are three parts of Post-processing. Such as:

a. Feature Extraction
b. Identification of ROI
c. Classification

a. Feature Extraction

To identify the region of interest shape, size and color features are required. As this research study used gray-scale images therefore; color features were ignored. Pathology laboratory experts' provided features were used in this research study which includes regions area, circularity, solidity, roundness, radius and diameter.

b. Identification of ROI

Using the features mentioned in the previous section identification is carried out to identify the region of interests.

c. Classification

After identification it was found that the accuracy rate was above 90% and the accuracy rate was accepted by the pathology laboratory experts. However; this research study tested with SVM classifier to check if it is possible to increase the accuracy rate. Accuracy rate did not increase and thus it is not necessary to use classifier at all as a fact that already the results have been accepted by the experts.

IV. Experimental Results

A. Result Analysis of the proposed method

In total 3 different datasets were used in this research study. Table 1 show that in batch 1 in total 130, in batch 2 in total 123 and in batch 3 in total 110 test images were analyzed.

.TABLE 1
Result Analysis of the proposed method

Dataset (No. of Images)	True Positive (TP)	True Negative (TN)	False Positive (FP)	False Negative (FN)	Accuracy
Batch1 (130)	122	1	4	3	93.85%
Batch2 (123)	112	2	4	5	91.06%
Batch3 (110)	104	1	2	3	94.55%
Total = 363	338	4	10	11	**Average = 93.11%**

Table 1 also indicates the accuracy rate for each batches used in this research study where for batch 1 the accuracy rate is 93.85%, for batch 2 the accuracy rate is 91.06% and for batch 3 the accuracy rate is 94.55%. According to pathology experts information, ultrasound image accuracy rate ±17% is acceptable. This proposed method was capable of obtaining an average accuracy rate of 93.11% which is acceptable in the pathology laboratory.

B. Experimental Outcomes

The experimental results of this proposed method is shown in Fig. 5. Three breast cancer test images with there results are shown in Fig 5.

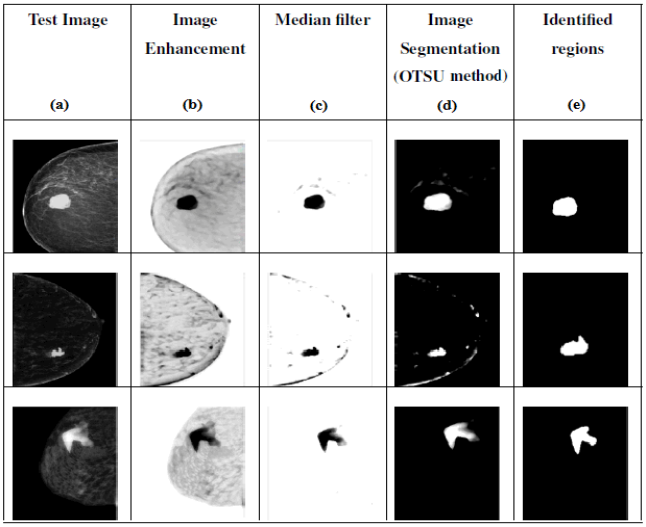


Fig. 5: (a) Test Image (b) Image Enhancement (c) Median Filter (d) Image Segmentation (OTSU Method) (e) Identified Regions

C. Comparative Results analysis

Using the images of an existing research study [14] images it was found that this study proposed approach was able to identify the breast cancer tissues accurately whereas; in [14] a number of unwanted regions were also identified which can confuse the experts' shown in Fig 6.



Fig. 6: Comparative Cancer detection using our proposed method and other existing approaches.

Although the accuracy rate is acceptable by the pathology laboratory experts' but is it still possible to increase the accuracy rate further if and only if the intra and inter-observation variability issues among pathology experts' can be minimized.

V. Conclusion

Manual detection of breast cancer tissue is more time consuming and the success rate normally depends on the person's skill. In this study, an automated approach has been proposed which is less complicated and takes less time with high precision. The aim of this research was to identify the breast cancer tissues correctly while maintaining the pathology experts' accuracy rate. Pathology laboratory experts already accepted the test results and the accuracy rate. However; in future this study will investigate further if it is possible to come up with another idea where it will be possible to increase the accuracy rate further.

Acknowledgment

This study delightedly acknowledges Dr. Sadequel Islam. Talukder (asst. Professor. Kishorgonj Medical College, Bangladesh) a pathology expert for providing the test datasets and necessary feature information for this study. This study also would like to acknowledge Green University of Bangladesh for relentless technical support.

References

[1] "Cancer - Cancer causes, Types, Treatment", medicinenet.com, 2018. [Online]. Available: https://www.medicinenet.com/cancer/article.htm.

[2] "Breast cancer: prevention and control - WHO— Breast cancer: prevention and control", who.int, 2018. [Online]. Available: http://www.who.int/cancer/detection/breastcancer/en/index1.html.

[3] M. Murtaza, A.Nornazirah, FA.Salih, EL.Illzam, M.Suleiman and AM.Sharifa, "Breast Cancer: Detection Markers, Prognosis, and Prevention", IOSR Journal of Dental and Medical Sciences, vol. 15, no. 8, pp. 73-80, 2016.

[4] "5 signs of breast cancer that arent a lump— CTCA - Cancer center", cancercenter.com, 2018. [Online]. Available: https://www.cancercenter.com/discussions/blog/5-signs-of-breast-cancer-that-aren't-a-lump/.

[5] M. Xian, Y. Zhang, H.D. Cheng, F. Xu, Z. Boyu and J. Ding, "Automatic Breast Ultrasound Image Segmentation", Sci. Diect, vol. 79, pp. 340-355, 2018.

[6] S. G. Marta, O. Weronika, W. Agata, J. Jan and C. Joanna, "Breast Cancer Segmentation Method in Ultrasound Images", Springer International Publishing AG, pp.23-31. 2017.

[7] D. Selvathi and A. AarthyPoornila, "Performance Analysis of Various Classifiers on Deep Learning Network for Breast Cancer Detection", International Conference on Signal Processing and Communication, pp. 359-363, 2017.

[8] C. Zhantao, D. Lixin, Y. Guowu, Y. Ting, C. Qin, F. Huazhu and X. Yanwu, "Breast Tumor Detection in Ultrasound Images Using Deep Learning", Springer International Publishing, pp. 21-128, 2017.

[9] A. N. Hanung, T. Yuli, R. Made and A. Igi, "Breast Ultrasound Image Segmentation Based on Neutrosophic Set and Watershed Method for Classifying Margin Characteristics", 2017

[10] B. Yash, V. Kesaari, P. Lipismita and P. S. Satya, "Automated Boundary Detection of Breast Cancer in Ultrasound Images Using Watershed Algorithm", Springer Nature Singapore Pte Ltd, pp. 729-738, 2018.

[11] T. Goh, S. Basah, H. Yazid, M. Aziz Safar and F. Ahmad Saad, "Performance analysis of image thresholding: Otsu technique", Measurement, vol.114, pp.298- 307, 2018.

[12] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, "Digital Image Processing Using MATLAB.", 2nd Ed. Hardcover–2009.

[13] Eric Olson, "Particle Shape Factors and Their Use In Image Analysis", Journal of GXP Compliance, vol.15, Issue 4, 2011.

[14] K. A. Singh and B. Gupta, "A Novel Approach for Breast Cancer Detection and Segmentation", ScienceDirect, vol.54, pp.676–682, 2015.

[15] Sazzad, T. M. Shahriar et al. "An Automated Detection Process to Detect Ovarian Tissues Using Type P63 Digitized Color Images." 2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI) (2015): 278-285.

[16] Sazzad, T. M. Shahriar et al. "An automated approach to detect human ovarian tissues using type P63 counter stained histopathology digitized color images." 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI) (2016): 25-28.

[17] Sazzad, T. S., Armstrong, L. J., & Tripathy, A. K. (2016, July). Type P63 non-counter stained digitized color images performs better identification than other stains for ovarian tissue analysis. In 2016 20th International Conference Information Visualisation (IV) (pp. 361-366). IEEE.

[18] Sazzad, T. S., Islam, S., Mamun, M. M. R. K., & Hasan, M. Z. (2013). Establishment of an efficient color model from existing models for better gamma encoding in image processing. International Journal of Image Processing (IJIP), *7*(1), 90.

An Automated Dermatological Images Segmentation Based on a New Hybrid Intelligent ACO-GA Algorithm and Diseases Identification Using TSVM Classifier

Md. Humayan Ahmed
Department of Computer Science and Engineering (CSE)
Khulna University of Engineering and Technology (KUET)
Khulna, Bangladesh
humayan036@gmail.com

Romana Rahman Ema
Department of Computer Science and Engineering (CSE))
Jashore University of Science and Technology
Jashore, Bangladesh
romanacsejstu@gmail.com

Tajul Islam
Department of Computer Science and Engineering (CSE))
Khulna University of Engineering and Technology (KUET)
Khulna, Bangladesh
tajulkuet09@gmail.com

***Abstract*—Dermatological diseases are the common and most extensive diseases worldwide. Diagnosis of dermatological diseases as well as skin cancer diseases is extremely difficult during its early to mid-stages by visual examination. Current researches propose efficient approach which can identify maximum 3- 9 types of skin diseases. It is necessary to develop an automated system which can identify more than 9 types of skin diseases. In this paper, we propose a new automated system for various kinds of dermatological diseases detection. In this automated system, a new hybrid intelligent ACO-GA algorithm including Ant Colony Optimization Algorithm (ACO), Genetic Algorithm (GA), and tabu list is used for different types of skin lesion segmentation of dermatological images and Transductive Support Vector Machine (TSVM) is used for dermatological diseases identification. Ant Colony Optimization Algorithm and Genetic Algorithm work simultaneously to search most optimal cluster centers in the problem space. After image segmentation, pixels are divided into different classes. In testing phase, the testing samples are compared with training samples using TSVM classifier and features extraction which can give the accurate disease detection result. The result of the proposed system successfully detects 24 different types of dermatological diseases with an accuracy rate of 95%.**

Keywords—digital image processing, genetic algorithm, ant colony optimization algorithm, mathlab code, transductive support vector machine.

I. INTRODUCTION

Dermatology is the branch of medical science dealing with the skin, hair, nails with their diseases and it is one of the most unpredictable and difficult terrains to diagnose due to its complexity. Nowadays, people are suffering from different types of skin diseases such as scabies, psoriasis, leprosy, acne, tinea corporis, rosea, melanoma skin cancer, benign skin cancer, kaposi sarcoma, seborrheic keratosis, squamous cell, basal cell etc. Among them, more than 125 million people are suffering from psoriasis (Chronic skin disorder). Skin cancer for example melanoma skin cancer, basal and squamous cell skin cancer, lymphoma of the skin etc. rate is increasing over the few decades. Melanoma is most diversifying skin cancer and 75% of deaths associated with melanoma skin cancer [1]. Every year, in globally between 2 and 3 million people are suffering from non-melanoma skin cancer and 132,000 people are suffering from melanoma skin cancer in accordance with World Health Organization (WHO) [2]. Researchers have been attracted to diagnosis this type of problem. Dermatophytosis is high at rural area especially in developing countries like Bangladesh. In the developing countries like Bangladesh, it is expensive for a large number of people to consult a dermatologist for their skin diseases. Every year a large number of populations in the developing countries like Bangladesh suffer due to different types of skin diseases and it is a great challenge for us to overcome this problem. For this reason, we propose an automated skin disease detection system for the patients as well as dermatologists where a new hybrid intelligent ACO-GA algorithm is used for automatic colored skin image segmentation and TSVM classifier is used for automatic skin disease detection. In the hybrid algorithm, an intelligent ant colony optimization algorithm and evolutionary genetic algorithm is used for searching the most optimal cluster centers by segmenting the image pixels into different cluster centers using updated pheromone values and tabu list. After that, we use color co-occurrence method for features extraction of the segmented image. After features extraction, transductive support vector machine is used as a classifier to detect the skin diseases automatically.

II. LITERATURE REVIEW

Nowadays, medical science and its field largely depend on computer aided diagnosis. It is more complex to detect skin disease at early stage. Different types of skin diseases detection from color image is a very challenging task in computer vision. Digital image processing can play an important role for skin disease diagnosis and it is possible without any physical contact with skin. Not only image processing but also machine learning plays an essential role in the medical field for the automation as well as to predict what type of diseases present on the skin. Researchers have been attracted by this technology. That is why researchers have been proposed various kinds of skin diseases segmentation and classification systems. Wei, Li-sheng,

Quan Gan and Tao Ji [3] have proposed skin disease recognition method based on image color and texture features where SVM is used as a classifier for finding the diseases. Sumithra, R., Mahamad Suhil and D. S. Guru [4] have proposed a segmentation and classification method of skin lesions for disease diagnosis. They used SVM and K-NN classifier for the classification using the extracted features. Yasir, Rahat, Md. Ashiqur Rahman and Nova Ahmed [5] have developed different types of image processing algorithms for feature extraction. For diseases identification, they used feed forward artificial neural network in training and testing phase. Shoieb, Doaa A., Sherin M. Youssef and Walid M. Aly [6] have implemented a computer aided model for skin diagnosis using deep learning. In their research, they use convolution neural network for feature extraction and multiclass linear support vector machine (SVM) is used as a classifier to detect the diseases. Kumar, Vinayshekhar Bannihatti, Sujay S. Kumar and Varun Saboo [7] have proposed a dermatological disease detection system using image processing and machine learning. In their experiment, they used different types of machine learning algorithms such as ANN, Decision tree and K-NN for skin disease detection. Aswin, R. B., J. Abdul Jaleel and Sibi Salim [8] have introduced a hybrid genetic algorithm- artificial neural network classifier for skin cancer detection in where, ANN classifier is used for classification and it is optimized by genetic algorithm to improve the accuracy. L. G. Kabari and F. S. Bakpo [9] have made an artificial neural networks system for predicting the skin diseases and their accuracy rate is 90%. M. Shamsul Arifin, M. Golam Kibria, Adnan Firoze, M. Ashraful Amin and Hong Yan [10] have proposed an automated skin disease diagnostic system. In their research, they used different pre-processing algorithms and feed forward back propagation artificial neural network is used for diseases identification. F. Tushabe, E. Mwebaze and F. Kiwanuka [11] have proposed a classification method to classify the image as a bacterial or viral skin infection using image processing techniques. Nidhal K. Al Abbadi, Nizar Saadi Dahir, Mushsin A. Al-Dhalimi and Hind Restom [12] have developed a psoriasis diseases detection system using feed forward artificial neural networks. M. Ganesh Kumar and J. J. B. Vasanthi [13] have implemented the melanoma disease detection technique by using preprocessing and edge detection. Md. Khaled Abu Mahmoud, Adel Al-Jumaily and Maen Takruri [14] have designed an automatic skin cancer classification system using neural network classifier based on wavelet and curvelet features.

From the literature, we can see that most of the researchers consider only 3-9 types of skin diseases in their research work. The remaining skin diseases have the equal importance in the medical field. We get another motivation from this and so, we consider 24 types of skin diseases including the previous 3-9 types of skin diseases.

III. Limitation of Existing Work

The limitation of the existing works are: 1) consider only 3-9 types of skin diseases, 2) Lack of huge amount of training samples in database, 3) Computational complexity is very high in some cases, 4) low segmentation accuracy in color image 5) low detection accuracy in accordance with the limitation of proper classification system. For this reason, we develop a new hybrid intelligent ACO-GA algorithm for colored skin image segmentation and disease detection using TSVM classifier so that we can achieve more optimization, low computational complexity and high segmentation and detection accuracy.

IV. Proposed Methodology

We know real time problem is optimization problem. In our proposed system, optimization problem is developed. A new hybrid ACO-GA algorithm is used in this system because of its much more intelligent background with easier implementation and very low computational complexity.

Our implemented hybrid intelligent algorithm for dermatological colored image segmentation and type of skin disease identification using TSVM classifier that consists of some steps are given below:

A. Image Acquisition

The first step of our automated color image analysis system is image acquisition. In this step, a good quality of skin image is captured using a digital camera and it is then transferred into image preprocessing step for color transformation.

B. Image Preprocessing

Noisy image may lead algorithm towards incorrect result and it is necessary to have quality image without any noise to get accurate result. There are a variety of image preprocessing techniques such as median filter, Gaussian filter, contrast enhancement technique, region of interest technique for removing the noise from the image. We remove the noise using median filter and contrast enhancement technique. By using median filter and contrast enhancement technique, the image will be very smooth and sharp.

C. Hybrid ACO-GA Algorithm for Clustering in Image Segmentation Phase

The intelligent ACO algorithm is used for finding the most optimal cluster centers from the problem space. In ACO phase, swarm of ants is generated randomly and all ants are walking through the image pixels. Because, image consists of pixels and all the ants move from one pixel to another by accumulating pheromone value. However, they create cluster by their intelligent behavior. Every ant on the pixels calculates its transition probability and moves to the chosen pixel by the equation (1).

$$P_{mn}(t+1) = \tau_{mn}{}^{\alpha}(t)\ \eta_{mn}{}^{\beta}(t) / \sum_{i \epsilon\, s} \tau_{mi}{}^{\alpha}(t) \eta_{mi}{}^{\beta}(t) \quad (1)$$

Where, $\Delta\tau_{mn}^{k}$ denotes the pheromone value and η_{ij} denotes the heuristic information. In each iteration, every ant chooses one pixel of its adjacent eight pixels as its next shift and records the pixels into the tabu list. However, all the ants move from one pixel to another and record the pixels into the tabu list. It will be continued until total move is finished. After finishing all ants' path construction in each iteration, we get the best solution and record the best solutions' pixels into tabu list. After that, pheromone is updated for obtaining best solution by the equation (2).

$$\tau_{mn}(t+1) = (1-\rho)\tau_{mn}(t) + \sum_{k=1}^{k=N} \Delta\tau_{mn}^{k} \quad (2)$$

The flowchart of hybrid intelligent ACO-GA algorithm is shown in Fig. 1.

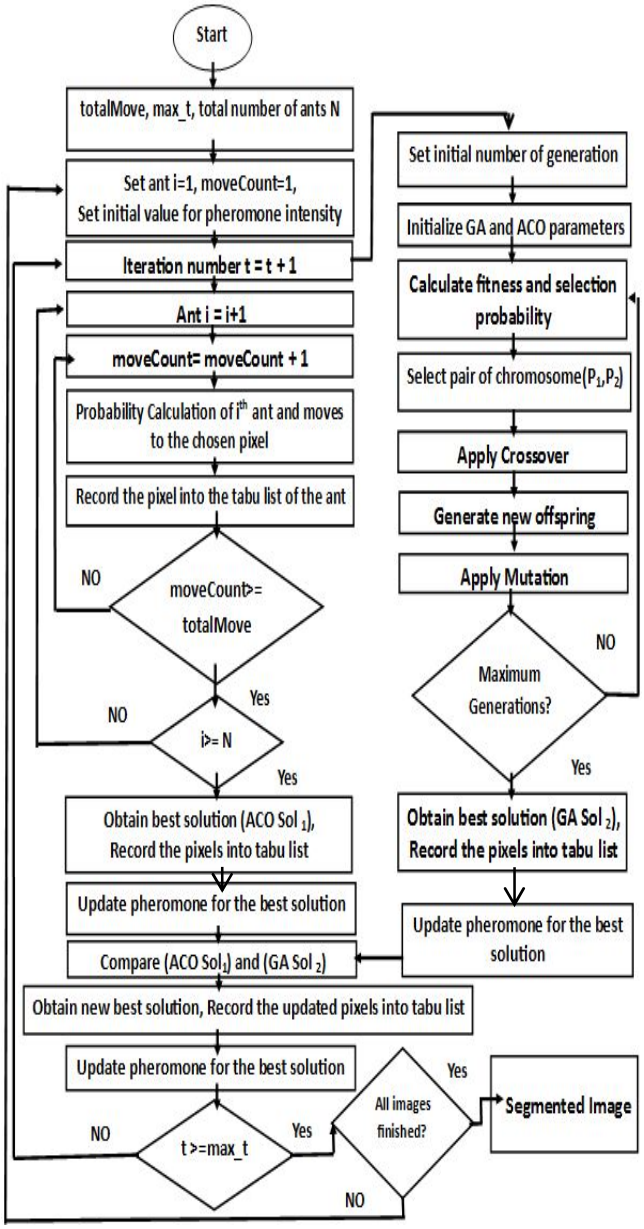


Fig. 1. Flowchart of hybrid ACO-GA algorithm for color image segmentation.

The evolutionary genetic algorithm is also used simultaneously with ACO algorithm for finding the most optimal cluster centers from the problem space. It is possible by measuring the distance using Euclidean distance. In this phase, a random initialized population is generated of any desired size at first. It is mandatory to initialize GA and ACO parameters in this phase. At each iteration, randomly produced initial population is generated with ant colony optimization in parallel manner. In each iteration of GA phase, Chromosome is generated randomly with m cluster centers where, chromosome is represented as string. Pixels are clustered by chromosome encoding scheme. The pixels are closest to the cluster center will be kept in the same cluster center. For searching most optimal cluster center, we need the equation of Euclidean distance d_i is given below:

$$d_i = \sum_{i=1}^{m} \sum_{(for\ each\ pixel\ p_j \epsilon\ cc_i)} \left|\left|p_j - cc_i\right|\right|$$

From the Euclidean distance, we get the fitness function as follows:

$$F = \frac{1}{d_i} = \frac{1}{\sum_{i=r}^{m} \sum_{(for\ each\ pixel\ p_j \epsilon\ cc_i)} |(p_j - cc_i)|}$$

After calculating fitness function, selection probability is used to select the fittest individuals from the population. The selection probability equation is as follows:

$$P_C = \frac{fitness\ of\ individual\ chromosome\ \ or\ cluster}{Total\ fitness\ of\ the\ current\ chromosomes\ or\ clusters}$$

In this GA phase, we use 2-point crossover for producing new offspring from two chromosomes and shift mutation [15] is used for creating genetic diversity.
New offspring is continuously generated until maximum generation is reached. However, we get the best solution (image pixels) from genetic algorithm and record the pixels into tabu list for storing the pixels into memory. For obtaining best solution, the pheromone value of the adjacent pixels is updated using ACO parameter in GA phase.

Both sided clustering is performed simultaneously. The obtaining best solution from GA phase is compared with the obtaining best solution from ACO phase. However, we obtain the new best solution and store the pixels into tabu list. After storing the pixels into tabu list, pheromone is updated for obtaining optimal solution. All the process of ACO and GA phase is continued until reached maximum iteration. Therefore, we get the segmented image. After image segmentation, TSVM is used for disease identification using segmented image features extraction.

D. Feature extraction

Feature extraction is a necessary step to get accurate result in biomedical image processing. The infected region of skin is determined in feature extraction phase. Instead of choosing the whole set of pixels, we can choose only which are necessary and sufficient to describe the whole of the segmented image. Using image enhancement and region of interest (ROI) technique, we can find out the necessary infected region. The frequency of occurrence of certain value in the whole of the image is obtained by histogram calculation. So, it is an important feature for frequency analysis. Afterward frequency analysis is taken by co-occurrence which is transmitted into next level. In the next level, the intensity occurrence of two pixels together is stored in the matrix. However, we get color co-occurrence matrix. The features such as Contrast, Energy, Correlation, Homogeneity and Entropy are extracted from color co-occurrence matrix data. The formulae of the features are given below:

Contrast: $\sum_{i=0}^{i=k-1} \sum_{j=0}^{j=k-1} |\mathrm{i} - \mathrm{j}|\, 2C(i,j)$

Energy: $\sum_{i=0}^{i=k-1} \sum_{j=0}^{j=k-1} C(i,j)\, 2$

Correlation: $\sum_{i,j} \frac{(i-\mu)(j-\mu)C(i,j)}{\phi_i \phi_j}$

Homogeneity: $\sum_{i=0}^{i=k-1} \sum_{j=0}^{j=k-1} C(i,j)\,/(1 + |i-j|)$

Entropy: $-\sum_{i=0}^{i=k-1} \sum_{j=0}^{j=k-1} C(i.j) \log C(i.j)$

E. Dermatological Diseases Identification

Transductive support vector machine is used not only for finding the hyper plane but also finding the labeled samples/data from unlabeled samples/data. We use TSVM classifier for detecting which type of disease present on the skin using features extraction. The TSVM objective function is as follows:

$$min_{y^*,Z,b,\upsilon,\eta^*} 0.5||Z|| + F\sum_{i=1}^{i=n} \upsilon_i + F^*\sum_{j=1}^{j=m} \eta_j{}^*$$

Algorithm 1 shows our TSVM training process.

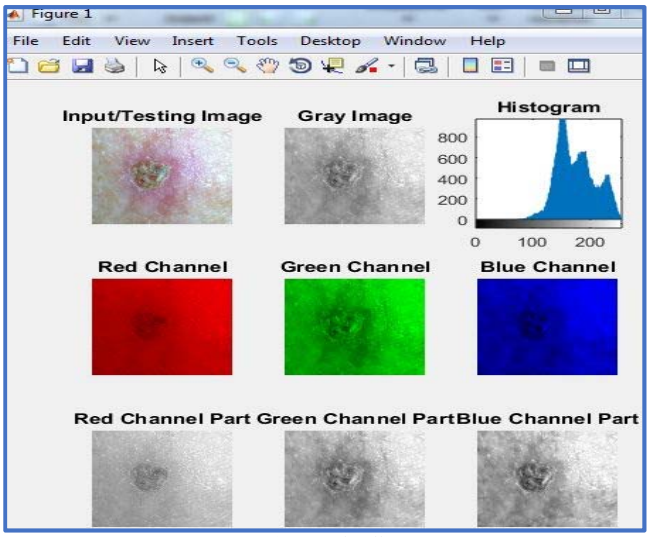
Algorithm 1: TSVM training process algorithm

1. Initialize parameter T and T^*
2. Apply inductive SVM on training data
3. Set positive label samples number
4. Find label samples for all unlabeled samples using decision boundary function value in where label samples are positive for the highest value of decision boundary function
5. Set T^*_{tmp}
6. Retrain SVM over the whole samples of labeled and unlabeled samples
7. Shift unlabeled samples to labeled samples and make the value of objective function
8. If all the data satisfy the shifting condition then $T^*_{tmp} = T^*_{tmp} + 1$, else go to step 7
9. If $T^*_{tmp} > T^*$ then obtain the output data, else go to step 6
10. Stop

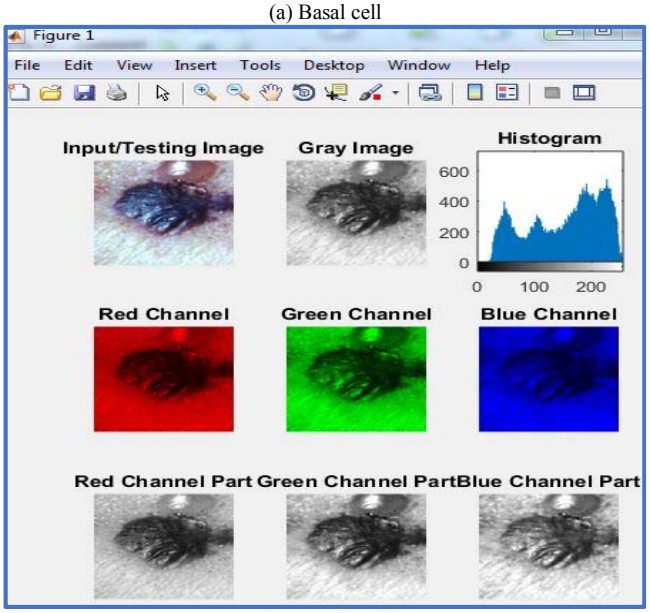
In testing phase, the features of the input image are calculated and the obtained data is then compared with TSVM training data. However, skin diseases are classified and detected using TSVM classifier.

V. Results and Discussion

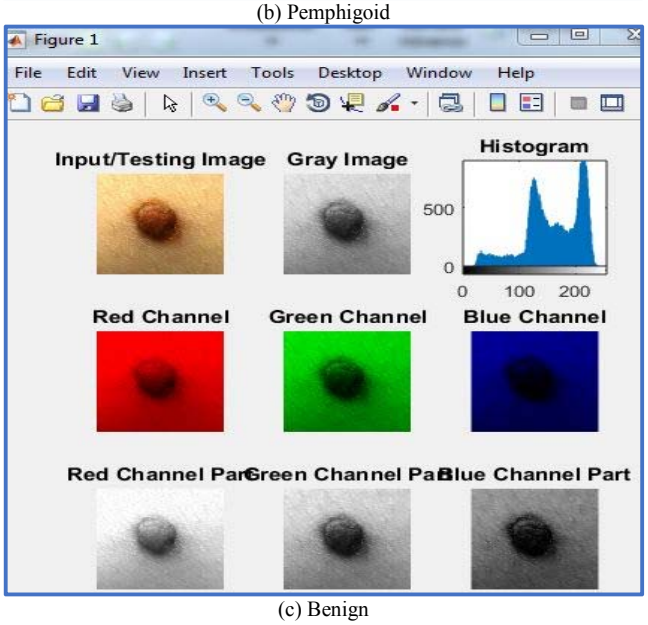
We implement our proposed hybrid intelligent ACO-GA algorithm using MATLAB program which can segment dermatological images accurately. We also implement transductive support vector machine algorithm using MATLAB which can detect the skin diseases accurately. In this experiment, a total of 812 colored skin images from Khulna Medical College, Department of Dermatology are collected out of 24 different types of skin diseases such as scabies, psoriasis, seborrheic dermatitis, tinea, corporis, chronic dermatitis, eczema, foot ulcer, acne, rosea, vitiligo, leprosy, lichen planus, bullae, pityriasis rosea, pityriasis rubra pilaris, shingles, melanoma skin cancer, benign skin cancer, kaposi sarcoma, seborrheic keratosis, squamous cell, basal cell and merkel cell skin cancer. Above diseased skin images are used as input dataset of the algorithm. Input query image is taken into different color channel (RGB). Gray scale image is also computed as well as its histogram information. We have also shown the segmentation and detection of two new skin diseases such as Basal cell and pemphigoid skin diseases which is not researched yet by the researchers. The output of color transformation system of input image is shown in Fig. 2.



(a) Basal cell



(b) Pemphigoid

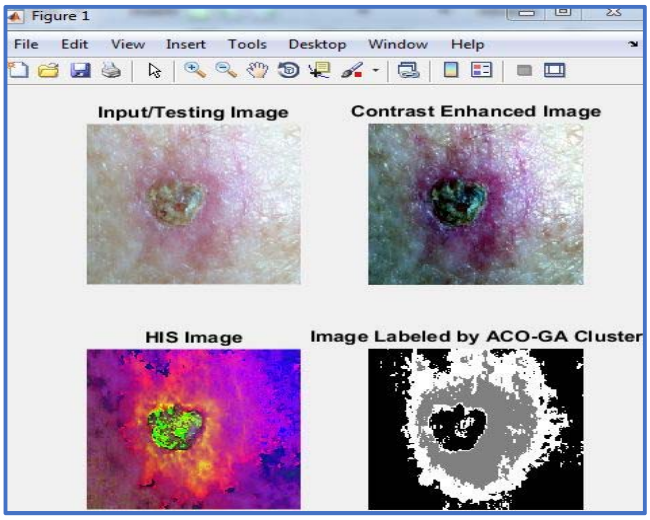


(c) Benign

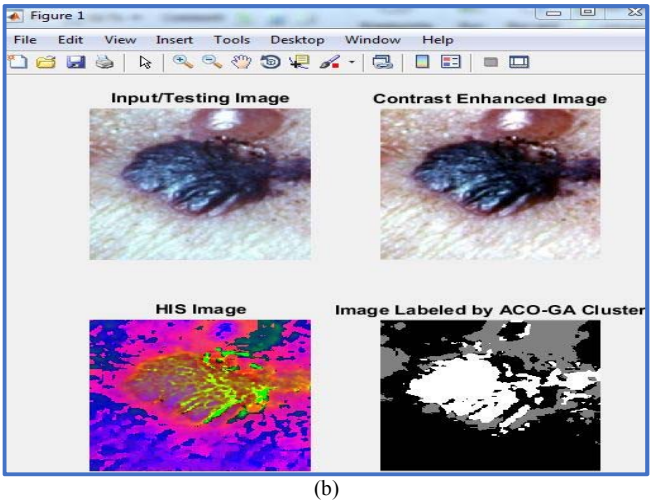
Fig. 2. Color transformation system of different infected skin images with different diseases.

Noise of the input infected skin image is removed by contrast enhanced technique and Median filter. The input

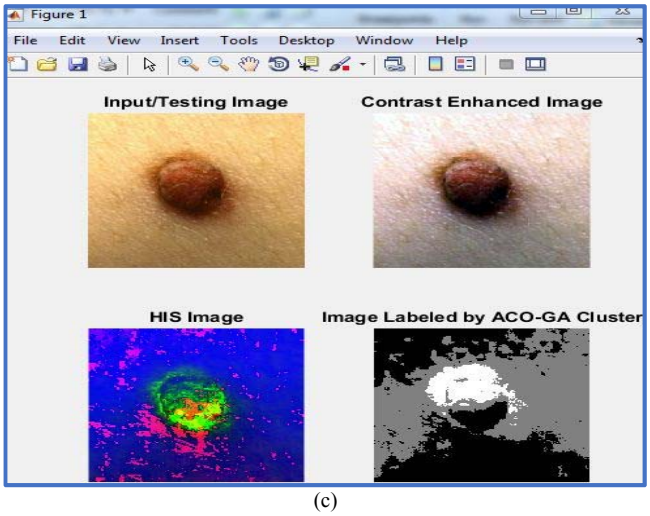
image is converted into HIS and the disease affected area is sharply appeared in Fig. 3. Afterwards, image is labeled and segmented by hybrid ACO-GA clustering.



(a)



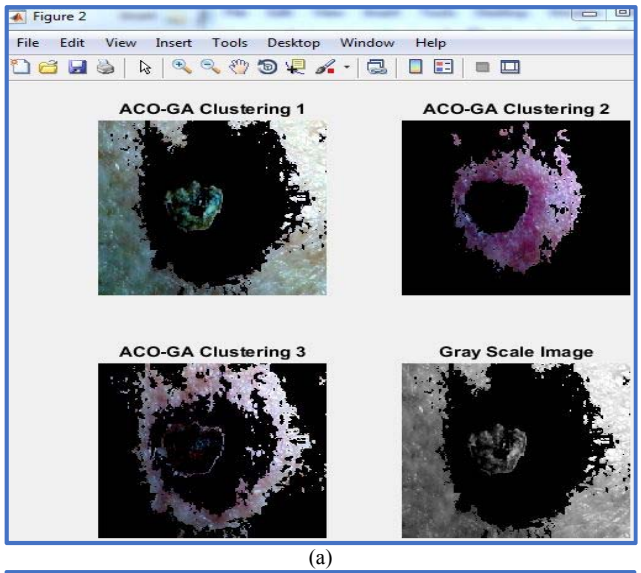
(b)



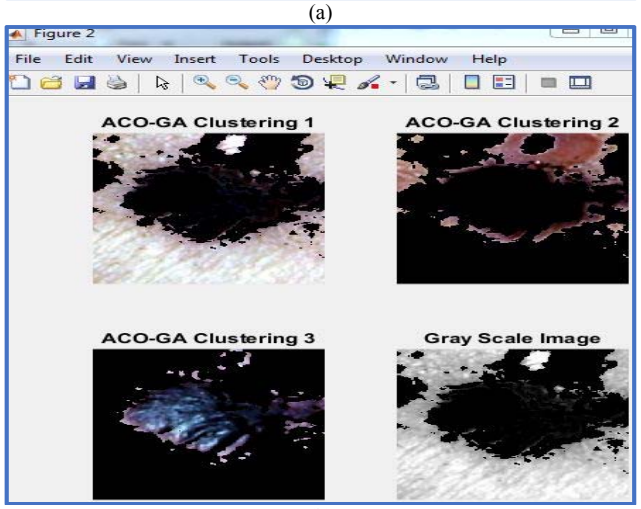
(c)

Fig. 3. Input infected skin image conversion to contrast enhanced image, HIS image and labeled image by ACO-GA clustering.

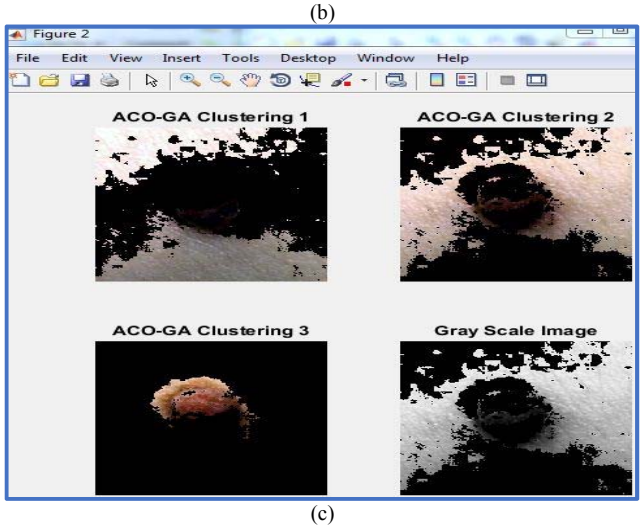
The hybrid ACO-GA clustering algorithm segments the input infected skin image by clustering process and gives us optimal solution. In this process, we use three clusters in our implemented hybrid ACO-GA algorithm. In ACO-GA clustering system, the total number of pixels from the segmented diseased area as well as skin area is detected. Fig. 4 shows the segmented output image performed by hybrid intelligent ACO-GA algorithm.



(a)



(b)



(c)

Fig. 4. Input infected skin image segmentation by hybrid ACO-GA clustering.

In training phase, The TSVM training process is performed for all the images to obtain the high detection

accuracy. The segmented image is entered into testing phase. After calculating segmented image features, the data is then compared with TSVM training data. However, the segmented diseased image is classified into different disease using TSVM classifier. Fig. 5 shows the dermatological skin diseases detection using TSVM classifier.



(a) Benign skin cancer identification



(b) Impetigo skin disease identification

Fig. 5. Dermatological skin diseases identification using TSVM classifier.

Two papers in this domain [8], [5] propose hybrid GA-ANN and Feed Forward BP-ANN respectively for identification of different types of skin diseases. The detection accuracy of their system was 88% and 90% respectively. They research with only 5-9 types of skin disease that is very inadequate. On the other hand, our proposed system of using hybrid ACO-GA algorithm with TSVM classifier has given very promising results in identification of different types of skin diseases with accuracies of up to 95%. The accuracy comparison result is shown in Fig. 6.

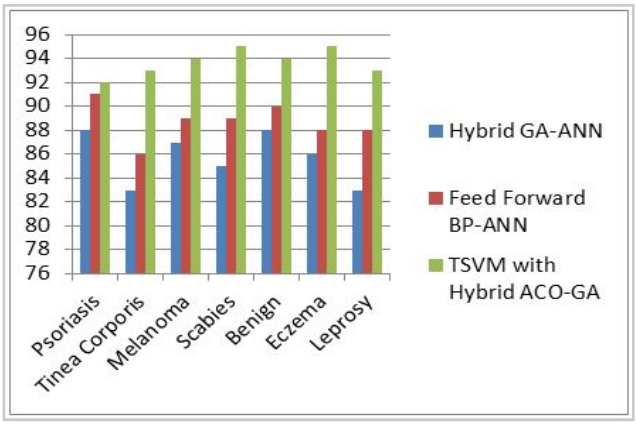


Fig. 6. Accuracy comparison of Hybrid GA-ANN, Feed Forward BP-ANN and our proposed TSVM with hybrid ACO-GA algorithm.

VI. Conclusion

The implementation of our proposed system segments the colored skin image using hybrid ACO-GA algorithm and also detects not only dermatological skin disease but also skin cancer using TSVM classifier. In this process, hybrid ACO-GA algorithm is used for finding the most optimal cluster centers from the problem space. This hybrid ACO-GA algorithm then supports TSVM to find the accurate skin diseases. In the comparison of our work with related works in this domain, there is no other research and solution to identify the skin diseases which can handle 24 different types of skin diseases. The advantage of our proposed system is to give high performance with a very low computational complexity, to achieve segmentation accuracy of about 94% and skin disease detection accuracy of about 95%. We suggest some recommendations for our future work: collect large amount of dermatological images for disease identification, give prescribed medicine information for treatment to the rural people in accordance with different types of skin diseases and also support dermatologists for skin diseases diagnosis.

References

[1] A. F. Jerant J. T. Johnson, C. D. Sheridan, and T. J. Caffrey, "Early detection and treatment of skin cancer," in American Family Physician, vol. 62, no. 2, pp. 1-6, July 2000.

[2] World Health Organization. [Online]. Available: http://www.who.int/en./

[3] Wei, Li-sheng, Quan Gan and Tao Ji., "Skin Disease recognition method based on image color and texture features," in Computational and mathematical methods in medicine, 2018.

[4] Sumithra, R., Mahamad Suhil and D. S. Guru, "Segmentation and classification of skin lesions for disease diagnosis," in Procedia Computer Science 45 (2015): 76-85.

[5] Yasir, Rahat, Md. Ashiqur Rahman and Nova Ahmed, "Dermatological disease detection using image processing and artificial neural network," in 8th International Conference on Electrical and Computer Engineering, IEEE, 2014.

[6] Shoieb, Doaa A., Sherin M. Youssef and Walid M. Aly, "Computer-aided model for skin diagnosis using deep learning," in Journal of Image and Graphics 4.2 (2016): 122-129.

[7] Kumar, Vinayshekhar Bannihatti, Sujay S. Kumar and Varun Saboo, "Dermatological disease detection using image processing and machine learning," in 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR), IEEE, 2016.

[8] Aswin, R. B., J. Abdul Jaleel and Sibi Salim, "Hybrid genetic algorithm-artificial neural network classifier for skin cancer detection," in 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), IEEE, 2014.

[9] L. G. Kabari and F. S. Bakpo, Member, IEEE, "Diagnosing Skin Diseases Using an Artificial Neural Network," c_2009 IEEE.

[10] M. Shamsul Arifin, M. Golam Kibria, Adnan Firoze, M. Ashraful Amin and Hong Yan, "Dermatological Disease Diagnosis Using Colour-skin Images," Proceedings of the 2012 International Conference on Machine Learning and Cybernetics, Xian, 15-17 July, 2012.

[11] F. Tushabe, E. Mwebaze and F. Kiwanuka, "An image based diagnosis of virus and bacterial skin infections," in The International Conference on Complications in Interventional Radiology, 2011.

[12] Nidhal K. Al Abbadi, Nizar Saadi Dahir, Mushsin A. Al-Dhalimi and Hind Restom, "Psoriasis Detection Using Skin Colour and Texture Features," in Journal of Computer Science 6(6): 626-630, 2010, ISSN 1549-3636, 2010 Science Publications.

[13] M. Ganesh Kumar and J. J. B. Vasanthi, "Skin disease identification using image segmentation," in International Journal of Innovative Research in Computer and Communication Engineering, vol. 5, no. 1, pp. 154-160, 2017.

[14] Md. Khaled Abu Mahmoud, Adel Al-Jumaily and Maen Takruri, "Wavelet and Curvelet Analysis for Automatic Identification of Melanoma Based on Neural Network Classification," in International Journal of Computer Information Systems and Industrial Management (IJCISIM), ISSN 2150-7988, vol. 5 (2013), pp. 606-614.

[15] Md. Humayan Ahmed, Romana Rahman Ema, Tajul Islam, Rana M Pir, "Application of an Efficient Genetic Algorithm for Solving n*m Flow Shop Scheduling Problem Comparing it with Branch and Bound Algorithm and Tabu Search Algorithm," in American Scientific Research Journal for Engineering, Technology and Sciences, Vol. 41, Issue 1, pp. 12-25, 2018.

Design and Development of a Hand Gesture Based Virtual Mouse

Kabid Hassan Shibly
Dept. of Computer Science & Engineering
Dhaka International University
Dhaka, Bangladesh
khshibly00@gmail.com

Samrat Kumar Dey
Dept. of Computer Science & Engineering
Dhaka International University
Dhaka, Bangladesh
samratcsepstu@gmail.com

Md. Aminul Islam
Dept. of Computer Science & Engineering
Dhaka International University
Dhaka, Bangladesh
aminuloet@gmail.com

Shahriar Iftekhar Showrav
Dept. of Computer Science & Engineering
Dhaka International University
Dhaka, Bangladesh
shahriarshowrav@gmail.com

***Abstract*— Method of establishing a process of interaction between human and the computer is evolving since the invention of computer technology. Though Wireless and Bluetooth mouse technology are invented still, that technology is not completely device free. A Bluetooth mouse has the requirement of battery power and connecting dongle. Presence of extra devices in a mouse increases the difficulty to use it. Our proposed virtual mouse model is beyond this limitation. This paper proposes a virtual mouse system based on HCI using computer vision and hand gestures. Gestures captured with a built-in camera or webcam and processed with colour segmentation and detection technique. The users are allowed to control some of the computer cursor functions with their hands that bear coloured caps on fingertips. Primarily, a user can perform left clicks, right clicks, and double clicks, scrolling up or down using their hand in different gestures. This system captures frames using a webcam or built-in cam and processes the frames to make them track-able and after that recognizes different gestures made by users and perform the mouse function. Therefore, the proposed mouse system eliminates device dependency in order to use a mouse. Therefore, it can be proved beneficial in order to develop HCI technology.**

Keywords—HCI (Human-Computer Interaction), HSV (Hue Saturation Value), Hand Gesture, Color Detection, Gesture Recognition.

I. INTRODUCTION

As the technologies are developing day by day the devices becoming compact in size. Some devices have gone wireless, some of them gone latent. This paper proposes a system that could make some the devices go latent in the future that is the future of HCI (Human-Computer Interaction). The proposal is to development of a Virtual Mouse using Gesture Recognition. The aim is to control mouse cursor functions using only a simple camera instead of a traditional or regular mouse device. The Virtual Mouse works as a medium of the user and the machine only using a camera. It helps the user to interact with a machine without any mechanical or physical devices and control mouse functions. In this gesture recognition system, it is very possible to capture & track the fingertip of hand with a webcam or built-in cam which is bearing a color cap or color sticky note paper and the system track the color and movement of the hand & move cursor with it.

Typically we use a mouse, keyboard or other interacting devices which is mainly compact with the computer machine. The wireless devices also need a power source and connecting technologies, but in this paper, the user's bare hand is the only input option using a webcam. Therefore, it is a very interactive way to control the mouse cursor.

This system is implemented in Python programming language using the Computer Vision based library OpenCV. This system has the potential to replace the typical mouse and the remote controller of machines. The only barrier is the lighting condition. That is why the system still cannot be enough to replace the traditional mouse as most of the computers are used in poor lighting conditions.

A. Problem Description & Overview

To track fingertips as a movable object, and to utilize it for mouse functions, the camera should be positioned in a way so that it can see the user's hands in the right positions. This can be used in space-saving situations, for those patients who don't have control over their limbs and for other similar cases. It's a virtual mouse instead of a physical mouse which will work only based on webcam captured frames & tracking colored fingertips.

B. Significance in Real World Application

Video conferencing is very popular nowadays. For this reason, most of the computer users use a webcam on their computer and most of the laptops have a built-in webcam. The proposed system which is webcam based, might be able to eliminate the need of a mouse partially. The process of interaction with a computer using hand gesture is a very interesting & effective approach to HCI (Human-Computer Interaction). There is some really good research on this interest. The hand gesture recognition technology is also popular in sign language recognition.

C. Objective

The objective is to develop and implement an alternative system to control a mouse cursor. The alternative method is hand gesture recognition using webcam and color detection method. The ultimate outcome of this paper is to develop a system which recognizes hand gesture and controls mouse cursor using color detection method of any computer.

II. RELAVENT WORK

Cursor control application using hand gesture is used in many ways, but most of the time, it requires of wearing a DataGlove. This reduces the efficiency of performance

between the user and the system. And system complexity is also an issue in this process.

There could be two possible gesture recognitions for HCI, one is hardware based & another one is computer vision based. One of the early hardware-based systems was proposed by Quam (1990) in which, the user had to wear a bulky DataGlove to use the system [1]. Though this method gives a high accuracy control, it is very difficult to use as some gesture is not meant for everyone and also very impractical for mass users in the everyday world.

Meanwhile, the vision-based hand gesture recognition is also two types: marker-based & non-marker based. Non-marker based recognition generally is not as accurate as the marker based recognition. And the maker based recognition has better accuracy compared to the other gesture recognition systems, though the user has to use a simple color cap on the fingertip. But this way is lighter and almost nothing in comparison with the DataGlove of the hardware-based system.

Gesture recognition might be a futuristic way to the computers to understand human gesture (Body Language). It will build a greater interaction between human & computer machines rather than primitive text-based interaction.

Most of the marker based gesture recognition mouse uses at least two color markers to track. Because of detecting multiple colors, the system gets slow and some laggings appear on the system during the time of performance.

In 2010, ChenChiung Hsieh and Dung-Hua Liou proposed a paper "A Real-Time Hand Gesture Recognition System Using Motion History Image" [2] based on adaptive skin color model & motion history image (MHI). In their work they used an adaptive skin color model and a motion history image-based hand moving direction detection method. The prime limitation of the paper is a problem with working for more complicated hand gestures recognition.

In 2011, Chang-Yi Kao and Chin-Shyurng Fahn published a paper "A Human-Machine Interaction Technique: Hand Gesture Recognition Based on Hidden Markov Models with Trajectory of Hand Motion" [3] that is basically learning based interaction between human & machine. Their work is very accurate but it worked only in high configuration computers.

In 2013, Angel, Neethu. P.S proposed a paper named "Real Time Static and Dynamic Hand Gesture Recognition" [4] in which design, develop and study a practical gesture recognition that can be used in a variety of human-computer interaction applications framework for real-time. But, it was unable to work at a complex background and was computable only under good light.

In 2013, Ashwini M. Patil, Sneha U. Dudhane, Monika B. Gandhi proposed a paper titled "Cursor Control System Using Hand Gesture Recognition" [5] where they developed a machine-user interface which implements hand gesture recognition using simple computer vision and multimedia techniques. But a major limitation is before working with gesture comparison algorithms, skin pixel detection and hand segmentation from stored frames needs to be done.

In 2014, Abhik Banerjee & Abhirup Ghosh proposed "Mouse Control using a Web Camera based on Color Detection" [6] titled paper where the methodology is Hand gestures were acquired using a camera based on color detection technique. The limitations of their work are the operating background has to be light and no bright colored objects are present. It works well on certain computers of high configuration.

In 2016, Yimin Zhou, Guolai Jiang & Yaorong Lin published "A novel finger and hand pose estimation technique for real-time hand gesture recognition" [7] based on directly extract fingers from salient hand edges. Considering the hand geometrical characteristics, the hand posture is segmented and described based on the finger positions, palm center location and wrist position. But this method is only compatible with high configuration computer machines.

In 2016, Pooja Kumari, Saurabh Singh & Vinay Kr. Pasi developed "Cursor Control using Hand Gestures" [8] is based on multiple color bands where different colors perform different actions. The number of colors is the key to perform mouse functions. But, multiple colors were used to control the system. It is relied on the number of colors to perform a function instead of different gestures.

In 2017, Aashni Haria, Archanasri Subramanian, Nivedhitha Asokkumar, Shristi Poddar, Jyothi S Nayak developed "Hand Gesture Recognition for Human-Computer Interaction" [9] based on background extraction and contours detection system. But it is very slow to work with.

In 2018, Abhilash SS, Lisho Thomas, Naveen Wilson, Chaithanya C published a paper on "Virtual Mouse Using Hand Gesture" [10] which designed to work with the color detection system, functions works on the number of colors detected. But it can perform only a few mouse actions, and does not work without static background.

III. METHODOLOGY

The methods used in each & every part of the system proposing in this paper are explained separately.

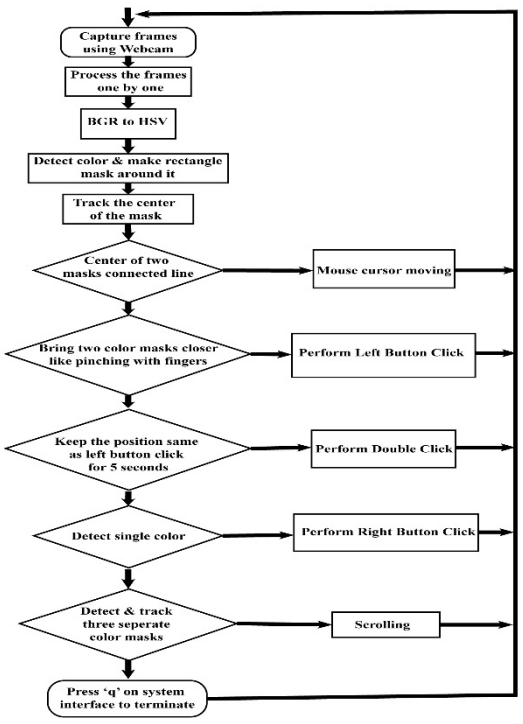


Fig. 1. Flowchart of The Methods of Gesture Based Mouse

A. Camera

The system works on the frames captured by the webcam on the computer machine or built-in camera on a laptop. By creating the video capture object, the system will capture video using webcam in real-time. In order to use a single camera with this system the device index would be "0". To add additional camera device index would be 1, 2 and so on. This camera will capture frame by frame and will pass it to the system.

B. Capturing

By using an infinite loop the webcam captures each and every frame till the program termination. The frames from the real-time video are processed from BGR to HSV color space.

C. Color Detection & Masking

In this proposed system, color detection is done by detecting color pixels of fingertips with color caps from the frames that were captured by the webcam. This is the initial and fundamental step of the proposed system. The outcome of this step will be a grayscale image, where the intensity of the pixels differs the color cap from the rest of the frame and the color cap area will be highlighted. Then rectangle bounding boxes (masks) will be created around the color cap and the color cap will be tracked. The gesture will be detected from the tracking of these color caps.

IV. Gesture Recognition

A. Mouse Movements

At first, calculation of the center of two detected color objects which is done by the coordinates of the center of the detected rectangle. To create a line between two coordinates, the built-in OpenCV function is used and to detect midpoint equation given below is used:

$$M = \left(\frac{Xa+Xb}{2}, \frac{Ya+Yb}{2}\right) \quad (1)$$

This midpoint is the tracker for the mouse pointer and the mouse pointer will track this midpoint. In this system, the coordinates from camera captured frames resolution is converted to screen resolution. A predefined location for the mouse is set, so that when the mouse pointer reaches that position, the mouse started to work and this may be called open gesture. This allows the user to control the mouse pointer.

B. Mouse Clicking

The proposed system uses close gestures for clicking events. When the rectangle bounding boxes come closer to another rectangle, bounding box is created with the edge of the tracking bounding boxes. When the newly created bounding box becomes 20% of its creation time size then the system performs left button click and it can be clicked. By holding this position more than 5 seconds, the user can perform a Double click. In addition, for the right button click again the open gesture is used. To perform the right button click, the single finger is good enough. The system will detect one fingertip color cap then it performs a right button click.

C. Mouse Scrolling

To scroll with this system, user needs to use the open gesture movement with three fingers with color caps. If the users use their three fingers together & changes its position to downwards, it will perform scrolling down. Similarly, if its position is changed to upwards, it will perform scrolling up. When three fingers move up or down the color caps gets a new position and new coordinates. By the time all three color caps get new coordinates, it performs scrolls. If their y coordinate values decrease, it will perform scrolling down and if the values increase, it will perform scrolling up.

V. Result & Evaluation

In this work, we have used HCI (Human Computer Interaction) and computer vision in order to contribute in future vision based interaction between the machine and human.

The proposed paper is on controlling the mouse functions using hand gestures. The main functions are mouse movement, left button click, right button click, double click and scrolling up or down.

In this system, the users can pick any color from multiple colors. There are a few color bands defined & the users can pick any color from the colors according to the backgrounds and the lighting conditions. This may vary in a different background. For example, when a user starts the system, it will provide the choice to pick a color from multiple colors (Green, Yellow, Red, Blue & two others). The user will have to pick one color such that, the color would not match with the background or else, he has to choose the color which will be highlighted in the current background.

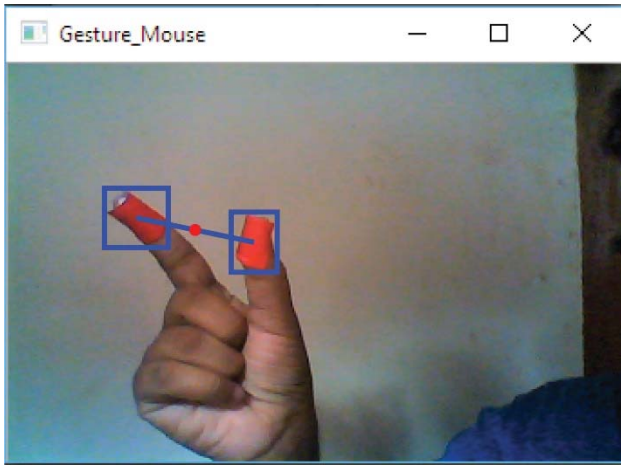


Fig. 2. Mouse Movement (Open Gesture)

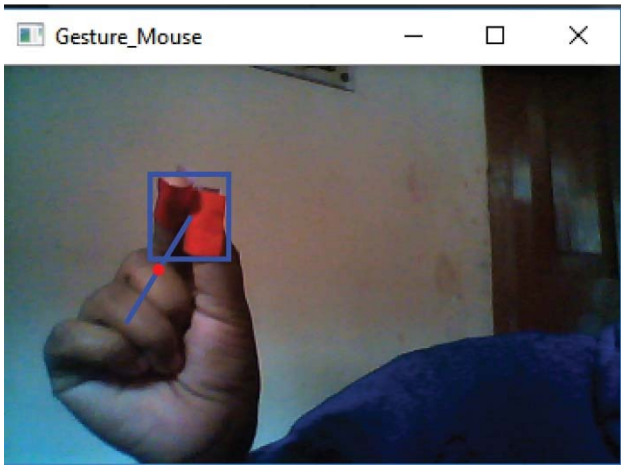


Fig. 3. Left Button Click (Close Gesture)

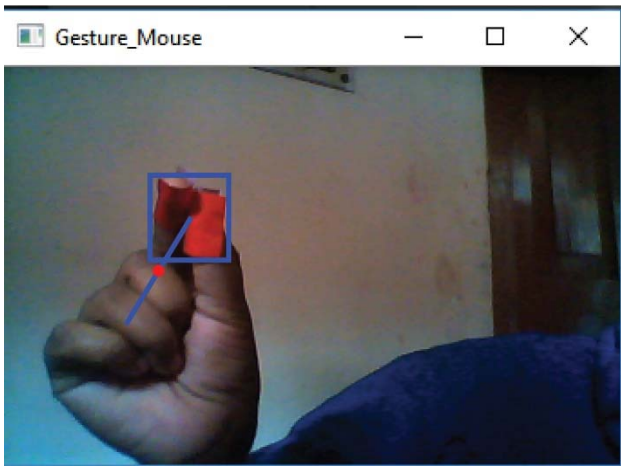


Fig. 4. 5 Seconds Perform Double Click (Close Gesture)

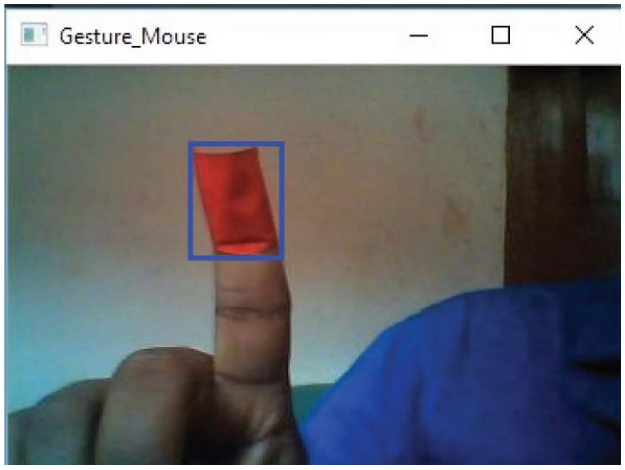


Fig. 5. Right Button Click (Open Gesture)

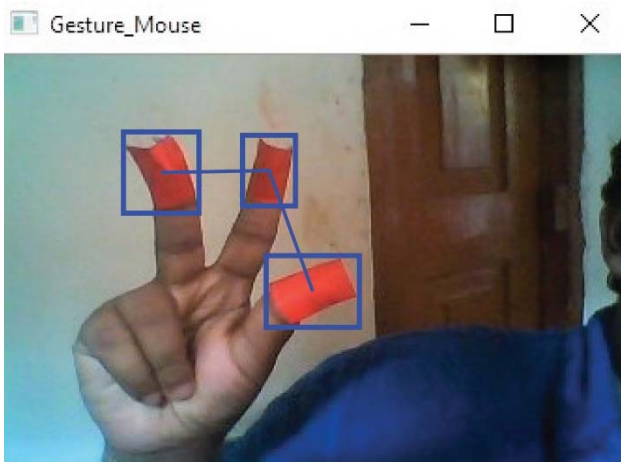


Fig. 6. Scrolling Up & Down (Open Gesture)

These testing are done with a plain background. It works very fast and the accuracy is 78-90 percent. It is also tested with some complex backgrounds, such as check printed shirt, colorful t-shirt (Though most of the people sit in front of the laptop or computer wearing regular shirts or t-shirts), complex home backgrounds, indoor daylight environments, indoor home light environments etc. These events are evaluated with the accuracy, and which is showed in a table.

TABLE I. MOUSE EVENTS & EVALUATIONS

Gesture Input	**Evaluation**		
	Mouse events	***Accuracy with plain background (in %)***	***Accuracy with complex/ non-plain background (in %)***
Two color caps (Open Gesture)	Mouse Movement	91	41
Closer two color caps (Closed Gesture)	Left Button Click	88	40
Keep closer for 5 seconds (Closed Gesture)	Double Click	87	42
Single color cap (Open Gesture)	Right Button click	95	79
Three color caps (Swipe Up/Down)	Scrolling up or down	78	40

Clearly, the system is not much efficient in the complex and rough background. However, our system is faster than the other gesture-based mouse control systems. This accuracy can be increased by using high definition camera. The readings of the table are taken from a built-in laptop VGA (Video Graphics Array) camera. Though this system is developed in Windows OS, in some events, the accuracy rate might be better in Linux base or Macintosh OS computer rather than Windows OS computer.

TABLE II. COMPARISON WITH EXISTING SYSTEMS

Existing Methods	**Comparison**		
	Recognized Gestures	***Average Accuracy (in %)***	***Control Type (Static/Dynamic)***
Color tracking and counting [8]	Five (05) : Mouse Movement, Left Click, Right Click & Double Click etc.	Not Available	Dynamic
Finger and palm tracking [9]	Seven (07) : Opening Media Player, Web Page, Launching Powerpoint etc.	78	Static
Color masking and pinchflag [10]	Three (03) : Mouse Movement, Left Button Click, File Transfer	Not Available	Dynamic
Our proposed method	Five (05) : Mouse Movement, Left Click, Right Click, Double Click, Scrolling	78	Dynamic

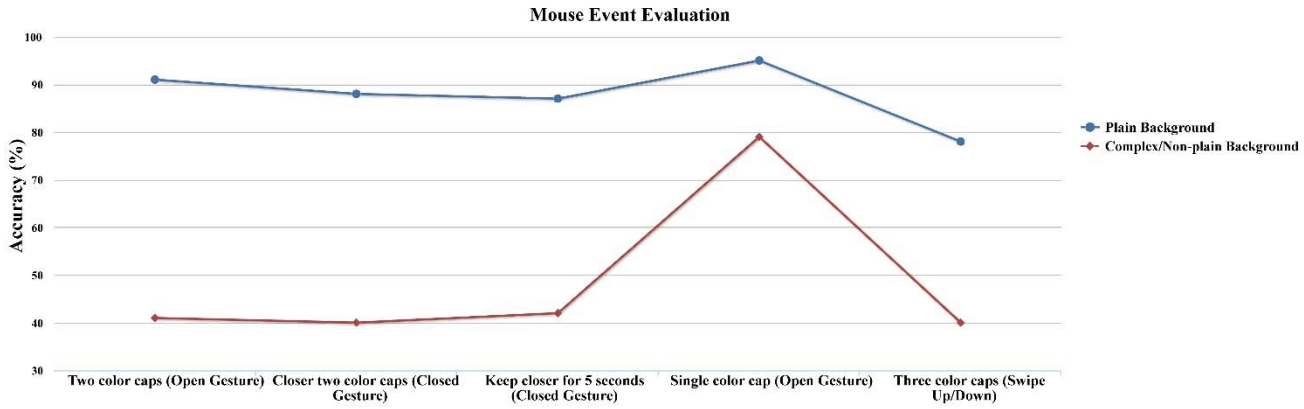


Fig. 7. Mouse Event Accuracy (%) Evaluation

VI. Application

This project can be useful for presentations and for reducing workspace and burden of extra hardware devices. In critical events such as battleground, operation theater, this project will be able to withstand itself, as it is wispier than any other contemporary system in the market for computer interaction. Since it removes the burdens of devices, it brings the user and the workspace more closer than before.

Major applications:

- Robot controlling is one of the major applications of this system. Controlling robots without machines or extra devices can be an excellent addition to this technology.
- Digital artist can draw 2D or 3D images using this gesture-based mouse on digital canvases. It will allow more freedom more dynamic more space to artists to create their art.
- Critical events like a battleground, operation theater, mining fields can be controlled by gesture mouse.
- Virtual Reality or Augmented Reality based games can be played without any extended or wireless devices with bare hands more comfortably.
- For patients with no limb control, this system can be proved very useful and effective.
- This mouse system can be applicable for sign language for the dumb & deaf people. It can help them to interact with computing machines.

VII. Conclusion

Virtual gesture control mouse is a system that profound to guide the mouse cursor and execute its task using a real-time camera. We implemented mouse navigation, selection of icons and its functions and tasks like left, right, double click and scrolling. This system is based on image comparison and motion detection technology to do mouse indicator movements and selection of the icon. Analyzing results, it can be anticipated that if we provide enough light, decent camera, the algorithms can work at any domain. Then our system will be more systematized. In future, we want to merge more features such as interaction in multiple windows, enlarging and shrinking windows, closing window, etc. by using the palm and multiple fingers.

Acknowledgement

We would like to express our special thanks of gratitude to our teacher, supervisor Samrat Kumar Dey who gave us the golden opportunity to do this wonderful project on the topic Virtual Mouse using Gesture Recognition, which also helped us in doing a research work and we came to know about so many new things for which we are really thankful. Secondly, we would also like to thank our parents and friends who helped us a lot in finalizing this project within the limited time frame.

References

[1] D.L. Quam, “Gesture recognition with a Data Glove,” 10.1109/NAECON.1990.112862, vol. 2, pp. 755 - 760, 1990.

[2] C.-Chiung Hsieh, D.-Hua Liou, & D. Lee, , “A real time hand gesture recognition system using motion history image,” Proc. IEEE Int'l Conf. Signal Processing Systems (ICSPS), 2. 10.1109/ICSPS.2010.5555462, 2010.

[3] C.-Y. Kao, C.-S. Fahn, "A Human-Machine Interaction Technique: Hand Gesture Recognition Based on Hidden Markov Models with Trajectory of Hand Motion," in Procedia Engineering 15 pp. 3739 – 3743, 2011.

[4] P.S. Neethu,”Real Time Static & Dynamic Hand Gesture Recognition,” International Journal of Scientific & Engineering Research, Vol. 4, Issue3, March 2013.

[5] A. M. Patil1, S. U. Dudhane1, M. B. Gandhi1,” Cursor Control System Using Hand Gesture Recognition”, International journal of advanced research in computer and communication engineering, vol. 2, issue: 5, May 2013.

[6] A. Banerjee, A. Ghosh, K. Bharadwaj, “Mouse Control using a Web Camera based on Color Detection,”IJCTT,vol.9, March 2014.

[7] Y. Zhou, G. Jiang & Y. Lin, “A novel finger and hand pose estimation technique for real-time hand gesture recognition,” in Pattern Recognition vol. 49, pp. 102-114, January 2016.

[8] P. Kumari, S. Singh & V. Kr. Pasi, "Cursor Control using Hand Gestures," Recent Trends in Future Prospective in Engineering & Management Technology 2016, International Journal of Computer Applications (0975 – 8887), 2016.

[9] A. Haria, A. Subramanian, N. Asokkumar, S. Poddar & J. S. Nayak, “Hand Gesture Recognition for Human Computer Interaction," 7th International Conference on Advances in Computing & Communications, ICACC-2017, Cochin, India, Procedia Computer Science 115 (2017), pp.367–374, 22-24 August 2017.

[10] S. S. Abhilash, L. Thomas, N. Wilson & C. Chaithanya, "VIRTUAL MOUSE USING HAND GESTURE," International Research Journal of Engineering and Technology (IRJET), vol. 05 issue: 4, April 2018.

Nonparametric development of the ranked tree structure for the accurate facial landmarking localization

1st Supratip Ghose
Dept. of Computer Science & Engineering
University of Information Technology & Sciences
Dhaka, Bangladesh
SGresearch@gmail.com

2nd Habibur Rahman
Dept. of Computer Science & Engineering
University of Information Technology & Sciences
Dhaka, Bangladesh
habib.rech@gmail.com

***Abstract*—Reliable face recognition starts with the analysis of accurate and robust facial localization component and poses an unresolved challenging computer vision problem in unconstrained settings in the backdrop of clutter, and large head poses variations. The landmark localizer needs to meet robust reasonable point coordinates in correspondence with the face alignment despite to the extent the image in consideration, is under scenarios with face variability. Correspondingly, in this paper; the task is viewed as non-parametric development of local estimator that produces learned features in the leaf node of a tree with an adopted ranking in a supervised clustering way. Then, the proposed approach helps to implement a sequence of estimators based on ensembles of regression trees. The trees use simple scale invariant, asymmetirc Multi-Scale shape indexed pixel difference at each split in the internal nodes using adopted label ranking clustering methods, and this makes them able to rapid detection of the process in the region of interest. The developed system is then tested on several publicly available datasets and analyzed on the sensitivity of the experimental occlusion. The results show that the performance of our method has practical value for commercialization aspect overcoming slowness in the state-of-the-art methods and evaluates to significant accuracy compared to recent methods.**

***Index Terms*—: Facial landmark localization, clustering, detection process, intermediary face shape**

I. Introduction

The cornerstone of all applications of computer vision face analytics; Facial Landmarking with spatial and dynamic configurations of landmarks is a subset of the shape prediction problem, characterized as the detection and localization predefined descriptor points of fiducial salient **(reference gradient)** points for facial expressions or variations due to gestures [15]. In recent years, localizing the landmark points on the face images captured under unconstrained conditions have garnered avenues and witnessed a monumental growth from both scientific challenges in research and industrial practices in the backdrop of pose, illumination and occlusions of face variability [1]. Variation in studies include model fitting methods with an active appearance model and active shape model respectively in [4], [5] tend to build templates that fit the region of interest of test images. Dollar [12] et al. uses parametric model which minimizes model parameters in the training. However, smaller parametric errors do not lead to small landmark errors for alignment. The above landmark research showed a number of optimization criteria fits with the effect to have the minimized residuals between the face appearance and synthesized face template. However, due to the existence of nonlinearity between the input face and output shape, the parametric shape models suffer significant variations in the wild with limited expressive power to deal with.

On the other hand, Regression-based models measure the landmark location by local (or global) regression using image features, for example, in [7], the author uses a local image patch with SVM regressors. Belhumeur et al. [2] combines the output of SIFT-based face part detectors as local experts or regressors for constrained local models with a non-parametric global shape model for the part or entire of the face locations. The nonparametric approaches avoids the iterative alignment methods, use sequential regressors to better convergence and fast regularization. However, the downside of this approach is its low processing speed, the presence of handcrafted noisy features causes overfitting in testing. We avoid the shape constraint in the post-processing steps, instead, generate local descriptors directly for landmark localizations. Li et al. [14] proposed a special feature called SURF in a cascade of AdaBoost settings, and area under the curve (AUC) criterion to scale up the speed of face detection algorithm which in our method is adopted as multi-scale Pixel difference features with cascaded realboost settings. Cao [3] proposed a boosted regression-based method realizing nonparametric shape constraint for progressive face alignment in which features are extracted as the pixel intensity difference, thus avoids the flexibility concern in parametric model [7]. Then, local evidence with regression employed boosting method where the few prominent landmark points are initialized, and in the global stage image and shape estimation is computed from sequential regressors. We used tree-based estimators with the Decision Tree algorithm for label ranking proposed by Cheng [8]. Our work in the contribution perspective is different from the earlier researchers is that current cluster central ranking induces instances to clusters of landmarks in

a tree growing iteratively with the mechanism of supervised clustering and these new instances to clusters gives rise to reassigned cluster central rankings. Later we analyze it in the framework designed by earlier researchers [16] and found to have significant F-measure compared to the baseline algorithm. The paper is organized as follows. Section II described the adopted pixel difference features are described with the description of one stage regressor as a ranked list of greedy tree construction with the ranked leaves as output. A model of supervised decision tree in the form of clustering label is proposed. Section III describes the dataset and experimental set up and results. Finally, section IV describes the conclusion.

II. Methodology

A. Formation of a localize feature

The methodology of our systems inspired by the Viola-Jones [9] method of the learning features of pixel difference with a little variations in scaling. The adopted idea in the similar research is to learn the local tree structure with standard regression forest that estimates the increment of each $L\oplus$ facial point of the image with a cascade of binary classifiers at all reasonable positions and scales using the multi-scale shape indexed pixel difference features (MPDF). Despite split test for an internal node, we boost the shape indexed MPDF in [3] to multi-scale. Specifically, we downsample the input image to half and one fourth by scaling it upto three scales of the image. Soliciting a desirable feature starts with the selection of two random facial points as described in (1), retrieve two random offsets with those points. Then, we take the difference of these two points as the feature and thus this attribute is coined as pixel difference features is invarient to various illumination conditions as it expresses the related difference of pixel intensity values and carrying ordinal relationship in sign changes. We used the variations of Chen et al. in [13] computes a Weber Fractions of pixels in a settings of 3 3 window which is a local image descriptor, called Webers Law Descriptor for face recognition in which is scales as normalized difference.

The construction of MPDF adopted in the paper between two pixels in an image is defined as below:

$$f(\mathbb{P},\mathbb{N}) = \frac{\mathbb{P}-\mathbb{N}}{\mathbb{P}+\mathbb{N}} \tag{1}$$

where $\mathbb{P},\mathbb{N} \geq 0$ are intensity values of the two pixels and $f(0,0)$ is defined as 0 when $\mathbb{P},\mathbb{N} = 0$.

We found the MPDF is more robust to noises and necessary for detection learning for fiducial points [1]. Asymmetry in (1) brings reduced feature space.

B. Weak Classifiers as greedy regression tree

We used a label ranking clustering scenarios [8] to recursively growing decision trees grown in a supervised way. The advantage of label ranking is that it involves strict label order relations, rather than single values in the leaf node. Specifically, in the scenarios with label rankings, a set of features $\varkappa$, represent each instance. This set of features maintains a ranked order of labels $\amalg$ in a form of total (e.g. $\amalg$= 5,3,1,4,2) or partial (e.g. $\amalg$=5,3,2) order over a finite set of class labels $\aleph$ (e.g., $\aleph$=1.2.3.4.5). The label ranking problem in our paper, comprises of representing a model that maps instances $\varkappa$ to a total label order $f : \varkappa \rightarrow \amalg$.

Let us assume from the labeled faces, training data $\mathcal{D} = (\varkappa_i, \aleph_i), n = 1, , N$, is the distribution in which it is to be represented as $\Bbbk$ clusters by learning cluster representation in the feature space in the following:

$\oplus_{\Bbbk} = (\nu_{\Bbbk}\eth_{\Bbbk}) \rightarrow \eth_{\Bbbk}$ is an estimate of $\hat{\amalg}$, in which $\Bbbk$ -th cluster is a representation of the $\nu_{\Bbbk}$ clusters $\Bbbk = 1, 2, 3, ..., \Bbbk$ as following: $\nu_{\Bbbk} \rightarrow$ represents a cluster centroid, say, for example, feature abstraction that spans a certain cluster, for example, in this scenario, it outcomes as a Decision Tree. Unassigned instances can get membership of a cluster based on the feature vector $\varkappa_\nu$ by measuring a ranking of central tendency and is associated with a certain central ranking $\eth_{\Bbbk}$, $\Bbbk = 1 \cdot \cdot \cdot \cdot \cdot \cdot , \Bbbk$, both to be determined from dataset $\mathcal{D}$.

$\amalg_u \rightarrow$ In the label ranking scenario, a single instance $\varkappa$, described by dimension $\lceil$ is associated with a total order or assigned class labels. Total order is assumed by transitive and asymmetric relation $\succcurlyeq_\varkappa$ on a finite set collection of labels $\curlyvee_i \succcurlyeq_\varkappa \curlywedge_i$ given $\varkappa$

$d_e(\varkappa_\nu, \nu_{\Bbbk}) \rightarrow$ is the new unlabeled instance x_ν is assigned to the $m-th$ cluster feature space abstractions $\nu_\varkappa$, and as such is assigned a ranking $\amalg$. Kendall tau correlation [8] provides an correspondence between the two measurable quantities of label ranks and represents a rank correlations as distance measure by counting the number of mutually inclusive and conflicting label pairs.

Clustering weight: The most common clustering measures taken in this paper is the within-cluster sum of squares (WSCC) as in (2)

$$\lceil_{mean} = \frac{1}{k}\sum_{k=1}^{k}\sum_{n\in\oplus_{\Bbbk}} \|\varkappa_n - \nu_{\Bbbk}\| \tag{2}$$

in which $\nu_{\Bbbk}$ is the mean of points in cluster $\odot_{\Bbbk}$. However,$\mathcal{K}$-means algorithm is the direct cause of reduction of this measure. The main disadvantage of this approach is that it assumes global clusters. A modification is of d_knh as from [11] is obtained by simply calculating the average distance of each example $\varkappa_n$ from cluster $\odot_m$ to its $\mathcal{K}$ neighbors in each cluster $\odot_m$.

$$d_{knh}(n,\oplus) = \frac{1}{k}\sum_{i\in knh} \|\varkappa_n - \nu_{\Bbbk}\|^2 \tag{3}$$

C. Ranking Model

We have adopted the Mallows sampling model, which finds a central ranking among all the available label rankings using the maximum likelihood principle. For any distributions, the distance based probabilistic model proposed by Mallows yields the likelihood of any permutation $\eth$ of $\mathcal{L}$ labels is given in terms of a permutation distance $\lceil$ as follows

$$\mathbb{P}(\eth|\Theta,\amalg) = \frac{exp(-\Theta d(\amalg,\theta))}{z(\Theta,\eth)} \tag{4}$$

in which, $\eth \in \Re$ is reflected candidate form of permutation, $\amalg$ is central form of ranking, $\theta \in \Upsilon$ is a parameter for the measure of dispersion, $z(\theta, \amalg) = \sum_{\eth \in \Re} exp(-\Theta d(\eth, \amalg))$ is a normalization constant, $\lceil$ is assumed to be the $Kendalls$ tau distance d_τ. The maximum probability is assigned to the central ranking $\diamond$.

Several recent papers [8], [10] proposed to use decision Trees with binary splits for solving Label ranking problems. The proposal contains a recursive induction of decision tree by finding optimal splits that make the subsets as pure as possible regarding label ranking at each stage [8]. In our system, the construction of the optimized binary decision is formed by MPDF features of the regression tree. The tree uses variations of pixel difference as its comparison as a binary test in the internal nodes. Formally, given data subsample $\Re$, which iterates and splits into subsets by $\Re^+$ and $\Re^-$ by generalizing the sets trained on Mallows model and generating the ranks by formulating the variance in each subset. The results are the subset central ranking $\eth^+$ and $\eth^-$ proposed by Mallow model and the spread parameters θ^+ and θ^- that estimate the variance. Training data is a set $(\clubsuit_{\mathfrak{S}}, \Psi_{\mathfrak{S}}, \Delta_{\mathfrak{S}})$ in which $\mathfrak{S} = 1, 2, , \mathfrak{S}$ in which $\Psi_{\mathfrak{S}}$ is the ground truth value associated with image, $\clubsuit_{\mathfrak{S}}$ and the weight $\Delta_{\mathfrak{S}}$ is its importance factor (weight) calculated from the (3). Likewise, the binary classification of ground truth yields a value $+1$ and -1. The weights Δ_{st} enable us to assign different importance values to each different sample. The optimal split is the one that minimizes error and maximizes spread parameter obtained when the incoming training data is split by the test as in (5).

$$|\Re|^{-1}(\Re^+|\theta_{\Re^+}|\Re^-|\theta_{\Re^-}) \tag{5}$$

in which, $\Re^-$ and $\Re^+$ are the representative clusters of training instances in which binary check on an associated image were 0 and 1, respectively from S_m of cluster (4).

Recursively clustered the training data in this way happens until some termination condition is met. This is performed by minimizing the following quantity:

We measure the rank losses at each node $\Re$ as in (6) and prunes leaves of parent node with minimum value of losses

$$\sum_{n \in \Re} d_\tau(\amalg_n, \eth_\Re) - d_\tau(\amalg_n, \eth_\Re) + (\sum_{n \in \Re^-} d_\tau(\amalg_n, \eth_\Re^-)) \tag{6}$$

The disadvantage is that the final tree will have a fewer number of clusters-splits to adequately cluster the feature space if it would form a partial rank. Our algorithm starts by randomly selecting $\Bbbk$ ranks and growing a tree limited to the assigning one to the $\Bbbk$ ranks to the leaves set $\Re$ is iteratively split into two data subsets $\Re^+$ and$\Re^-$ by trying each of the $\Bbbk$ rankings for the subset central rankings $\eth^+$ and $\eth^-$ and selecting the best separation which reduces

$$\sum_{n \in \Re^+} d_\tau(\amalg_n, \eth_{\Re^+}) + (\sum_{n \in \Re^-} d_\tau(\amalg_n, \eth_\Re^-)) \tag{7}$$

and thus achieves the full rank. After the completion of the constructed tree, the $\Bbbk$ ranks are recalculated by finding a central ranking of examples assigned to each of the $\Bbbk$ clusters using the modified Borda count. The bound box of MPDF is purturbed by repeated sampling of two locations from a uniform distribution on a square $[-1, +1] \times [-1, +1]$.

A more generalized version of AdaBoost in the form of RealBoost [12], [14] is taken as a degenerative decision tree which is called a cascade inspired by the work [3], [10], [12], [14]. The earlier model adopts a boosted decision stump as a tree classifier which degenerates into two leaves constrained by the presence of threshold. However, higher order information curtails and less intractable feature dimensions are obvious problems. Each regressor modeled as an ensemble of trees which learns a greedy optimization of the weak learners to minimize the regressor loss function and to improve the reduction of the alignment error bounded by an optimal threshold. Each stage is obtained by confidence score of the soft output of the decision stump classifier of the previous stage. Thus multiple weak classifiers with reasonable accuracy are attained by rank based sampling of the outputs of the leaves of all classification stages in the cascade.

III. Dataset Description & Experimental set up

We have taken the most relevant face databases suitable for landmarking studies with Ground truth data. Ground truth data provides the manually landmarked spatial positions of the landmark points suitable for the referenced landmarking performance.

A. LFPW & LFW

The Labeled Face Parts in-the-wild (LFPW) database was created in [2] contains 1,287 images downloaded from google.com, Flickr.com, and yahoo.com 1132 training images and 300 test images taken under unfavorable unconstrained conditions as the challenging task (in the wild) with significant variations in the pose, expression, illumination and with partial occlusions. From the initial dataset described in LFPW, we use only the 811 training images and the 224 test images provided by the ibug website Ground truth annotations of 68 facial points. We divide the available training samples into two sets: a validation set (135 samples) and a training set (676 samples). The LFW dataset is with 87 annotation was created by Huang et. al. [17] with 5,749 individuals from 13,233 images. Out 5,749 images, 4079 have distinct face images.

B. BIOID

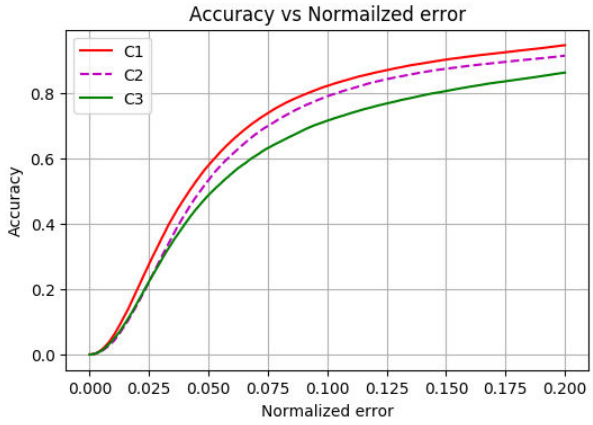
The BioID dataset is composed of grayscale face images with different sizes, illuminations, and backgrounds from 1521 approximate frontal faces was first proposed in Jesorsky et. al referred in [3], [15]. We used 17 of 20 facial landmarks manually annotated. We selected a subset of BioID that is of BioID-2 which consists of 1455 images, some of which include glasses, beards and blinking eyes; 790 images were used for training, 260 for validation, and 405 for evaluation.

Thus, we evaluate performance according to these four characteristics. Accuracy is evaluated according to the distance between the selected optimal landmark and the manually annotated landmarks as ground-truth. In both cases, we used

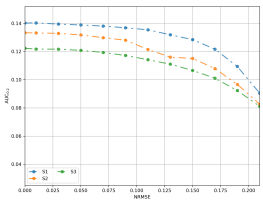
annotation in the dominant contours of faces, that is, in the locations of eyebrows, nose, upper and lower lip, mouth and eye corners.

C. Experimental result

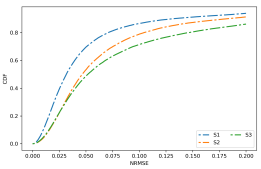
In our experiments, the normalized root-mean-squared error (NRMSE) relative to the ground truth is adopted as the error measurement for the accuracy of the facial landmark detection. The NRMSE is given as a percentage, computed by dividing the root mean squared error by the distance between the two eye centers. The cumulative distribution function (CDF) is a typical cumulative error distribution (CDE) in which normalized error is represented by values in hte x-axis and proportion of test images are represented by values in the y-axis and thus, CDE of NRMSE is used to evaluate the performance of facial landmark detection algorithm.



(a) Accuracy Representation



(b) Face Shift VS AUC



(c) Face centre shift in CDF

Fig. 1. Graphs for Accuracy and Face Centre Shift

For the test set performance and overfitting minimization, Data augmentation happens with rotations, scale and translation and mirroring the imageset.

Each landmark point position estimation structure is learned independently in our framework. A progressive cascade constitutes of 6 stages with 18 trees having a depth equal to 9 yields a good results in practice. The bounding box estimates the position of a particular landmark point induced by supervised clustering tree in the ensembles. The learning of the landmark point is individually estimated. The local learning of each next stage is mapped on a training set with features generated by shrinking the bounding box by 0.7 and updating its center by moving it at the position output by the ensemble of the previous stage. In this way, invariance regarding scale and orientation are formed to scale the system robustness. times until The refining the predictions points are done in this way by several times of the process repetition.

First, we compare our system to the state-of-the-art methods based on a novel method using Boosted Regression coupled with Markov Networks [6] and local binary pattern applied [7]. The proposed method was observed to be performed significantly better than all previous methods in the below settings. We set the protocols from [3] to obtain the normalized errors for five facial landmarks with facial features like eye centers, mouth corners, and the tip of the nose as facial landmarks presents in the vicinity of the dominant contours around the facial features. The results are reported in the Fig. 1 in Fig. a. We can see that our method labelled as C1 outperforms the benchmark works in accurate landmark localization ($\varepsilon \approx 0.75$). One reason for this good performance is that the tree indexed learned to map in the leaf and provide exact image representation anticipating the ranked features in the leaf. Secondly, the regressor uses prior point information to localize the regression point very effectively in each state. Third very few features are required to mask the image.

Moreover, our method is approximately thousand times faster. It takes around 3.4 ms per image on a Pentium core $\rangle$ı with 3.5 GHz speed machine which is faster than OpenCV based systems on the same machine, in which it runs on nearly 16.4 ms and 17.8 ms for a detection of faces larger than 100100 windows of pixels in aspect around 640480 size image with 29 landmark detectors. It even outperform Belhumeur [2] as the system uses expensive local detectors (SIFT+ SVM) for consensus exemplars approach and it takes more than 10 seconds for 29 landmark detectors.

Fig. (b) in Fig. 1 shows that the maximum shift of bounding box from the ground truth is shrunk from about $\backsim$ .40 to .10. This is due to the small percentages $\backsim$ 1 to 2% of the ground truth rejection rate in each stage.

However, CDF is good for sensitivity analysis and AUC is a better choice according to literature for sensitivity analysis. Therefore, we provide Fig. 1 to have a good metrics for performance evaluation. $AUC_{0.2}$ is seen as the upper bound of the definite integration in the form $\int_0^{\upsilon} f(e)de$ in which $\upsilon = .2$. Fig. 1 shows the result in Fig. b and Fig. c respectively for face center shift for AUC and CDF.

The Later test is the nonparametric development of rank order label clustering taking the comparison with the experimental design in [16] in which, knn approximation methods are called in terms of F-measure over the dataset LFW in [17]. We induces a certain form of adjacency matrix with ground-truth 120 numbers of clusters taking the distance measure as in equation (4). The results are shown in Table I.

For k-Means the clustering algorithm defined in equation (4) produces clusters with ground truth number of clusters, where in ranked-order clustering labels depends on the distance threshold as defined in section II and in equation (7). Ranked order labeling induces more diverse discriminative

TABLE I
RANKED CLUSTERING IN THE LFW DATASET

Nonparametric development	Clusters	F-Measure
K-Means	*120*	.36
Our Method Rank order	*6000*	.83

features as fixed number of faces in large fraction of nearest neighbor lists and thus has to be evaluated at several effective values of clusters. The result of k-means algorithm shows poor clustering performance with lower value of F-measure, whereas, in rank-order clustering, the distance threshold which induces the cluster label with best overall F-measure close to the ground truth identity and thus scores significant F-Measure overall than K-Means.

IV. CONCLUSION

In this work, we have presented a clustering approach to learning local binary features for highly accurate and speedy face alignment for commercial perspective. The proposed approach helped to implement a sequence of estimators based on ensembles of regression trees. Individual estimators find the ranked result in the decision tree with better accuracy and sensitivity. Our method significantly performed better than the baseline methods for localization of landmarks accurately for the normalized root mean square and sensitivity analysis. Besides rank order clustering is effectively analyzed to score over the baseline algorithm. In the future, we will include more experimentation on effectiveness of ranked estimator which ultimately proves the commercial advantage of using this proposal.

REFERENCES

[1] eliktutan, Oya, Sezer Ulukaya, and Blent Sankur. "A comparative study of face landmarking techniques." EURASIP Journal on Image and Video Processing 2013, no. 1 (2013): 13.

[2] Belhumeur, Peter N., David W. Jacobs, David J. Kriegman, and Neeraj Kumar. "Localizing parts of faces using a consensus of exemplars." IEEE transactions on pattern analysis and machine intelligence 35, no. 12 (2013): 2930-2940.

[3] Cao, Xudong, Yichen Wei, Fang Wen, and Jian Sun. "Face alignment by explicit shape regression." International Journal of Computer Vision 107, no. 2 (2014): 177-190.

[4] Matthews, I., & Baker, S. (2004). Active appearance models revisited. International journal of computer vision, 60(2), 135-164.

[5] Cootes, Tim F., Mircea C. Ionita, Claudia Lindner, and Patrick Sauer. "Robust and accurate shape model fitting using random forest regression voting." In European Conference on Computer Vision, pp. 278-291. Springer, Berlin, Heidelberg, 2012.

[6] Valstar, Michel, Brais Martinez, Xavier Binefa, and Maja Pantic. "Facial point detection using boosted regression and graph models." In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pp. 2729-2736. IEEE, 2010.

[7] Zhu, Xiangxin, and Deva Ramanan. "Face detection, pose estimation, and landmark localization in the wild." In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, pp. 2879-2886. IEEE, 2012.

[8] Cheng, Weiwei, Jens Hhn, and Eyke Hllermeier. "Decision tree and instance-based learning for label ranking." In Proceedings of the 26th Annual International Conference on Machine Learning, pp. 161-168. ACM, 2009.

[9] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1, pp. I-I. IEEE, 2001.

[10] Burgos-Artizzu, Xavier P., Pietro Perona, and Piotr Dollr. "Robust face landmark estimation under occlusion." In Proceedings of the IEEE International Conference on Computer Vision, pp. 1513-1520. 2013.

[11] Kremer, Hardy, Philipp Kranen, Timm Jansen, Thomas Seidl, Albert Bifet, Geoff Holmes, and Bernhard Pfahringer. "An effective evaluation measure for clustering on evolving data streams." In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 868-876. ACM, 2011.

[12] Dollr, Piotr, Peter Welinder, and Pietro Perona. "Cascaded pose regression." In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pp. 1078-1085. IEEE, 2010.

[13] Chen, Jie, Shiguang Shan, Chu He, Guoying Zhao, Matti Pietikainen, Xilin Chen, and Wen Gao. "WLD: A robust local image descriptor." IEEE transactions on pattern analysis and machine intelligence 32, no. 9 (2010): 1705-1720.

[14] Li, Jianguo, Tao Wang, and Yimin Zhang. "Face detection using surf cascade." In Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, pp. 2183-2190. IEEE, 2011.

[15] Zafeiriou, Stefanos, Cha Zhang, and Zhengyou Zhang. "A survey on face detection in the wild: past, present and future." Computer Vision and Image Understanding 138 (2015): 1-24.

[16] Otto, C., Wang, D., & Jain, A. K. "Clustering millions of faces by identity." IEEE transactions on pattern analysis and machine intelligence, 40(2), 289-303.

[17] Huang, Gary B., et al. "Labeled faces in the wild: A database forstudying face recognition in unconstrained environments." Workshop on faces in'Real-Life'Images: detection, alignment, and recognition. 2008.

A Vision Based Three-Layer Access Management System with IoT Integration

Nafize Ishtiaque Hossain
Department of Electrical and Electronic Engineering
Chittagong University of Engineering and Technology
Chittagong, Bangladesh
nafize.ishtiaque@gmail.com

Ali Reza Galib
Department of Electrical and Electronic Engineering
Chittagong University of Engineering and Technology
Chittagong Bangladesh
argalib.ac@gmail.com

Raihan Bin Mofidul
Department of Electrical and Electronic Engineering
Khulna University of Engineering and Technology
Khulna, Bangladesh
raihanbinmofidul@outlook.com

Dhiman Kumar Sarker
Department of Electrical and Electronic Engineering
Chittagong University of Engineering and Technology
Chittagong, Bangladesh
sagor.eee.cuet@outlook.com

***Abstract*— In developing countries, traditional access management systems ubiquitously use either keypad based password protection or radio frequency identification (RFID) card based protection. With the increased number of threats in recent years, these systems are becoming more vulnerable. If the password or the RFID card is somehow compromised, any unauthorized person can breach the system with ease. Considering and analyzing these issues, a cost-effective prototype of a vision based three-layer access management system with IoT connectivity was developed. In this paper, an access management system architecture is proposed based on the fusion of radio frequency identification, back propagation based face recognition and password protection. The system is also connected to a Node JS based web server. Whenever an access is granted or any unauthorized access is detected, an SMS and an email are sent to both the user and the system administrator. Data security is ensured by exploiting AES encryption on the system side and AES decryption on the server side. In case of face recognition, the access is granted if the achieved confidence level is over 80%.**

Keywords—AES encryption; back propagation; face recognition; Node JS; password; RFID;

I. INTRODUCTION

In today's world, significant priorities are being given to research and development of secure access management systems. As the level of security increases, the cost of systems increases accordingly. In developing countries like Bangladesh, it is quite overwhelming for small businesses and corporations to afford high-level and complex security systems. In most small businesses, access control systems are either based on password or radio frequency identification (RFID) cards or a combination of both. But it is seen that most of the time the access systems are quite vulnerable to potential threats. Unauthorized access is possible if the password or the RFID card or both are compromised somehow. To prevent such an inconvenience, a three-layer authentication based cost-efficient access control system integrating the internet of things (IoT) environment is developed. In section II literature review, section III system architecture, and section IV software architecture is described. Section V, VI and VII describe performance evaluation, comparative analysis and conclusion respectively.

II. LITERATURE REVIEW

Face recognition and gait analysis based identification system for access control is shown in [1]. From a video, at first human detection is done then after background subtraction and human silhouette, the silhouette is normalized to analyze a gait. Both frequency representation and binary representation is shown in [1]. Research in [2] shows face, fingerprint and palm vein based biometric security systems for the internet of medical things (IoMT) platform. A noninvasive and continuous biometric signal collection, as well as authentication from wearable medical device for access control, are shown in research [3]. Smart card based control system where all system data is protected by cryptography is proposed in a research [4]. A near field communication (NFC) and digital signature based access control system is in literature [5]. Here the access is granted after the validation of NFC and digital signature. Electromagnetic lock and modified RFID based system was also shown in some research [6] where electromagnetic lock is interconnected with a system which will grant access for some specific RFID tag. An artificial neural network (ANN) controlled face recognition based access control system has been developed recently [7]. Local binary pattern based face recognition for access control is developed in [9] were after conversion to binary image, analysis of the pattern of the binary image is done for face recognition. An FNN and RFID based access system for internal use is shown in research [8]. A stereo face recognition system is proposed previously by some researchers [10] which can be implemented only in low light indoor environment. Principle component analysis based face recognition is shown in some access systems [11] where the captured image can be seen from a web browser and at same time, a GSM modem sends this information to the user and the administrator. The incorporation of password, fingerprint and RFID card for several stage authentications is implemented in one of the recent researches [12]. PCA and back propagation based face recognition is shown in some

research [14] which can be very useful in future access control system. Again, logistic regression based face image recognition can be a potential resource to use in accesses control systems. In [15], a smart phone application based smart lock system is introduced.

III. System Architecture

The proposed system hardware includes a raspberry Pi 3, an RFID reader, a USB webcam, and a 4X4 keypad. The system block diagram is given in the following Fig. 1.

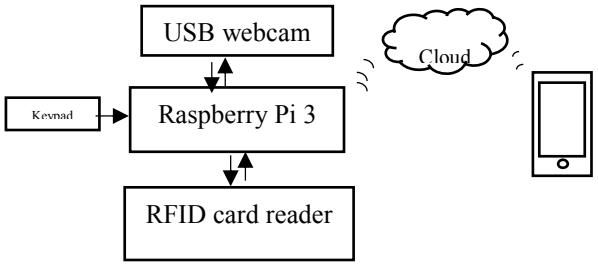
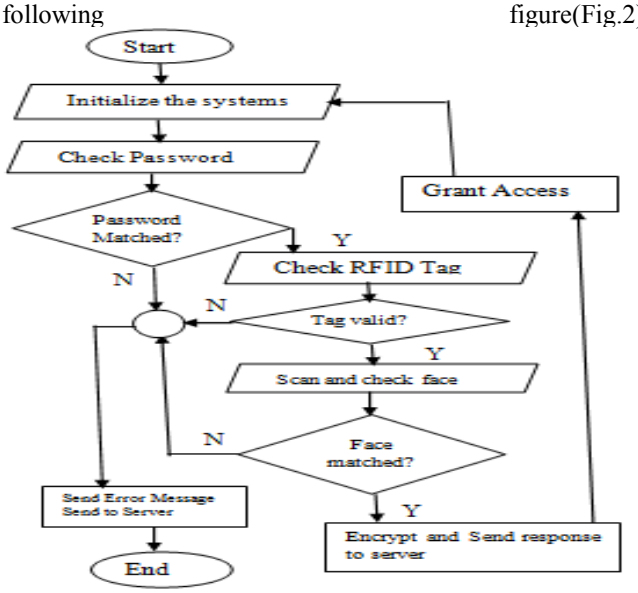


Fig. 1. System Architecture

The main processing part of authentication system is done in the Raspberry Pi 3, which is a single board computer with 1GB RAM, and the Linux operating system. To reduce power consumption, a 125kHz passive RFID tag is used and a low power RFID reader RDM600 [16] is connected with the Raspberry Pi 3 via serial communication, where the baud rate was 9600. The resolution of the USB web cam is 5M pixels, and the frame rate is 60fps. The focal length is manually adjustable from 3.5cm to 8cm. A lin-log CMOS image sensor is used in this camera.

If the user wants to access the facility, at first, the password has to be given via the keypad. After that, an RFID card has to be placed in front of the RFID reader. If the RFID card and the password are matched, then the face of the user will be scanned via the USB webcam. If all three of the systems are matched with the stored data, then the access is granted and an AES encrypted grant response is sent to the Node JS based cloud server. Upon receiving this message and decrypting it, the web server will send a verification SMS to the user and a notification SMS to the system administrator. The flow chart of the system is given in the following figure(Fig.2).



. Fig 2. Flow chart for access management

If any one of the step is mismatched, the system will send an error message to the web server which will be encrypted too. When the web server receives the error message, it will send an alert SMS to both the user and the administrator.

IV. Software Development

The software development can be divided into two main parts and they are system side software development and the server side software development. These two developments are described in the following two sub sections. The system side software was developed with python language and the server side software was developed in Node JS platform.

A. System Side Software Development

The main architecture of the software is based on the flow chart shown in the figure 2. The password verification, RFID match, and face recognition is employed via polling algorithm. The main section of the flow chart is the face recognition part. To recognize the face, back propagation based neural network is employed. At first, the image is captured and the face is detected. Then, the image is converted in to grayscale image. After that feed forward propagation is employed where each node of the network is characterized by sigmoid function denoted by,

$$g(x) = 1/(1+\exp(-z)) \quad (1)$$

Since sigmoid function is used, it has the probability to meet at the local minima so the modified cost function is defined as

$$J(\theta) = (1/m) * \{\Sigma\, y^{(i)} \log[g(x^{(i)})] + (1-y^{(i)}) \log[a-g(x^{(i)})]\} \quad (2)$$

The learning rate is chosen such that it is neither too small nor too large. If the learning rate is high, it has the risk to overshoot the minima. On the other hand, if the learning rate is too small then the it will cause prolonged convergence making the system impractical. To prevent oscillation, the momentum factor is chosen accordingly. The learning rate and momentum **f**actor took part in the modification of the weight.

The activation function of any layer is a_j^l and denoted by

$$a_j^{(l)} = g(\Theta^T.x) \quad (3)$$

Here Θ is the weight vector and Θ^T is the transpose of the weight vector.

If the error in jth node in layer l is denoted by

$$\delta_j^{(l)} = (\Theta^{l)\,T}\, \delta_j^{(l-1)}\ .\ *g^l(z^l) \quad (4)$$

Where $g^l(z^l) = a^{(l)}\ .*(1-a^{(l)})$.

After computing the error term, $\Delta_{i,j}^l$ will be modified as

$$\Delta_{i,j}^l := \Delta_{i,j}^l + a_j^{(l)}\, \delta(l\text{-}1) \quad (5)$$



Fig. 3.Face recognition dataset.

B. *Server Side Software Development*

The server side software development is more straight forward. In this development, a Node JS based IBM's server is used. In the subsequent two blocks after the decision taking block namely error response block in the flow chart in Fig. .4, Twilio's API is used for sending SMS to the user and the administrator. MongoDB API used for server side database development. It is realized by the following flow chart.

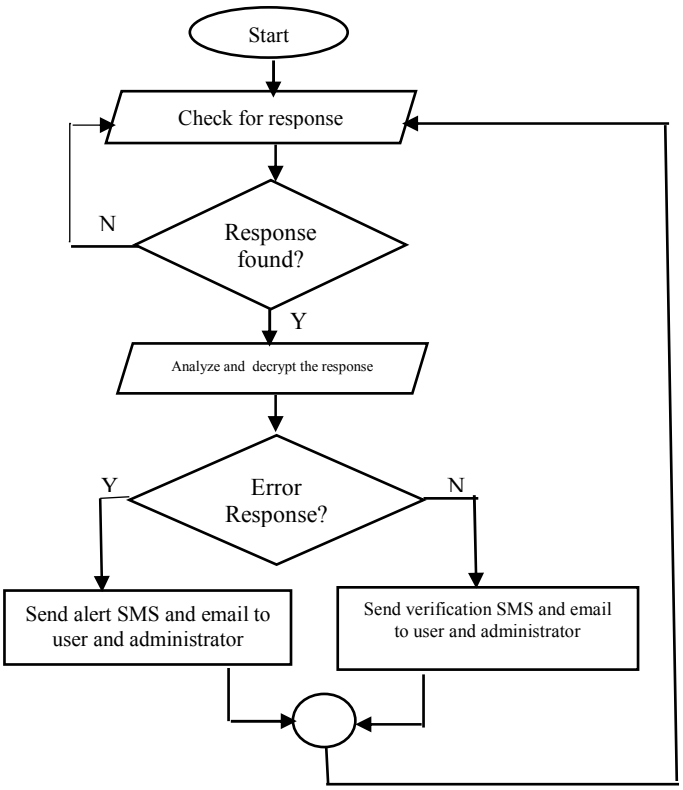


Fig. 4. Flow chart for web server side development

The practical implementation of the automated SMS and email in response to the defined situation is shown in the Fig. 5. The first two sub figures are email responses from the web server to the user about access notification and alert notification (Fig 5(a) and Fig 5(b)). The third sub figure is the SMS response to the user which notifies them about suspicious activity in their residence. (Fig 5c). Authorized face recognition and access management system hardware setup is shown in Fig. 6 and Fig. 7 respectively.

V. PERFORMANCE ANALYSIS

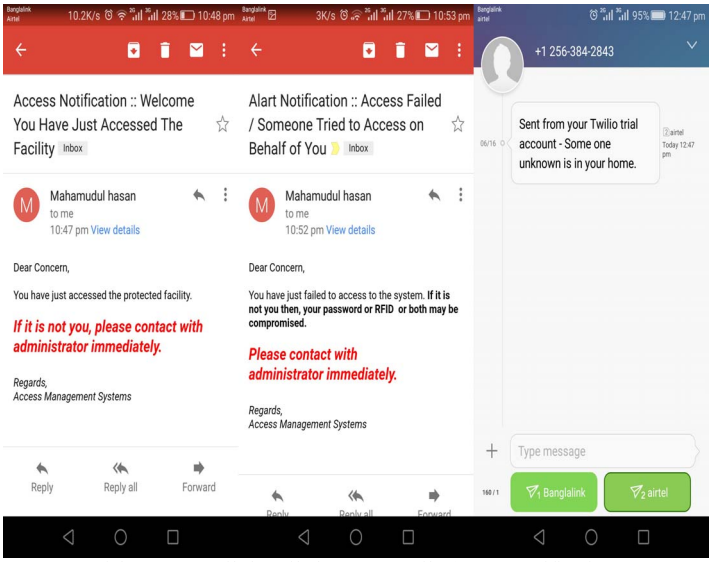
Performance analysis can be done in various ways. In this paper, performance is evaluated both qualitatively and quantitatively. These two analyses are described in the following subsections.

A. *Qualitative analysis:*

The qualitative analysis is done based on several parameters. These parameters are described in the following subsections.

a) Latency:

In real time embedded systems, latency is one of the key factors by which system's performance can be evaluated. The time interval between the application of input to the system and obtaining output from the system can be defined as system latency. There are three types of latencies in this system. They are Raspberry Pi 3 side system latency, IBM cloud server side latency and network side system latency. For simplicity, Raspberry Pi 3 side system latency and total system latency is measured since cloud side latency and network side latency are inconsistent over time. From the average of 50 instances, it is found that Raspberry Pi 3 side system latency is about 120ms in which the face recognition time is about 45ms as the total latency is about 3.5s.



(a) Successful access email (b)Failed access email (c) SMS notification

Fig. 5. E-mail and SMS based notification on mobile device

b) Data Security:

If the data of the system is stored properly and cannot be changed or altered, the security of the system is primarily ensured. The data is encrypted via AES 256 in system side and decrypted on server side. The key is being changed at a regular time interval to provide system integrity.

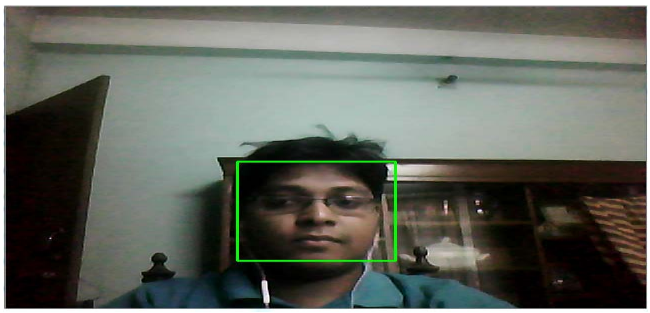


Fig. 6. Authorized face recognition



Fig. 7. Access management system

B. *Quantitaitve analysis:*

The quantitative analysis of the system is based on the performance of the back propagation of the neural network of the face recognition section. The following table shows the test data of three different samples. It is to be noted that each test is conducted 65 times. Every time true positive (TP), true negative (TN), false positive (FP) and false negative (FN) are calculated, and the recognition rate is denoted by the following equation.

$$RR = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

TABLE I. RECOGNITION RATE

	Test 1 $(\%RR)_{avg}$	*Test 2 $(\%RR)_{avg}$*	*Test 3 $(\%RR)_{avg}$*
Sample 1	83.31	81.32	85.11
Sample 2	85.32	82.71	87.31
Sample 3	84.74	81.57	89.74

For optimal performance of the face recognition system, the intensity of the light should be above 300 lux. All efficiencies (Test 1 and Test 2) of the table 1 are measured on 350 to 400 lux in daytime indoor condition. The system also recognizes the faces in day time outdoor environment(Test 3) where the light intensity is above 1000 lux.

Although it is tough to find homogeneous systems resembling the proposed system immaculately, the overall performance analysis of the access management system with some of the predecessor systems is shown in the following table.

TABLE II. PERFORMANCE ANALYSIS WITH RELEVENT SYSTEMS

Authors	*Latency*			*$(\%RR)_{avg}$*	
	System	*NN*	*Overall*	*Face*	*Others*
A. Derbel et al.[1]	N/A	N/A	N/A	**97.4%**	N/A
Y. Xin et al.[2]	N/A	N/A	N/A	**72.9%**	**70.4%(fingerprint) 68.2(palm vein)**
A. Mosenia et al.[3]	N/A	N/A	N/A	N/A	**91.1%(FRR) 91.8%(FAR)**
S.-H. Lee et al.[7]	N/A	N/A	N/A	**96.88%**	N/A
J.-J. Lin et al.[9]	N/A	N/A	N/A	**83%**	N/A
X. Peng et al.[10]	N/A	**3.6s**	**3.6s**	**92.9%**	N/A
F. Mahmud et al.[14]	N/A	N/A	N/A	**96.25%**	N/A
N. Xue et al. [15]	**0.6072s**	N/A	**.6074s**	N/A	N/A
Proposed work	**120ms**	**45ms**	**3.5s**	**89.74**	N/A

The cloud connectivity and integration of three layers of authentication in cost effective platform is unprecedented in the previously developed systems. Due to cost effective hardware and multistage authentication, the overall system latency is comparatively larger.

VI. COMPARATIVE ANALYSIS

The comparative analysis is accomplished based on comparing the developed system to some of the previously developed systems. The comparative analysis is shown in the TABLE III .

TABLE III. COMPARATIVE ANALYSIS WITH DIFFERENT SYSTEMS

Authors	*Password*	*RFID*	*Face recognition*	*IoT*	*Others*
A. Derbel et al.[1]	no	no	**yes**	no	**Gait**
Y. Xin et al.[2]	no	no	**yes**	no	**Finger print and palm vein**
A. Mosenia et al.[3]	no	no	no	no	**Bio signals**
L. Malina et al.[4]	no	**yes**	no	no	**Cryptography**
X. Wang et al.[6]	no	**yes**	no	no	No
S.-H. Lee et al.[7]	no	no	**yes**	no	no
PAN Xiang et al.[8]	no	**yes**	**yes**	no	no
J.-J. Lin et al.[9]	no	no	no	no	**NFC, digital signeture**
X. Peng et al.[10]	no	no	no	**yes**	no
M. Sahani et al.[11]	no	no	**yes**	no	**Zigbee GSM**
D. K. Sarker et al. [12]	**yes**	**yes**	**yes**	no	no
Proposed work	**yes**	**yes**	**yes**	**yes**	**SMS, Email, AES encryption**

In the proposed work, three layers namely, password, RFID and face recognition, based security system is developed and the IoT connectivity makes it easy for both the user and the administrator to verify the access. As the SMS and email are sent to user and administrator server, the access record is also preserved.

It is quite evident that most of the previously developed systems rarely concerned with more than two parameter access authentication and connectivity of those systems are seldom taken into consideration. Though some systems dealt with biosignals and zigbee based communication systems for ensuring systems data security, those systems showed seemingly complex method which will not only contribute to higher system latency but also will be prone to data taping if PAN ID of zigbee network is compromised or bio signal pattern is recognized via deep neural network. To prevent data tapping AES algorithm is employed in this work and the key is modified at a defined time interval which ensures data security.

In summary, few previously developed systems exploit three layers of hardware authentication, one layer of software based security for prevention of data tapping and three different connectivity schemes for access management system.

VII. CONCLUSION AND FUTURE WORKS

In this paper, three layer authentication based access management where data security is ensured via AES algorithm is developed. Implementation of three hidden layer back propagation based face recognition feature for both low light(indoor application) and day light(outdoor application) condition, which had recognition rate (RR) ranging from 81.32%(at low light) to 89.74%(at day light), improves the performance of the access management system. Incorporation of IoT environment for connectivity with AES algorithm for data security in software layer makes the system both flexible and secure simultaneously.

Presently, this system is using IBM's free IoT cloud server bluemix on trial basis which offers generic and limited development on their server. For dedicated smart

phone application stable api of the cloud server is required which is not available in this trial cloud platform. The free bluemix server also has limitations that prevent to employ good lossless data compression algorithm with fairly large blocks of code in server side. This free server also does not support over the air (OTA) update on remote devices. In future, when dedicated server for this access management system will be developed then, cloud connected smart phone application, exploitation of lossless data compression and OTA update on remote devices will be possible.

REFERENCES

[1] A. Derbel, D. Vivet and B. Emile, "Acess control based on gait analysis and face recognition", Electronics Letters, Vol. 51, No. 10, pp. 751–752, 14 May 2015

[2] Y. Xin, L. Kong, Z. Liu ,C. Wang, H. Zhu, M. Gao, C. Zhao, and X. Xu, "Multimodal Feature-Level Fusion for Biometrics Identification System on IoMT Platform" IEEE Access, vol. 6,pp. 21418 – 21426, 13 March 2018

[3] A. Mosenia, S. Sur-Kolay, A. Raghunathan, and N. K. Jha, "CABA: Continuous Authentication Based on BioAura" IEEE Transactions on Computers, Vol. 66, no. 5 , pp. 759 – 772, 27 October 2016

[4] L. Malina, V. Benes, J. Hajny, P. Dzurenda, "Efficient and Secure Access Control System Based on Programmable Smart Cards" in proc. of 40th IEEE International Conference on Telecommunications and Signal Processing (TSP), 5-7 July 2017,pp. 32-36

[5] C.C. Wu, C.-W. Hsu R.-S. Cheng"The digital signature technology for access control system of mobile" in proc. Of *IEEE International Conference on Applied System Innovation,* 13-17 April 2018 *pp.896-898*

[6] X. Wang, Y. Wang, " An office intelligent access control system based on RFID" in proc. Of *IEEE* Chinese Control And Decision Conference (CCDC), 9-11 June 2018,pp. 623-626

[7] S.-H. Lee, C.-S. Yang, "An Intelligent Home Access Control System Using Deep Neural Network" " in proc. Of *IEEE* International Conference on Consumer Electronics - Taiwan (ICCE-TW) 12-14 June 2017 *,pp. 281-282.*

[8] PAN Xiang "Research and implementation of Access Control System based on RFID and FNN Face Recognition" in proc. Of *IEEE* 2nd International Conference on Intelligent System Design and Engineering Application, 6-7 Jan. 2012, pp.716-719

[9] J.-J. Lin, S.-C. Huang "The Implementation of the Visitor Access Control System for the Senior Citizen Based on the LBP Face Recognition" in proc. Of *IEEE* International Conference on Fuzzy Theory and Its Applications (iFUZZY), 12-15 Nov. 2017 ,pp.1-6

[10] X. Peng, C. Van "Design of Access Control System Based on Stereo Face Recognition" in proc. Of IEEE International Conference on Computer Application and System Modeling (lCCASM 2010), 22-24 Oct. 2010,pp.557-561

[11] M. Sahani, C. Nanda, A. K. Sahu and B. Pattnaik "Web based online embedded door access control and home security system based on face recognition." in proc. Of IEEE International Conference on Circuit, Power and Computing Technologies, 19-20 March 2015 ,pp.1-6

[12] D. K. Sarker , N. I. Hossain , I. A. Jamil , "Design and implementation of smart attendance management system using multiple step authentication" in proc. Of IEEE International Workshop on Computational Intelligence (IWCI), 12-13 Dec. 2016, pp.91-95

[13] Vanlalhruaia , Y. K. Singh ; N. D. Singh "Binary face image recognition using logistic regression and neural network" in proc. Of IEEE International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS) , 1-2 Aug. 2017. ,pp. 3883-3888

[14] F. Mahmud , S. Afroge , M. A. Mamun , A. Matin "PCA and back-propagation neural network based face recognition system" in proc. Of IEEE18th International Conference on Computer and Information Technology (ICCIT), 21-23 Dec. 2015,pp. 582-587

[15] N. Xue, L.Liang, J. Zhang, X. Huang " An Access Control System for Intelligent Buildings"in proc of 9th EAI International Conference on Mobile Multimedia Communications, 18 - 20 June 2016 ,pp. 11-17

[16] "RDM 6300 datasheet " Seeed Studio, China.

Computer Vision and Smoke Sensor Based Fire Detection System

Nahid Chowdhury
Department of Computer Science
American International University – Bangladesh
Dhaka, Bangladesh
nahid.chowdhury1995@gmail.com

Dewan Ruhul Mushfiq
Department of Computer Science
American International University – Bangladesh
Dhaka, Bangladesh
mushfiq20@gmail.com

AZM Ehtesham Chowdhury
Department of Computer Science
American International University – Bangladesh
Dhaka, Bangladesh
ehtesham@aiub.edu

***Abstract*— Unwanted fire causes many disasters. Thus, fire detection has been an important phenomenon to save human life. In this paper, we proposed a hybrid model that incorporates computer vision and smoke sensor-based fire detection into a single system to devise a more accurate and smart fire detection system. Our vision-based fire detection model utilizes color and motion attributes of fire combustion. Rule-based color segmentation is performed for classifying fire color regions and a modified image difference technique is used for detecting the motion of fire regions in consecutive frames. The proposed dynamic threshold technique is used for mitigating the false positive rate of the vision-based model. After performing a vision-based analysis, the proposed algorithm uses the MQ-2 smoke sensor to detect environment smoke and gas caused by fire combustion. Our implemented system raises a fire alarm in the form of remote notification after analyzing vision and smoke sensor-based model data. Real life experimental data shows that our system can detect fire with 86.67% of accuracy. In addition, the proposed system can be incorporated in existing surveillance systems at an adequately low extra expense.**

Keywords—Fire flame detection, Fire color pixel segmentation, Fire region estimation, Dynamic threshold, Motion rate.

I. INTRODUCTION

Ensuring safety for the people and goods are one of the major challenges of 21st century. Among other causes one extremely dangerous cause that can ruin lives is fire. According to the National Fire Protection Association's (NFPA) report published on April 23, 2013 the United States fire department had to respond to an estimated annual average of 366,600 home structure fires from 2007-2011 [1]. And estimated 7 people died each day! One of the primary reasons for so many fire deaths is not immediately being aware of the fire hazard.

In the context of 21st century, surveillance cameras are extremely common. It is used in homes, offices, shops, industries and everywhere. Vision based fire detection that can be done with existing surveillance cameras can go a long way to reduce the amount of severe fire accidents. Even though only smoke based fire alarms are most popular but a lot of times they do not necessarily work the way they are intended. An integrated system that can check smoke as well as recognize the fire flame using existing surveillance cameras and immediately notify the authorized persons can reduce the fire loss immensely.

Fire has shown to have diverse, multispectral marks, a few of which have been used to devise diverse strategies for its detection. The vast majority of the techniques can be classified into smoke, heat, or radiation detection. According to Wikipedia "Fire detectors sense one or more of the products or phenomena resulting from fire, such as smoke, heat, infrared and/or ultraviolet light radiation, or gas" [2]. However, each of the fire detection technique is better suited for a distinct scenario and it is quite difficult to devise a solution which can detect all the phenomena caused by fire. Therefore we implemented a hybrid model which utilize computer vision as well as traditional smoke sensors for devising a more accurate and modern solution for fire detection in different environment. The objective of this work is to provide a fire detection model that utilize existing infrastructure but way more accurate and smarter than existing sensor based fire detection techniques.

The rest of the paper is organized as follows. The following section highlights some related works of sensor and computer vision based fire detection. Section III describes the proposed fire detection method as a whole. Section IV presents the architecture and the layout of our fire detector prototype. Section V describes the proposed vision based fire detection algorithm. It also discusses the methods in details. Section VI discuss about smoke sensor configuration in our prototype. Section VII shows the experimental data analysis in various real life environments and summarizes their result. Section VIII sheds light on why this system is beneficial for the detection of fire. Finally the last section IX concludes the paper.

II. RELATED WORK

Despite the growing interest in the field of automation, in the surveillance sector specially in fire detection, there is still not sufficient number of research conducted in the field of automated fire detection using real life scenarios. The authors [3] proposed a decent color analysis based fire detection method using HSI color space. They used image difference and color masking based approach for getting rid of spurious fire like region and also proposed a simple fire degree estimation technique. But the research only used color based distinctive feature of fire which is more error prone in real life scenarios.

Wireless sensor network based fire detection algorithm in video sequences is proposed by the authors [4]. They used motion and color feature of fire. For background extraction in image mixture of Gaussian is used and color based algorithm had been utilized for foreground extraction. The color detection algorithm is performed in RGB color space. Change map and blob's area of consecutive binary frames is computed for motion detection in fire region. But spurious fire color objects can also be detected as fire by the algorithm because the authors did not addresses the smoke based distinctive feature of fire.

A decent and architecturally complete fire alarm system is proposed by the authors [5] using raspberry pi, Arduino Uno and smoke sensors. The embedded system detects fire using presence of smoke in the air and generates a SMS notification with an image of surroundings. Thus it incorporates human intervention for confirming a notification to the fire fighters and mitigate false alarm. The shortcomings of this approach are it does not fully automate the fire alarm system and did not analyze the image captured by the system. A vision based model for analyzing surrounding image could be used for eliminating human intervention.

Therefore a complete fire detection system which incorporates vision and smoke sensor based fire detection in a single system is proposed for real life fire detection. This paper also provides a complete architectural design of a fire alarm system in great details.

III. Approach

Detection of fire using automated system is complex because fire has many distinctive features and all the feature cannot be detected using single type of fire detector. Therefore false positive rate of existing fire detection systems are very high.

Fire has many distinguishing features such as color, motion, shape, growth, heat, smoke and flame rays. In this paper we focused on features such as flame color, flame motion, generated smoke by fire and we hope to include additional features in our prototype in future work.

As mentioned earlier our algorithm is designed to work in static environment like CCTV surveillance areas therefore the first step of our algorithm is to dynamically define a threshold for the environment. After initializing the threshold the algorithm takes each frame from image source after T time and classifies fire like pixels using our color segmentation module.

In our prototype we have used IP camera as live image source and defined T = 2 sec. If the result of color segmentation find any fire like region then the algorithm proceed for motion detection of fire like region. After successful detection of motion in the fire like region the algorithm use the estimated amount of fire in the image and collect the data of smoke sensor for ensuring the presence of fire. If all the steps give positive result then our algorithm raise fire alarm in a form of remote notification.

IV. Architecture Of Fire Detector Prototype

"Fig 1" shows the components used in our prototype fire detector. "Fig 2" shows the architecture and full block diagram of the system. The structure of out fire detection

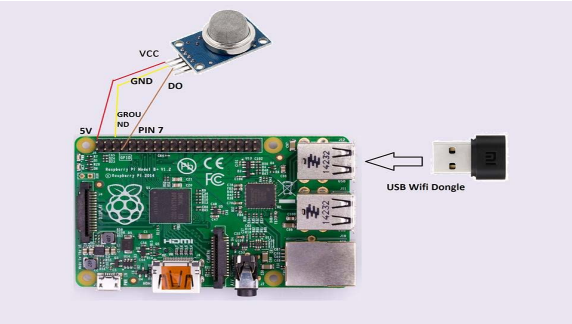


Fig. 1. Component diagram of the fire detector

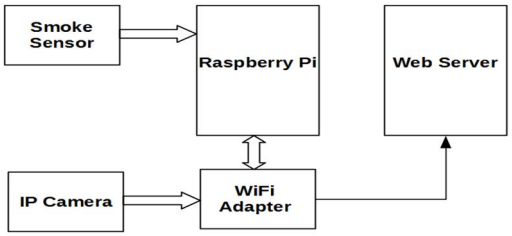


Fig. 2. Block diagram of the fire detector

system composed of five core components which are Raspberry Pi 2 Model B single board computer for data processing, MQ-2 smoke sensor for sensing gas or smoke, IP camera for capturing the image from the environment, WIFI adapter for accessing internet from Raspberry Pi and a web server for sending real time fire alarm. As a web server and backend API we have used IIS server and ASP.NET MVC as well.

V. Proposed Fire Flame Detection Method

The proposed fire flame detection method is based on RGB color space because the targeted hardware to run our algorithm is IOT boards like Raspberry PI or Arduino and for extracting fire colored pixels, RGB color space require less computation compared to other color space. The algorithm consists of three sub algorithms: fire like color pixels segmentation for extracting only fire color pixels from the image, fire like region estimation for estimating the amount of fire in the environment image and motion detection in fire like regions. But in real life scenarios we also need to define a threshold dynamically for each distinct environment for decreasing false positive rate. The following sections will describe the steps in more details

A. Fire like Color Pixels Segmentation

According to the literature review and based on our own experiments we noticed that fire has very distinctive color characteristics. It is a very powerful distinguishing feature of fire and almost every computer vision-based fire detection research used fire color feature extensively. The color of hydrocarbon flames which are the most common type of flame seen in nature belongs to red-yellow [6]. As mentioned earlier we used RGB color space for segmenting fire like color pixels from image. The authors [4, 7, 8, 9] also used RGB color space for aiding their fire color segmentation. For segmenting fire like pixels in RGB color space the two most commonly used rules shown in "(1)" and "(2)"

$$R(x,y) > R_{mean} \quad (1)$$

$$R(x,y) > G(x,y) > B(x,y) \quad (2)$$

Where x and y are the spatial coordinates of a frame at time *t*. In "(1)", *R(x, y)* is the intensity value of red channel at *x* and *y* location and R_{mean} "(1)" is the mean of red channel intensity values. In "(2)", *G(x, y)* and *B(x, y)* are the intensity values of green and blue channel at x and y location respectively. So, to classify a pixel as fire color pixel the value of red channel should be greater than the mean value of the red channel and the value of red channel supposed to be greater than the value

of green channel and the value of green channel supposed to be greater than the value of blue channel.

But in our color segmentation algorithm we modified the first rule "(1)" for decreasing false positive rate in the real-life environment. So, the modified version of Rule "(1)" shown in the following Rule "(3)".

$$R(x, y) > 200pt \quad (3)$$

Pixels that satisfy "(2): and "(3)" are considered to be fire like colored pixels in our color segmentation algorithm. The resulted image after performing the color segmentation can be represented as "(4)".

$$S(x, y) = \begin{cases} P(x, y), & if\ (2)\ and\ (3)\ is\ true \\ 0, & otherwise \end{cases} \quad (4)$$



Fig. 3. Fire image

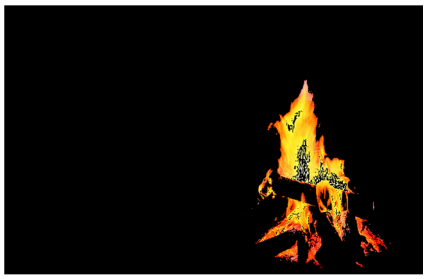


Fig. 4. Same fire image after color segmentation



Fig.. 5. Non fire image



Fig. 6. Same non fire image after color segmentation

As illustrated in "Fig. 3" and "Fig. 4", the fire areas are almost completely extracted. But only color segmentation is not sufficient for fire flame detection because "Fig. 3" and "Fig. 4" illustrates only the best-case scenario where there is no fire colored object in the background. Notice "Fig. 5" which is a non-fire image, but after color segmentation "Fig. 6" some areas are extracted because the pixels color belongs. to the range of hydrocarbon flame colors. So, the fire detection algorithms based on only color feature will be resulted in very high false positive rate.

B. Fire like Region Estimation

After color segmentation an approach is needed for estimating the burning degree of fire flames in the resulted image. A simple approach is proposed for roughly estimating the burning degree of fire flames which is similar to the approach of author [3]. In simple pseudo code, fire flame burning degree estimation can be expressed as,

```
counter = 0
for x = 0 to image.rows do
    for y = 0 to image.cols do
        if "(2)" and "(3)" equals true then
            counter++
        end if
    end for
end for
percentage = (counter / (image.rows * image.cols)) * 100
```

The percentage is used as the parameter of measuring the amount of fire in the image. Where 0 percentage means there is no fire found in the image and 100 percentage means the whole area is covered in fire.

C. Motion Detection

Motion detection algorithm runs on the result of color segmentation algorithm. The purpose of this sub algorithm is to remove static fire like colored object from the result and also gives a rough estimation of flame motion rate.

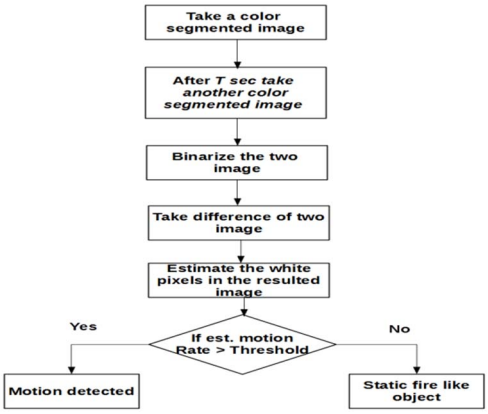


Fig. 7. Motion detection process

"Fig. 7" shows the motion detection algorithm takes two fire-color segmented images. The second frame should be taken after *T* sec of first frame. The next step is to binarize the images. The rule for binarization is if *P(x, y)* is greater than zero then the *P(x, y)* becomes white (255pt) otherwise the *P(x, y)* becomes black (0pt). After binarization the absolute

difference of two binarized frame is taken as resulted image. The result of motion detection algorithm is called motion rate which is nothing but the percentage of white pixels in the resulted image.

D. Dynamic Threshold

Despite of higher false positive rate of color-based fire detection algorithms, color-based approach is more popular because of its low false negative rate. Other vision-based fire detection algorithms that use color feature [3, 4, 7 - 13] also need to use other methods to decrease the higher false positive rate of color segmentation algorithm. Motion detection method can reduce the false positive rate of color segmentation but we also define a threshold at the very beginning of our fire flame detection algorithm. This method is also based on an assumption that when our algorithm first initializes the surrounding or the environment does not have any fire. On the basis of this assumption our algorithm takes a color segmented image of the environment and save its estimated fire pixel rate for using as threshold. So, whenever the algorithm estimates the fire pixel rate for any frame it deducts the threshold from its estimation. However, this technique is not applicable for dynamic environments where the background changes frequently. But this technique surprisingly decreases the false positive rate of the algorithm for static environment like surveillance camera's environment.

VI. SMOKE DETECTION

Smoke and gas are another important distinctive phenomenon caused by fire combustion. The gas produced during fire combustion highly dependent on combustion materials. Carbon monoxide is produced in quantity at most building fires because most organic materials contain carbon in their chemical structure. Materials that contain nitrogen, such as acrylic fiber, nylon, wool and urea-formaldehyde foam, could produce dangerous quantities of HCN in addition to CO [14]. Therefore, in our prototype we used MQ-2 smoke sensor which can sense smoke, LPG, Alcohol, Propane, Hydrogen, CO and even methane.

In our fire detection algorithm, the last step of verification of fire is smoke detection. After successful detection of fire flame by proposed fire flame detection algorithm, smoke detection take place. If smoke detection method gives positive result then our algorithm sends a strong fire alert to the web server otherwise it gives a mild fire alert because there is no smoke or gas in the air caused by fire combustion. The following section shows some experimental result of our fire detector in real life scenarios and a finally gives a verdict about its accuracy.



Fig. 8. MQ-2 smoke sensor

VII. EXPERIMENTAL RESULTS

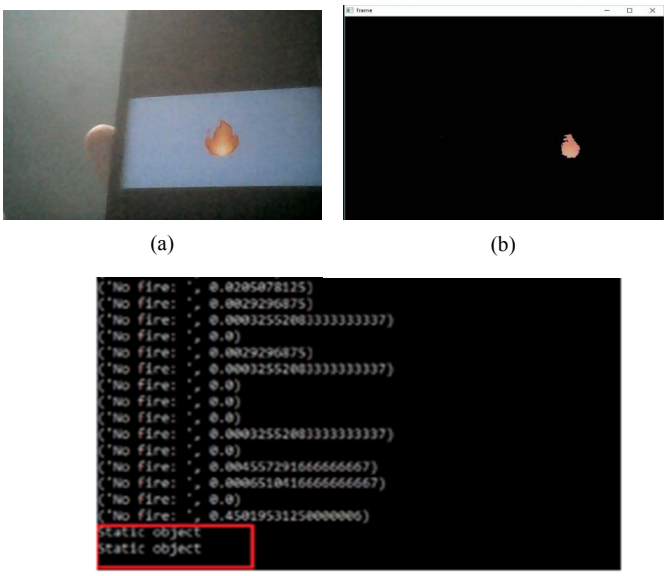
The proposed fire detection method was tested 60 times in variety of environment including indoor and outdoor. Some of the distinct scenarios mentioned here.



(a) (b)

Fig. 9. Test scenario with fire and smoke

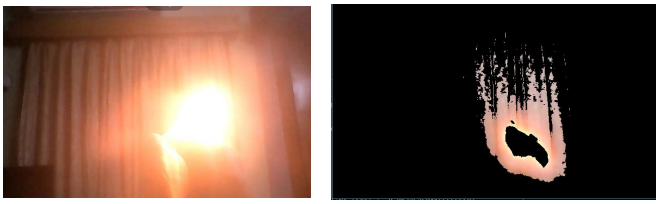
In "Fig. 9" (a) we have created some artificial fire and smoke. The fire detector successfully detected the flame area (b) and smoke and sent a strong fire alert to the web server.



(a) (b) (c)

Fig. 10. Static fire like object test

In "Fig. 10" (a) we tested our algorithm with static fire like object, at first it successfully segmented the fire color object as shown in (b) but due to absence of motion attribute of fire the fire detector classify it as a static fire like object.



(a) (b)

Fig. 11. Test scenario with fire

"Fig. 11" (a) shows an environment where we created artificial fire with match stick without artificial smoke. Our

fire detector successfully detected the fire as shown in (b) but due to absence of smoke in the environment the smoke detector does not sense any smoke therefore the fire detector sent a mild fire alert to our web server.

As mentioned earlier we conduct 60 tests in different environments and among 60 test cases 44 test cases contains actual fire and smoke and 16 test cases contains fake fire like objects without smoke. "Tab. I" shows the confusion matrix of experimental results where TP (true positive) means test cases where fire and smoke was present and algorithm detected it, FP (false positive) means test cases where fire wasn't supposed to be detected but the algorithm detected fire, FN (false negative) means test cases where fire was present but algorithm failed to detect it, and TN (true negative) means test cases where there was no fire and algorithm did not detect any.

TABLE I. CONFUSION MATRIX OF EXPERIMENTAL DATA

		Predicted Class	
		+	-
Actual class	+	41	3
	-	5	11

According to "Tab. I" the predictive accuracy of our proposed algorithm is about 86.67 %. The result is quite promising because all the test cases were done in real life environment. The author [4] achieved 92.8 % of correct classification result by their algorithm but all the test cases were on previously captured video sequences. Therefore, in real life scenarios 86.67 % of accuracy is quite sufficient in most of the cases.

VIII. ADVANTAGES OVER EXISTING SYSTEMS

The primary benefit of our proposed system over existing fire detection systems is the fact that whereas traditional fire detection system uses the smoke sensor to detect fire, our system not only detects the smoke but also can recognize fire combustion using computer vision. So, the detection is faster and more accurate than the traditional systems. The usability of our system is higher because of the real-time and two-step verification process. And the system is even more convenient since most offices and streets already have surveillance systems these days and our system can easily be integrated with those cameras that are already in place.

IX. CONCLUSION

In this paper, we proposed a hybrid fire detection system which incorporates computer vision and smoke sensor-based fire detection technique in a single system. Vision based algorithm use color and motion feature of fire combustion and MQ-2 smoke sensor detects smoke and gas produced by combustion. Our algorithm analyzes all the data produced by the vision and sensor-based model and raise alarm in form of remote Realtime notification. Our proposed system extends beyond fire detection and proposed a complete architectural reference of a fire alarm system. Our tests demonstrate that the proposed system distinguishes fire with high precision in real life CCTV surveillance environment. But the proposed system is not designed for the dynamic environment where the background changes frequently and the system is also not intelligent. So, in our future work we will endeavor to replace existing vision-based algorithm with deep learning based intelligent algorithm that will work equally in static and dynamic environment with better accuracy.

ACKNOWLEDGMENT

We would like to thank Mir Riyanul Islam, Faculty of Science and Information Technology, American International University Bangladesh for helping us optimizing our algorithm.

REFERENCES

[1] M. Ahrens, "Vehicle fire trends and patterns," Jun-2010. [Online]. Available: www.nfpa.org/News-and-Research/Fire-statistics-and-reports/Fire-statistics/Vehicle-fires/Vehicle-fire-trends-and-patterns. [Accessed: 15-Jul-2018].

[2] "Fire Detection," 23-Aug-2016. [Online]. Available: en.wikipedia.org/wiki/Fire_detection. [Accessed: 15-Jul-2018].

[3] W.-B. Horng and J.-W. Peng, "A Fast Image-Based Fire Flame Detection Method Using Color Analysis," *Tamkang Journal of Science and Engineering,* vol. 11, No. 3, pp. 273-285, 2008

[4] Y.-H. Kim, A. Kim, and H.-Y. Jeong, "RGB Color Model Based the Fire Detection Algorithm in Video Sequences on Wireless Sensor Network," *International Journal of Distributed Sensor Networks,* vol. 10, no. 4, p. 923609, 2014.

[5] M. S. B. Bahrudin, R. A. Kassim, and N. Buniyamin, "Development of Fire alarm system using Raspberry Pi and Arduino Uno," *2013 International Conference on Electrical, Electronics and System Engineering (ICEESE),* 2013

[6] C. E. Baukal, The John Zink combustion handbook, 1st ed. Boca Raton: CRC Press, 2001.

[7] X. Qi and J. Ebert, "A Computer Vision-Based Method for Fire Detection in Color Videos," Semantic Scholar, 01-Jan-1970. [Online]. Available: https://www.semanticscholar.org/paper/A-Computer-Vision-Based-Method-for-Fire-Detection-Qi Ebert/067f553df16650f7433a6e5b36207670cb16ab42. [Accessed: 12-Apr-2019].

[8] P. V. K. Borges and E. Izquierdo, "A Probabilistic Approach for Vision-Based Fire Detection in Videos," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 20, no. 5, pp. 721–731, 2010.

[9] Turgay Celik , Hasan Demirel, Huseyin Ozkaramanli, and Mustafa Uyguroglu, "Fire detection using statistical color model in video sequences," *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing,* Jun 2006.

[10] P. Gomes, P. Santana, and J. Barata, "A Vision-Based Approach to Fire Detection," *International Journal of Advanced Robotic Systems,* vol. 11, no. 9, p. 149, 2014.

[11] N. True, "Computer Vision Based Fire Detection." [Online]. Available: pdfs.semanticscholar.org/46f6/63bd1d32cdbec9b7ba206d471f94b9bbc03f.pdf. [Accessed: 22-Jul-2018].

[12] T. Çelik, H. Özkaramanlı, and H. Demirel, "Fire and Smoke Detection without Sensors: Image Processing Based Approach," *15th European Signal Processing Conference*, Sep. 2007.

[13] C.-B. Liu and N. Ahuja, "Vision based fire detection," *Proceedings of the 17th International Conference on Pattern Recognition*, 2004. ICPR 2004., 2004.

[14] "Toxic Gases and Vapours Produced at Fires," 01-Dec-1971. [Online]. Available:http://web.mit.edu/parmstr/Public/NRCan/CanBldgDigests/cbd144_e.html. [Accessed: 21-Jul-2018]

Implementation of Principal Component Analysis on Masked and Non-masked Face Recognition

Md. Sabbir Ejaz
Dept. of Computer Science and Engineering
Bangladesh Army University of Engineering and Technology
Natore, Bangladesh
cse.sabbir7@gmail.com

Md. Rabiul Islam
Dept. of Computer Science and Engineering
Rajshahi University of Engineering and Technology
Rajshahi, Bangladesh
rabiul_cse@yahoo.com

Md Sifatullah
Dept. of Computer Science and Engineering
Bangladesh Army University of Engineering and Technology
Natote, Bangladesh
sifatullahmsps@gmail.com

Ananya Sarker
Dept. of Computer Science and Engineering
Bangladesh Army University of Engineering and Technology
Natore, Bangladesh
ananya.ruet@gmail.com

***Abstract*— This paper represents an implementation of Principal Component Analysis (PCA) on masked and non-masked face recognition. Security is an essential term in our today's life. In various Biometric technology, face recognition is widely used to secure any system because it is better than any other traditional techniques like PIN, password, fingerprint etc. and most reliable to identify or verify a person efficiently. In recent years, face recognition is a very challenging task because of different occlusion or masks like the existence of sunglasses, scarves, hats and different types of make-up or disguise ingredients. The accuracy rate of face recognition is influenced by these types of masks. Many algorithms have been developed recently for non-masked face recognition which are widely used and give better performance. Still in the field of masked face recognition, few contributions has been done. Therefore, in this work a statistical procedure has been selected which is applied in non-masked face recognition and also apply in the masked face recognition technique. PCA is more effective and successful statistical technique and widely used. For this reason in this work, PCA algorithm has been chosen. Finally, a comparative study also done here for a better understanding.**

Keywords—Face recognition, PCA algorithm, Eigen face, Eigen value, Masked face.

I. INTRODUCTION

Face recognition is one of the most promising field of computer vision. Recognize a face and verifying a person automatically from images, known as face recognition system [1]. Face recognition plays an important role in our regular life. In a passport checking, ATM, credit card, voter verification, smart door, criminal or terrorist investigation and many other purposes face recognition is widely used to authenticate a person automatically and accurately. For those reasons, face recognition is the most popular than any other Biometric techniques.

In all automated personal identification system, face recognition has gained much attention as a unique Biometric recognition technique. For protecting the assets of many industries in the world are now trying to implement this authentication technique in their organizations. Throughout the world, many of the governments also interested to secure the public places such as railway stations, airports and bus stations etc. by using face recognition system. However for poor recognition rate, in real time recognizing remain unsuccessful. Recognition rate depends on quality of image. Recognition rate decreases for noisy and low quality image. That's why pre-processing is needed for better recognition rate. Some pre-processing techniques like cropping, resizing, sharpening, de-noising, normalizing, enhancing are used in face recognition process [2]. Recently many algorithms have been developed for reliable face recognition. Different techniques depend on different methods and they have different recognition accuracy. All of these algorithms, Principal Component Analysis (PCA) gives a better accuracy rate in normal or non-masked face recognition [3]-[5].

In present days masked face recognition is more important. Mainly terrorists and criminals covered their face with mask for disguise. Besides this, sunglass, hat, color festoon etc. also act like mask. Using different types of masks or occlusions the key features to identify a person is decreasing. Lower numbers of face features in the masked face cause difficulties than other normal face recognition technique [6]. Consequently the accuracy rate of recognition is decreasing. That's why masked face is being one of the majors concerned factor in the field of face recognition. Figure 1 shows some of masked faces.



Fig. 1. Human faces with different types of masks.

For face recognition, at first we need to detect the face from image. After detecting face, we can recognize the person. It is easy for human to detect a face but harder for a system to detect face. For this reason at first we need to train the system for detect face portion from images. Here we use Viola-Jones algorithm to detect face from an image. Viola-Jones algorithm

detect face using machine learning approach from any digital image [7]. After detecting face from input image, then we use PCA algorithm for feature extraction from image, which are used for training.

This work also presents a statistical difference of accuracy between masked face recognition and non-masked face recognition using the most successful algorithm Principal Component Analysis. All the result are shown in both graphically and tabulated form.

The paper is organized as follows: Section 2 presents our face detection and face recognition methodology. Section 2 describes about face recognition process. Section 3 signifies dataset. Section 4 shows experimental results and section 5 draws conclusions.

II. METHODOLOGY

Figure 2 demonstrate our proposed method to the problem of robust face recognition under mask and non-mask conditions. For face recognition at first detect the face portion from an image, find out the face features and finally train a system for face recognition. During the training phase, for detecting face portion we used Viola-Jones algorithm in our work and then PCA is used to efficiently represent the gallery images. In this way we obtain a PCA feature space which is called Eigen faces or ghost face. In the test case, a target face image (which can be masked or not) is given to be recognized. At first its PCA representation is computed. As a result, potential facial components are identified. Then, the Eigen faces are selected and used for recognition. The recognition is carried out by comparing the selected features from the target image against selected features from the corresponding template images. The nearest neighbor (NN) classifier distance is taken for the recognition.

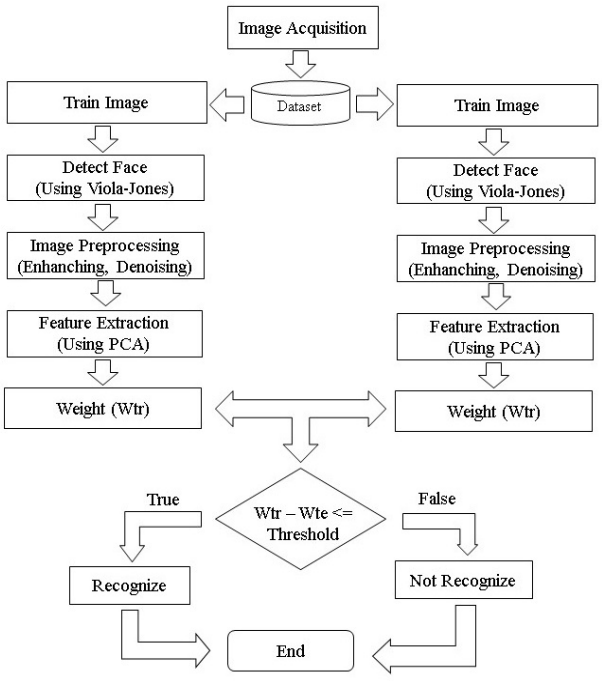


Fig. 2. Flow chart of face recognition process using PCA.

III. FACE RECOGNITION PROCESS

From a digital image or video sources identify or verify a person automatically by comparing facial features is called face recognition. Using various algorithms, features are extracted and then perform comparison between training image and test image for recognize a face.

A. Facial Image Acquisition

Image acquisition is the first stage of face recognition system. After obtaining the images, different methods of processing can be applied on the image. If the image has not been acquired correctly then the desired tasks are difficult to achieve. In this work we use faces from a database named ORL Database and some of our own captured images. We mix up our own masked and non-masked images to the database to enlarge the dataset so that our work is more reliable and efficient.

B. Detecting Face and Preprocessing

After acquiring the images, the first work is to segment out the face for further processing. We choose Viola-Jones face detection algorithm for detecting face region. Then implemented it in the MATLAB library to automatically segment the face region. Some of our face images quality are not much better as we needed. Up to expectation this face detector fails to detect face region in image. Faces which are heavily disguised or most of the important facial features are covered are subjects to fail detect face region. Every face image is normalized and apply image enhancement methods, so to achieve better results. This set of enhanced face images are therefore used to perform further experiments and evaluations.

C. Facial Feature Extraction using PCA

Principal Component Analysis (PCA) is an important statistical procedure and also defined as an orthogonal linear transformation. This algorithm emphasizes variation and brings out strong patterns in a dataset. It is used to minimize a big dataset to a small dataset still contains almost all the information as large dataset. PCA finds the data mean and principal components. It is popular as dimension reduction procedure. The technique is usually used for maximizing variance and seizing strong patterns of features in a dataset [8]. It was introduced by Karl Pearson in 1901. PCA is an effective statistical method.

For correlated data, Principal Component Analysis (PCA) works well. An image is also a highly correlated data. Hence extracting features from images, PCA performs better. Performing different operations on image matrix, it is transformed to a lower dimensional Eigen subspace. There after find out the covariance matrix from the lower dimension matrix. Relative variance between pixels in an image are represented by covariance matrix. Afterward Eigen vectors are calculated from this covariance matrix. Eigen vectors with the highest Eigen values are considered as the principal components.

Steps of Principal Component Analysis algorithm are mentioned below-

Step 1: Input data

Step 2: Calculate mean value of data

Step 3: Subtract the mean value from each input data

Step 4: Calculate Covariance matrix

Step 5: Calculate Eigen vectors and Eigen values

Step 6: Finding the greatest eigenvalue(s)

Step 7: Calculate Weight

D. Eigenvalue and Eigenvector

Eigenvalue and eigenvector comes from German word "Eigen" that means "Characteristic". In a square matrix eigenvalue is a scalar that represent by λ, which is a Greek letter. Eigenvector is represented by small letter *x* and it is a non-zero vector [9]. Here is an equation which is satisfied by all eigenvector and eigenvalue. Here, *S* is a given square matrix.

$$Sx = \lambda x \tag{1}$$

A huge calculation need in a large matrix and to solve this problem, eigenvalue and eigenvector method is used.

E. Eigen Face

Basically a set of eigenvector is called Eigen face which is used for face recognition of human. Eigen face extract the main feature from image and major Eigen face is selected for face recognition. In face recognition, Eigen face is used for represent the face image efficiently by principal component analysis. Images can be reconstructed as like as original image from a few number of weights and pictures [10]. Eigen face in a face recognition system must be less or equal to the total size of dataset. In figure 3 shows some of Eigen faces.



Fig. 3. Some Eigen faces.

F. Mathematical Analysis

For face recognition we need to train a dataset and then apply Principal Component Analysis on train dataset step by step. At first, we have to convert train dataset into face vector. Each 2D (*p* by *q*) train image is converted into 1D ($p \times q$ by 1) face vector. After converting all images, they are loaded which is called dataset. In this research work, total images K= 12 (face of our dataset). Here K depends on the number of train images and I is a variable which loads all images.

$$\mathrm{I} = \{\Gamma_1,\ \Gamma_2,\ \Gamma_3, \cdots\cdots\cdots, \Gamma_K\} \tag{2}$$

After load the dataset next procedure is to calculate average or mean face. Here average face is $\bar{X}$.

$$\bar{X} = \frac{1}{K}\sum_{n=1}^{n=K}\Gamma_K \tag{3}$$

Then, subtract the mean face from each image of the dataset which is called normalization. Here ϕ_n is a variable which stores the result of normalization. Then new matrix D is generated,

$$\phi_n = \Gamma_n - \bar{X}$$
$$\mathrm{D} = \{\phi_1,\ \phi_2,\ \phi_3, \cdots\cdots\cdots, \phi_K\} \tag{4}$$

After normalizing, we get a normalized face vector. Then we calculate covariance matrix of normalized vector.

$$C = DD^T\mathrm{s}$$
$$= \{D_{(p\times q)}D^T{}_{(q\times p)}\}_{(p\times p)} \tag{5}$$

Covariance matrix C is a large square matrix. For efficient and accurate calculation eigenvalue and eigenvector is calculated for reducing a huge face space vector. For recognition, we need to calculate weight which is compared with test image weight.

$$\omega_K = \mu^T(\Gamma_n - \bar{X}) \tag{6}$$

Here, ω = weight, μ = eigenvector, Γ = face and $\bar{X}$ = average face.

IV. DATASET AND EXPERIMENTAL ENVIRONMENT

In this work ORL face database [13] is used, which contains a set of faces taken between April 1992 and April 1994 at the Olivetti Research Laboratory in Cambridge, UK. This dataset contains 10 different images of 40 individuals. The files are in PGM format. Each image size is 92x112, 8-bit grey levels. For masked faces we created masked face images dataset using ORL face image and our own captured images. In this experiment, total 500 images used. 300 as a training image and 80-200 images used as test images on different test cases. Figure 4 shows some dataset images of our experiments. All the work done using MATLAB. Here use image processing toolbox of MATLAB 2017a version on windows platform.

V. EXPERIMENTAL RESULT

Proposed methodology is applied on our dataset images and found the resultant images. First step of face recognition is created a training dataset which contains pre-processed input images. After creating training set, average face is calculated for normalized training set. For normalize the dataset, subtracting an average face from each original face. Figure 5 shows the normalized faces of original face image. Then calculate the Eigen face which are shown in figure 6. Finally test some images for recognition. Here create some masked face images and also test them if it can be recognized

the person properly. Both masked and non-masked face testing images shown in figure 7.



Fig. 4. Some example images from training dataset.



Fig. 5. Some normalized faces



Fig. 6. Some Eigen faces



(a) (b)

Fig. 7. Testing (a) Masked face and (b) Non-Masked face

A statistical comparison between masked face recognition and non-masked face recognition using PCA algorithm shown in table-I. This table shows the recognition accuracy of both masked and non-masked face images. A graphical representation also shown in figure 8 for a better data visualization.

TABLE I. MASKED AND NON-MASKED FACE RECOGNITION ACCURACY

Test no.	Total train image	Total test image	Image types	Accuracy (%)
Test 01	300	80	Non-masked	96.25
	300	80	Masked	73.75
Test 02	300	120	Non-masked	95.83
	300	120	Masked	72.50
Test 03	300	160	Non-masked	95.05
	300	160	Masked	67.13
Test 04	300	200	Non-masked	95.62
	300	200	Masked	68.75

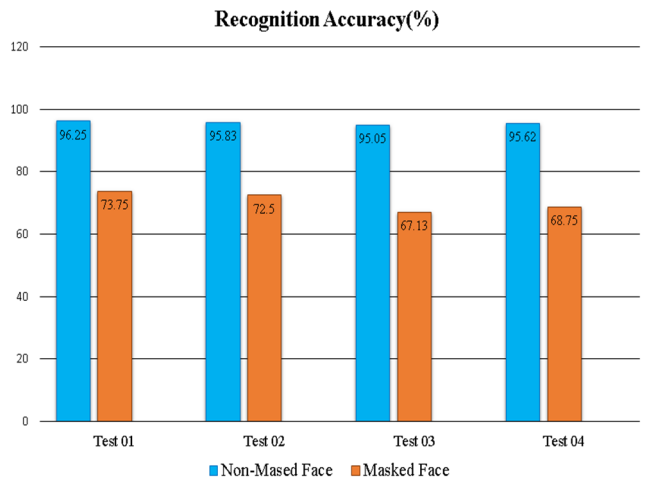


Fig. 8. Graphical representation of both masked and non-masked face recognition accuracy

From the above table, accuracy of masked face image recognition is on average 72% where non-masked face is on average 95%. So PCA gives poor recognition rate for masked face images rather than non-masked faces.

VI. CONCLUSION

This paper analyzed non-masked face recognition and masked face recognition accuracy using Principal Component Analysis (PCA) to recognize a person. It is proved that, a face without mask gives better recognition rate in PCA based face recognition system. But when a person is wearing mask, facial recognition gives poor recognition rate. It is found that extracting feature from a masked face is less than non-masked face. Because of missing features for wearing mask which decrease the recognition rate. Finally, we conclude that traditional statistical algorithm Principal Component Analysis (PCA) is better for normal face recognition but not for masked face recognition. So in the future, our concern to improve the accuracy of masked face recognition using other sophisticated machine learning methods.

VII. REFERENCES

[1] Ankit V Ponkia, Jitendra Chaudhari, "Face Recognition Using PCA Algorithm," Inventi Rapid: Image & Video Processing Journal, vol. 4, issue 1, pp. 519-524, Jan. 2012.

[2] Krishna Dharavath, Fazal Ahmed Talukdar and Rabul Hussain Laskar, "Improving Face Recognition Rate with Image Preprocessing," in Indian Journal of Science and Technology, vol. 7(8), pp. 1170–1175, Aug. 2014.

[3] Jian Yang, David Zhang, Alejandro F. Frangi, and Jing-yu Yang, "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 1, pp. 131-13, Jan. 2004.

[4] W. K. T. M. Gunarathne, S. K. Arachchi, W. G. C. W. Kumara, "Face Recognition by Independent Component Analysis To Verify Student Identity," International Research Symposium on Engineering Advancements (RSEA), Apr. 2015.

[5] Fatma Zohra Chelali, A. Djeradi and R. Djeradi, "Linear Discriminant Analysis for Face Recognition," in International Conference on Multimedia Computing and Systems, Apr. 2009, pp. 1-10.

[6] Wei Bu, Jiangjian Xiao, Chuanhong Zhou, Minmin Yang, Chengbin Peng, "A Cascade Framework for Masked Face Detection," in International Conference on CIS & RAM, Nov. 2017, pp. 458-462.

[7] Paul Viola, Michael Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," in Computer Vision and Pattern Recognition (CVPR), Dec. 2001, pp. 511-518.

[8] Ian T. Jolliffe1 and Jorge Cadima, "Principal component analysis: a review and recent developments," Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, vol. 374, Apr. 2016.

[9] A. S. Syed navaz, t. Dhevi sri, Pratap mazumder, " Face recognition using principal component analysis and neural network," in International Journal of Computer Networking, Wireless and Mobile Communications (IJCNWMC), vol. 3, pp. 245-256, Mar. 2013.

[10] Matthew A. Turk, Alex P. Pentland, "Face Recognition Using Eigen faces," International Journal of Emerging Technology and Advanced Engineering (IJETAE), vol. 2, issue 3, Feb. 2013.

[11] M. S. Ejaz, M. A. Hossain, A. Matin and M. T. Ahmed, "Performance Comparison of Partition Based Clustering Algorithms on Iris Image Preprocessing," in International Conference on Electrical & Electronic Engineering (ICEEE), Dec. 2017, pp. 1-4.

[12] A. Matin, F. Mahmud, T. Ahmed and M. S. Ejaz, "Weighted score level fusion of iris and face to identify an individual," International Conference on Electrical, Computer and Communication Engineering (ECCE), Feb. 2017, pp. 1-4.

[13] AT&T Laboratories Cambridge, "The ORL Database of Faces," [Online]. Available:https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html. [Accessed on: Aug. 12, 2018].

Brain Tumor Detection Using Convolutional Neural Network

Tonmoy Hossain[1*], Fairuz Shadmani Shishir[2@], Mohsena Ashraf[3#]
MD Abdullah Al Nasim[4&], Faisal Muhammad Shah[5$]

Dept of CSE, Ahsanullah University of Science and Technology, Dhaka, Bangladesh
tonmoyhossain.cse@ieee.org[*], fsshishir95@gmail.com[@], mohsena.ria20@gmail.com[#]
nasim.abdullah@ieee.org[&], faisal505@hotmail.com[$]

Abstract— **Brain Tumor segmentation is one of the most crucial and arduous tasks in the terrain of medical image processing as a human-assisted manual classification can result in inaccurate prediction and diagnosis. Moreover, it is an aggravating task when there is a large amount of data present to be assisted. Brain tumors have high diversity in appearance and there is a similarity between tumor and normal tissues and thus the extraction of tumor regions from images becomes unyielding. In this paper, we proposed a method to extract brain tumor from 2D Magnetic Resonance brain Images (MRI) by Fuzzy C-Means clustering algorithm which was followed by traditional classifiers and convolutional neural network. The experimental study was carried on a real-time dataset with diverse tumor sizes, locations, shapes, and different image intensities. In traditional classifier part, we applied six traditional classifiers namely Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Multilayer Perceptron (MLP), Logistic Regression, Naïve Bayes and Random Forest which was implemented in scikit-learn. Afterward, we moved on to Convolutional Neural Network (CNN) which is implemented using Keras and Tensorflow because it yields to a better performance than the traditional ones. In our work, CNN gained an accuracy of 97.87%, which is very compelling. The main aim of this paper is to distinguish between normal and abnormal pixels, based on texture based and statistical based features.**

Keywords— CNN, FCM, Medical Image, segmentation, SVM

I. INTRODUCTION

Medical imaging refers to a number of techniques that can be used as non-invasive methods of looking inside the body [1]. Medical image encompasses different image modalities and processes to image the human body for treatment and diagnostic purposes and hence plays a paramount and decisive role in taking actions for the betterment of the health of the people.

Image segmentation is a crucial and essential step in image processing which determines the success of a higher level of image processing [2]. The primary goal of image segmentation in medical image processing is mainly tumor or lesion detection, efficient machine vision and attaining satisfactory result for further diagnosis. Improving the sensitivity and specificity of tumor or lesion has become a core problem in medical images with the help of Computer Aided Diagnostic (CAD) systems.

According to [3], Brain and other nervous system cancer is the 10th leading cause of death, and the five-year survival rate for people with a cancerous brain is 34% for men and 36% for women. Moreover, the World Health Organization (WHO) states that around 400,000 people in the world are affected by the brain tumor and 120,000 people have died in the previous years [4]. Moreover, An estimated 86,970 new cases of primary malignant and non-malignant brain and other Central Nervous System (CNS) tumors are expected to be diagnosed in the United States in 2019 [5].

A brain tumor occurs when abnormal cells form within the brain [6]. There are two main types of tumors- Malignant and Benign. Malignant brain tumors originate in the brain, grows faster and aggressively invades the surrounding tissues. It can spread to other parts of the brain and affect the central nervous system. Cancerous tumors can be divided into primary tumors, which start within the brain, and secondary tumors, which have spread from elsewhere, are known as brain metastasis tumors. On the other hand, a benign brain tumor is a mass of cells that grow relatively slowly in the brain.

Hence, early detection of brain tumors can play an indispensable role in improving the treatment possibilities, and a higher gain of survival possibility can be accomplished. But manual segmentation of tumors or lesions is a time consuming, challenging and burdensome task as a large number of MRI images are generated in medical routine. MRI, also known as Magnetic Resonance Imaging is mostly used for a brain tumor or lesion detection. Brain tumor segmentation from MRI is one of the most crucial tasks in medical image processing as it generally involves a considerable amount of data. Moreover, the tumors can be ill-defined with soft tissue boundaries. So it is a very extensive task to obtain the accurate segmentation of tumors from the human brain.

In this paper, we proposed an efficient and skillful method which helps in the segmentation and detection of the brain tumor without any human assistance based on both traditional classifiers and Convolutional Neural Network.

II. LITERATURE REVIEW

One of the most challenging as well as demanding task is to segment the region of interest from an object and segmenting the tumor from an MRI Brain image is an ambitious one. Researchers around the world are working on this field to get the best-segmented ROI and various disparate approaches simulated from a distinct perspective. Nowadays Neural Network based segmentation gives prominent outcomes, and the flow of employing this model is augmenting day by day.

Devkota et al. [7] established the whole segmentation process based on Mathematical Morphological Operations and spatial FCM algorithm which improves the computation time, but the proposed solution has not been tested up to the evaluation stage and outcomes as- Detects cancer with 92%

and classifier has an accuracy of 86.6%. Yantao et al. [8] resembled Histogram based segmentation technique. Regarding the brain tumor segmentation task as a three-class (tumor including necrosis and tumor, edema and normal tissue) classification problem regarding two modalities FLAIR and T1. The abnormal regions were detected by using a region-based active contour model on FLAIR modality. The edema and tumor tissues were distinguished in the abnormal regions based on the contrast enhancement T1 modality by the k-means method and accomplished a Dice coefficient and sensitivity of 73.6% and 90.3% respectively.

Based on edge detection approaches, Badran et al. [9] adopted the canny edge detection model accumulated with Adaptive thresholding to extract the ROI. The dataset contained 102 images. Images were first preprocessed, then for two sets of a neural network, for the first set canny edge detection was applied, and for the second set, adaptive thresholding was applied. The segmented image is then represented by a level number and characteristics features are extracted by the Harris method. Then two neural network is employed, first for the detection of healthy or tumor containing the brain and the second one is for detecting tumor type. Depicting the outcomes and comparing these two models, the canny edge detection method showed better results in terms of accuracy. Pei et al. [10] proposed a technique which utilizes tumor growth patterns as novel features to improve texture based tumor segmentation in longitudinal MRI. Label maps are being used to obtain tumor growth modeling and predict cell density after extracting textures (e.g., fractal, and mBm) and intensity features. Performance of the model reflected as the Mean DSC with tumor cell density- LOO: 0.819302 and 3-Folder: 0.82122.

Dina et al. [11] introduced a model based on the Probabilistic Neural Network model related to Learning Vector Quantization. The model was evaluated on 64 MRI images, among which 18 MRI images were used as the test set, and the rest was used as a training set. The Gaussian filter smoothed the images. 79% of the processing time was reduced by the modified PNN method. A Probabilistic Neural Network based segmentation technique implemented by Othman et al. Principal Component Analysis (PCA) was used for feature extraction and also to reduce the large dimensionality of the data [12]. The MRI images are converted into matrices, and then Probabilistic Neural Network is used for classification. Finally, performance analysis is done. The training dataset contained 20 subjects, and the test dataset included 15 subjects. Based on the spread value, accuracy ranged from 73% to 100%.

Concentrating on Region based Fuzzy Clustering and deformable model, Rajendran et al. [13] accomplished 95.3% and 82.1% of ASM and Jaccard Index based on Enhanced Probabilistic Fuzzy C-Means model with some morphological operations. Zahra et al. [14] performed with LinkNet network for tumor segmentation. Initially, they used a single Linknet network and sent all training seven datasets to that network for segmentation. They did not consider the view angle of the images and introduced a method for CNN to automatically segment the most common types of a brain tumor which do not require preprocessing steps. Dice score of 0.73 is achieved for a single network, and 0.79 is obtained for multiple systems.

III. Proposed Methodology

In our proposed methodology, there are two distinct model for segmentation and detection of Brain tumor. First model segmented the tumor by FCM and classified by traditional machine learning algorithms and the second model focused on deep learning for tumor detection. Segmentation by FCM gives better result for noisy clustered data set [15]. Though it takes more execution time, it retains more information.

A. Proposed Methodology of Tumor Segmentation and Classification Using Traditional Classifiers

In our first prospective model, brain tumor segmentation and detection using machine learning algorithm had been done, and a comparison of the classifiers for our model is delineated. Our proposed Brain image segmentation system consists of seven stages: skull stripping, filtering and enhancement, segmentation by Fuzzy C Means algorithm, morphological operations, tumor contouring, feature extraction and classification by traditional classifiers. The results of our work accomplished satisfactory results. The main stages of our proposed model (Fig. 1) will be illustrated in the following sections.

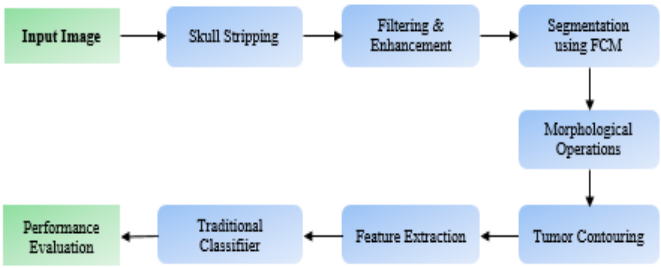


Fig. 1. Proposed methodology for classification using Traditional Classifiers

1) Skull Stripping: Skull stripping is a very important step in medical image processing because of the background of the MRI image not containing any useful information, and it only increases the processing time. In our work, we removed the skull portion from the MRI images in three steps. These three steps are:

a) Otsu Thresholding: For skull removal, at first we used Otsu's Thresholding method which automatically calculates the threshold value and segments the image into background and foreground. In this method, the threshold that is selected minimizes the intra-class variance, defined as a weighted sum of deviations of the two classes.

b) Connected Component Analysis: At the last stage of our skull stripping step, we used connected component analysis to extract only the brain region and as a consequence the skull part was removed.

2) Filtering and Enhancement: For better segmentation, we need to maximize the MRI image quality with minimized noise as brain MRI images are more sensitive to noise than any other medical image. Gaussian blur filter was used in our

work for Gaussian noise reduction existing in Brain MRI which prevailed the performance of the segmentation.

3) Segmentation using FCM: Fuzzy C-Means clustering algorithm was used for segmentation, which allows one piece of data to belong to two or more clusters. We got the fuzzy clustered segmented image at this stage, which ensured a better segmentation.

4) Morphological Operation: To segment the tumor, we only need the brain part rather than the skull part. For this, we applied morphological operations in our images. At first, erosion was done to separate weakly connected regions of the MRI image. After erosion, we will get multiple disconnected regions in our images. Dilation was applied afterwards.

5) Tumor Contouring: Tumor cluster extraction was done by an intensity based approach which is thresholding. The output of this image is the highlighted tumor area with a dark background.

6) Feature Extaction: Two types of features were extracted for classification. Texture-based features such as- Dissimilarity, Homogeneity, Energy, Correlation, ASM and Statistical based features including- Mean, Entropy, Centroid, Standard Deviation, Skewness, Kurtosis were extracted from the segmented MRI Images.

7) Traditional Classifiers: We used six traditional machine learning classifiers which are K-Nearest Neighbor, Logistic Regression, Multilayer Perceptron, Naïve Bayes, Random Forest, and Support Vector Machine to get the accuracy of tumor detection of our proposed model.

8) Evaluation Stage: Implementing other region-based segmentation methods and comparing it to our proposed segmentation technique, our model segments the ROI and segregates the tumor portion most accurately. An illustration of the whole process is depicted in Fig. 5. After segmentation and feature extraction from the tumor, we applied six classification techniques. Among them, we got the best result from SVM and obtained an accuracy of 92.42%.

B. Proposed Methodology Using CNN

Convolutional Neural Network is broadly used in the field of Medical image processing. Over the years lots of researchers tried to build a model which can detect the tumor more efficiently. We tried to come up with an exemplary which can accurately classify the tumor from 2D Brain MRI images. A fully-connected neural network can detect the tumor, but because of parameter sharing and sparsity of connection, we adopted CNN for our model.

A Five-Layer Convolutional Neural Network is introduced and implemented for tumor detection. The aggregated model consisting of seven stages including the hidden layers provides us with the most prominent result for the apprehension of the tumor. Following is the proposed methodology with a brief narration-

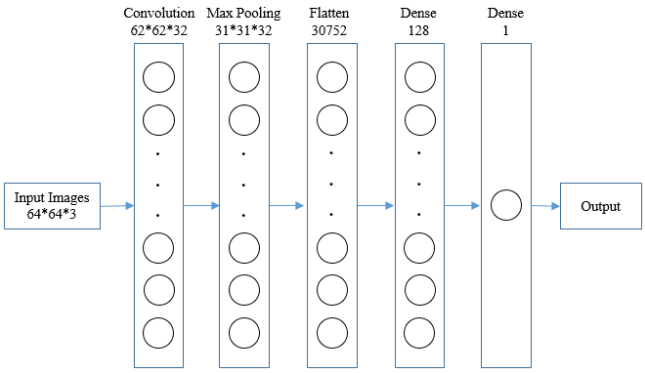


Fig. 2. Proposed Methodology for tumor detection using 5-Layer Convolutional Neural Network

Using convolutional layer as the beginner layer, an input shape of the MRI images is generated which is 64*64*3 converting all the images into a homogeneous dimension. After accumulating all the images in the same aspect, we created a convolutional kernel that is convoluted with the input layer — administering with 32 convolutional filters of size 3*3 each with the support of 3 channels tensors. ReLU is used as an activation function so that it's not corroborating with the output.

In this ConvNet architecture, progressively shorten the spatial size of the depiction for diminishing the chunk of parameters and computational time of the network. Working on the Brain MRI image can also cost the contamination of the overfitting and for this Max Pooling layer perfectly works for this perception. For spatial data which substantiate with our input image, we use MaxPooling2D for the model. This convolutional layer runs on 31*31*32 dimension. Because of divide the input images in both spatial dimensions, the pool size is (2, 2) which means a tuple of two integers by which to downscale by vertically and horizontally.

After the pooling layer, a pooled feature map is obtained. Flattening is one of the essential layers after the pooling because we've to transformed the whole matrix representing the input images into a single column vector and it's imperative for processing. It is then fed to the Neural Network for the processing.

Two fully connected layers were employed Dense-1 and Dense-2 represented the dense layer. The dense function is applied in Keras for the processing of the Neural Network, and the obtained vector is work as an input for this layer. There are 128 nodes in the hidden layer. Because the number of dimension or nodes proportional with the computing resources we need to fit our model we kept it as moderate as possible and for this perspective 128 nodes gives the most substantial result. ReLU is used as the activation function because of showing better convergence performance. After the first dense layer, the second fully connected layer was used as the final layer of the model. In this layer, we used sigmoid function as activation function where the total number of the node is one because we need to lower the uses of computing resources so that a more significant amount assuages the execution time. Though there is a chance of

hampering the learning in deep networks for using of the sigmoid as the activation function, we scale the sigmoid function, and the number of the nodes is much lesser and easy to handle for this deep network. In a summary, Fig. 3 shown the working flow of the proposed CNN model.

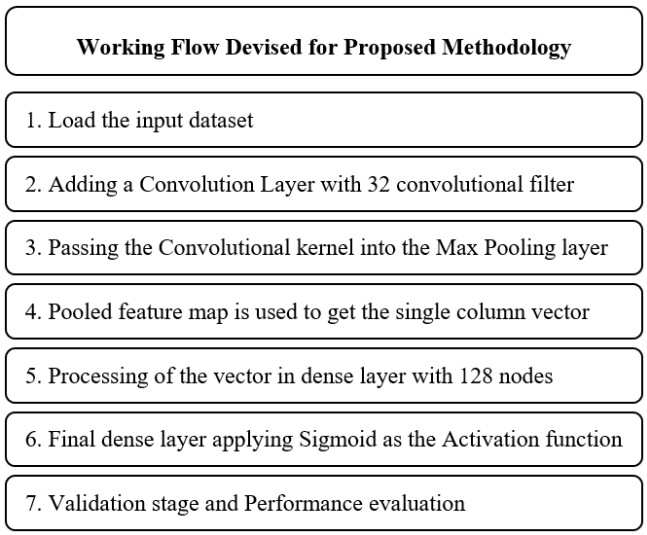


Fig. 3. Working flow of the proposed CNN Model.

Using Adam optimizer and binary cross-entropy as a loss function, we compiled the model and find the accuracy of detecting the tumor. An algorithm is depicted in Fig. 4 where we evaluated the performance of the model.

Algorithm 1: Evaluation process of CNN model

```
loadImage();
dataAugmentation();
splitData();
loadModel();
for each epoch in epochNumber do
    for each batch in batchSize do
        ŷ = model(features);
        loss = crossEntropy(y, ŷ);
        optimization(loss);
        accuracy();
        bestAccuracy = max(bestAccuracy, accuracy);
return
```

Fig. 4. Algorithm of the performance evaluation

All the hyper-parameters value are constituted in Table-I. Approximately 97.87% is achieved as the accuracy.

TABLE I. HYPERPARAMETER VALUE OF CNN MODEL

Stage	Hyper-parameter	Value
Initialization	bias	Zeros
	Weights	glorot_uniform
Training	Learning rate	0.001
	beta_1	0.9
	beta_2	0.999
	epsilon	None
	decay	0.0
	amsgrad	False
	epoch	10
	Batch_size	32
	steps_per_epoch	80

IV. EXPERIMENTAL RESULTS

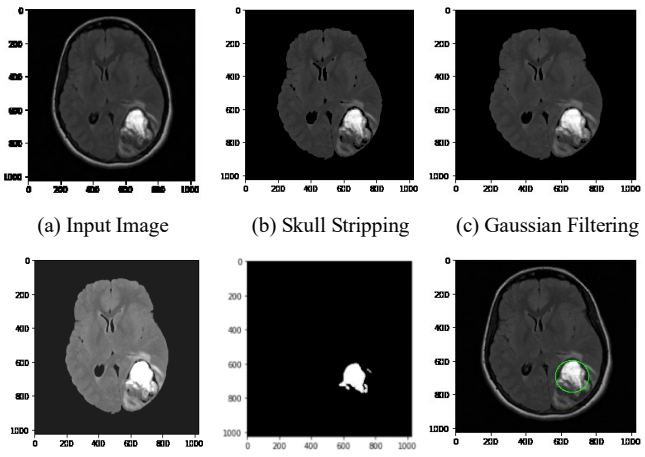
To justify our proposed model, steps of segmenting the tumor from 2D Brain MRI is illustrated (Fig. 5) and a comparative analysis of our proposed models of classification using machine learning, and deep learning is shown. We got 92.42% of accuracy using SVM and 97.87% of accuracy is achieved using CNN.

A. Experimental Dataset

For Performance Evaluation of our proposed model, we used the benchmark dataset in the field of Brain Tumor Segmentation, and that is BRATS dataset [16], consisting two classes'— class-0 and class-1 represents the Non-Tumor and Tumor MRI images. 187 and 30 MRI Images containing tumor and non-tumor respectively classified as class-1 and class-0. All the images are MRI images from different modalities like- T1, T2, and FLAIR. For traditional machine learning classifiers, we obtained the superlative result splitting the dataset by 70 to 30 in terms of training to testing images, and for CNN, we divided the dataset in both 70 to 30 and 80 to 20 formation and compared the outcomes.

B. Segmentation using Image processing techniques

Based on our proposed methodology, we segmented the tumor without loss of any subtle information. We removed the skull because for tumor segmentation the role of skull is approximately null and ambiguous in this process.



(a) Input Image (b) Skull Stripping (c) Gaussian Filtering

(d) Image Enhancement (e) Segmentation (f) Tumor Contouring

Fig. 5. Segmentation processes of an MRI

From the dataset, a 2D MRI was taken as an input image, Skull stripping technique is performed on the input image (Fig. 1b) followed by image enhancement (Fig. 1c) for understanding the features of the MRI properly. After that, Gaussian filter (Fig. 1d) is used for noise removal and finally simulating the FCM segmentation technique (Fig. 1e) followed by tumor contouring (Fig. 1f) to find out the ROI which is the tumor for Brain MRI. After the segmentation of the tumor, we classified the tumor based on different traditional Machine learning Algorithms.

C. Classification Using Machine Learning

Texture and Statistical based features are more popular for detecting the Region of Interest (ROI). Based on these features we can segregate the tumorous and non-tumorous MRI. We used texture and statistical based features for classification. Texture-based features like- Dissimilarity, Homogeneity, Energy, Correlation, ASM and Statistical based features including- Mean, Entropy, Centroid, Standard Deviation, Skewness, Kurtosis were extracted from the segmented Brain tumor. Further, we extracted the Area, Convex Hull Area and Diameter of the tumor. Extrapolating these features from the segmented MRI, we classified the image as the existence of normal and abnormal tissue. Table-II depicts the values of the features of some of the segmented MRI. After feature extraction, classification had been done. We adopt six classifiers which are- KNN, Logistic Regression, Multilayer Perception, Naïve Bayes, Random Forest, and SVM and achieved the best accuracy as the performance from SVM. Confusion Metrics' along with the performance of the classifiers is characterized in Table-III. The following factor evaluates the performance-

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$Sensitivity(recall) = \frac{TP}{TP+FN} \quad (2)$$

$$Speficity = \frac{TN}{TN+FP} \quad (3)$$

$$Precision\ (PPV) = \frac{TP}{TP+FP} \quad (4)$$

TABLE II. EXTRACTED FEATURES FROM SEGMENTED TUMOR

Image No	Contrast	Dissimilarity	Homogeneity	Energy	Correlation	ASM	Label
1	281.18	1.37	0.97	0.90	0.97	0.81	1
2	97.36	0.53	0.98	0.98	0.94	0.96	1
3	337.39	1.68	0.98	0.97	0.82	0.95	1
4	357.59	2.34	0.94	0.92	0.90	0.86	1
5	149.37	0.82	0.98	0.96	0.96	0.93	0
6	357.59	2.34	0.95	0.93	0.90	0.86	0

TABLE III. CONFUSION METRICS' OF THE CLASSIFIERS

Classifiers	Accuracy	Recall	Specificity	Precision	Dice Score	Jaccard Index
K-Nearest Neighnout	89.39	0.949	0.428	0.933	0. 941	0.889
Logistic Regression	87.88	0.949	0.286	0.918	0.933	0.875
Multilayer Perception	89.39	**1.000**	0	0.894	0.944	0.894
Naïve Bayes	78.79	0.797	**0.714**	**0.959**	0.870	0.770
Random Forest	89.39	0.983	0.167	0.903	0.943	0.892
SVM	**92.42**	0.983	0.428	0.935	**0.959**	**0.921**

From Table-III, we characterized that, among the six traditional machine learning classifiers, SVM gives the most prominent result and it is 92.42% in terms of accuracy. Though in terms of Precision and Specificity, Naïve Bayes gave the prominent outcome but the discrepancy with SVM was very subtle and also negligible considering the other performance metrics. From other performance metrics', it's also concluded that from SVM we obtained the pre-eminent result in terms of Jaccard Index, Dice Score, Precision, recall etc.

D. Classification Using CNN

The five-layer proposed methodology gives us the commendable result for the detection of the tumor. Convolution, Max Pooling, Flatten, and two dense layers are the proposed five layer CNN model. Data augmentation had been done before fitting the model as CNN is translation invariance. We evaluate the performance in two ways based on splitting the dataset. We accomplish 92.98% of accuracy for 70:30 splitting ratio where the training accuracy is 99.01%. Then at the second iteration, 80% of the images assigned for training and the rest of the images accredited for testing where we concluded 97.87% of accuracy and 98.47% of training accuracy. So our proposed model gives the best result when the division is 80:20. Table-IV represents the performance of the proposed methodology based on CNN.

We got 97.87% as accuracy which is remarkable in terms of using five-layer CNN. We analyzed with a different number of layers but the divergent of the outcomes were not very significant in terms of using this five-layer CNN model. Some of the aspects that we obtained when we increase the number of layers is- computation time, the complexity of the method batch size and steps per was immensely high. Further, we used 0.2 as the dropout value but did not commensurate the model as the accuracy flattened. As a result, this model provides the best accuracy without using dropout.

TABLE IV. PERFORMANCE OF THE PROPOSED CNN MODEL

No	Training Image	Testing Image	Splitting Ratio	Accuracy (%)
1	152	65	70 : 30	92.98
2	174	43	80 : 20	**97.87**

Fig. 6 depicts the training and validation accuracy of our model. It was calculated by the keras callbacks function. Operating with the different number of epochs we observed

the training and validation accuracy. We found that after 9 epochs model has the maximum accuracy for both training and validation.

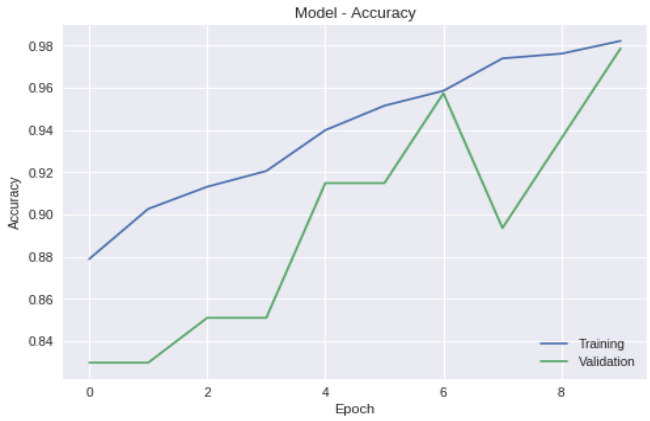


Fig. 6. Accuracy of the proposed CNN model.

E. Performance Comparison

Finally, we carried out a comparison between our proposed methodologies which are classification using traditional machine learning classifiers and CNN. We also compared our result with some other research articles which worked on the same dataset. In Seetha et al. [17], researchers got 83.0% accuracy using SVM based classification and 97.5% accuracy using CNN. Our proposed methodology provided an improved result for both machine learning and CNN based classification. Mariam et al. [18] got approximately 95% of dice co-efficient where we have 96% as the Dice score.

TABLE V. PERFORMANCE COMPARISON

Methodology	Accuracy (%)
Seetha et al [17]	97.5
Proposed CNN Model	**97.87**

CONCLUSION AND FUTURE WORK

Image segmentation plays a significant role in medical image processing as medical images have different diversities. For brain tumor segmentation, we used MRI and CT scan images. MRI is most vastly used for brain tumor segmentation and classification. In our work, we used Fuzzy C-Means clustering for tumor segmentation which can predict tumor cells accurately. The segmentation process was followed by classification using traditional classifiers and Convolutional Neural Network. In the traditional classifier part, we applied and compared the results of different traditional classifiers such as K-Nearest Neighbor, Logistic Regression, Multilayer Perceptron, Naïve Bayes, Random Forest, and Support Vector Machine. Among these traditional ones, SVM gave us the highest accuracy of 92.42%.

Further, for better results, we implemented CNN which brought in the accuracy 97.87% with a split ratio of 80:20 of 217 images, i.e. 80% of training images and 20% of testing images. In the future, we plan to work with 3D brain images, achieve more efficient brain tumor segmentation. Working with a larger dataset will be more challenging in this aspect, and we want to build a dataset emphasizing the abstract with respect to our country which will accelerate the scope of our work.

REFERENCES

[1] Kasban, Hany & El-bendary, Mohsen & Salama, Dina. (2015). "A Comparative Study of Medical Imaging Techniques". International Journal of Information Science and Intelligent System. 4. 37-58.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] D. Surya Prabha and J. Satheesh Kumar, "Performance Evaluation of Image Segmentation using Objective Methods", Indian Journal of Science and Technology, Vol 9(8), February 2016.

[3] Brain Tumor: Statistics, Cancer.Net Editorial Board, 11/2017 (Accessed on 17th January 2019)

[4] Kavitha Angamuthu Rajasekaran and Chellamuthu Chinna Gounder, Advanced Brain Tumour Segmentation from MRI Images, 2018.

[5] General Information About Adult Brain Tumors". *NCI*. 14 April 2014. Archived from the original on 5 July 2014. Retrieved 8 June 2014. (Accessed on 11th January 2019)

[6] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[7] B. Devkota, Abeer Alsadoon, P.W.C. Prasad, A. K. Singh, A. Elchouemi, "Image Segmentation for Early Stage Brain Tumor Detection using Mathematical Morphological Reconstruction," 6th International Conference on Smart Computing and Communications, ICSCC 2017, 7-8 December 2017, Kurukshetra, India.

[8] Song, Yantao & Ji, Zexuan & Sun, Quansen & Yuhui, Zheng. (2016). "A Novel Brain Tumor Segmentation from Multi-Modality MRI via A Level-Set-Based Model". Journal of Signal Processing Systems. 87. 10.1007/s11265-016-1188-4.

[9] Ehab F. Badran, Esraa Galal Mahmoud, Nadder Hamdy, "An Algorithm for Detecting Brain Tumors in MRI Images", 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, 2017.

[10] Pei L, Reza SMS, Li W, Davatzikos C, Iftekharuddin KM. "Improved brain tumor segmentation by utilizing tumor growth model in longitudinal brain MRI". *Proc SPIE Int Soc Opt Eng*. 2017.

[11] Dina Aboul Dahab, Samy S. A. Ghoniemy, Gamal M. Selim, "Automated Brain Tumor Detection and Identification using Image Processing and Probabilistic Neural Network Techniques", IJIPVC, Vol. 1, No. 2, pp. 1-8, 2012.

[12] Mohd Fauzi Othman, Mohd Ariffanan and Mohd Basri, "Probabilistic Neural Network for Brain Tumor Classification", 2nd International Conference on Intelligent Systems, Modelling and Simulation, 2011.

[13] A. Rajendran, R. Dhanasekaran, "Fuzzy Clustering and Deformable Model for Tumor Segmentation on MRI Brain Image: A Combined Approach," International Conference on Communication Technology and System Design 2011.

[14] Sobhaninia, Zahra & Rezaei, Safiyeh & Noroozi, Alireza & Ahmadi, Mehdi & Zarrabi, Hamidreza & Karimi, Nader & Emami, Ali & Samavi, Shadrokh. (2018). "Brain Tumor Segmentation Using Deep Learning by Type Specific Sorting of Images".

[15] Gupta, Gaurav and Vinay Singh. "Brain Tumor segmentation and classification using Fcm and support vector machine." (2017).

[16] Anam Mustaqeem, Ali Javed, Tehseen Fatima, "An Efficient Brain Tumor Detection Algorithm Using Watershed & Thresholding Based Segmentation", I.J. Image, Graphics and Signal Processing, 2012, 10, 34-39.

[17] Seetha, J & Selvakumar Raja, S. (2018). "Brain Tumor Classification Using Convolutional Neural Networks. Biomedical and Pharmacology Journal". 11. 1457-1461. 10.13005/bpj/1511.

[18] Mariam Saii, Zaid Kraitem, "Automatic Brain tumor detection in MRI using image processing techniques", Biomedical Statistics and Informatics, Vol. 2, No. 2, pp. 73-76, 2017.

BFAGC: A Bias-Free Adaptive Gamma Correction Method for Image Enhancement

Nazifa Akter, Sharmin Sultana, Rahat Hossain Faisal and Md. Mostafijur Rahman
Dept. of Computer Science and Engineering, University of Barishal, Barishal, Bangladesh
sharmin.cse1.bu@gmail.com, nazifa.cse1.bu@gmail.com, rhfaisal@gmail.com and mostafij.csebu@gmail.com

***Abstract*—One of the most interesting and visually appealing areas of image processing is image enhancement. In order to enhance the images, various methods have already been introduced. But, these existing enhancement methods are hardly able to enhance all types of images. To address this issue, Adaptive Gamma Correction (AGC) method has been introduced which classifies images into several classes based on the statistical information of the image and then applies an adaptive gamma correction method. However, AGC method can be affected by outlier pixels and fails to appropriately enhance the images. To solve this problem, we propose an improved version of Adaptive Gamma Correction (AGC) namely Bias-Free Adaptive Gamma Correction (BFAGC). The proposed BFAGC is rarely affected by the outlier pixels. The extensive experimental analysis has been performed to evaluate the efficacy of the proposed BFAGC method. Both the qualitative and quantitative evaluation metrics have shown that BFAGC produces comparatively better enhanced images than AGC and other existing state-of-the-art methods.**

***Keywords*—Contrast Enhancement, Image Classification, Image Enhancement, Adaptive, Gamma Correction.**

I. Introduction

Image enhancement is one of the well known issues in image processing field. Since, most of the images are affected by cloud, noise, low quality of the devices that capture images and the lack of operator expertise, image enhancement is needed to increase the visual quality and to remove unwanted noise of those images [1]. Besides this, image enhancement is necessary because of its expansive range of application in areas such as forensics [2], atmospheric sciences [3], astrophotography [4], oceanography [1], satellite image analysis [1], medical image processing [5], radar imaging [6], texture analysis and synthesis [1], remote sensing [7], person identification [8], face image analysis [9], machine learning [10], [11], [12], digital photography [13], surveillance and video processing applications [14].

In order to enhance the quality of an image, various contrast enhancement methods have been proposed. These methods involve Histogram Equalization (HE) [15], Brightness Preserving Bi-Histogram Equalization (BBHE) [16], Dualistic Sub-Image Histogram Equalization (DSIHE) [17], Recursively Separated and Weighted Histogram Equalization (RSWHE) [18], Contextual and Variational Contrast (CVC), Layered Difference Representation (LDR) [19], Gamma Correction with Weighting Distribution (AGCWD) [20] and Adaptive Gamma Correction (AGC) [21]. These methods are briefly described in the followings.

HE is one of the most fundamental image enhancement techniques because of its efficacy to give better contrast for all types of images [22]. It is a method that expanses the contrast by redistributing the gray level values equally [23]. Although HE improves the contrast and brightness of an image, it does not always give favorable result. It produces a significant amount of artifacts due to the mean-shift problem [23]. To reduce this problem, BBHE [16] and DSIHE [17] methods are introduced. BBHE [16] method solves the over-enhancement problem by dividing the input image into two sub-images based on the mean of the image. Then, HE is separately applied on both sides of the images. Afterward, two parts are merged to get the final output. For symmetric distribution of input image around its mean, BBHE provides better result. But it may fail for non-symmetric distribution [23]. Considering this issue, DSIHE [17] method is introduced which divides the histogram into two sub-images on the basis of median [17]. Then, each sub-image is equalized like BBHE. Though DSIHE method neglects the significant mean shift problem in some cases, it does not able to preserve mean brightness. It fails to improve some extents of image or to create artifacts on image [23]. Then, another method namely RSWHE [18] is introduced by combining BBHE [16] and DSIHE [17] methods, which improves the contrast and conserves the brightness of an image. It partitions the histogram into two parts and applies weights. At the last, it performs histogram equalization on each part of the weighted histograms. This method has the same time complexity but expands DSIHE by including multi-equalization to reduce the generation of artifacts.

Except these techniques, some other techniques have already been introduced. Among these, CVC [24] method uses inter pixel contextual information. This information is taken, and then enhancement is performed on its neighboring pixels which creates a 2D histogram [24]. For creating a better visual quality output image, this method is an effective one. But, the computational complexity of this method becomes higher. Another method, LDR [19] divides an image into multiple layers. Then, it derives a transformation function for each layer [19]. After performing those transformation functions, it merges all the layers to get final output. Another method, combining traditional gamma correction and histogram equalization methods are AGCWD [20]. This method first analyzes the histogram of the image then weighting distribution is used and then the Gamma Correction method is applied to the image to get the enhanced output image [20]. Although

AGCWD enhances the contrast, it may fail to provide desired output when an input image lacks bright pixels [23]. Since, the above methods rarely enhance the contrast of all types of images. To enhance both the contrast and brightness of an image, AGC method is introduced. This method classifies images into high contrast (Bright and Dark images) and low contrast (Bright and Dark images) class. Then, different transformation functions are applied for each class. Here, all the parameters are automatically set based on image information. But, sometimes this method is affected by outlier pixels and fails to enhance the image properly due to using arithmetic mean of the image. Therefore, in this paper we propose a modified adaptive gamma correction technique namely Bias-Free Adaptive Gamma Correction (BFAGC). The proposed BFAGC reduces the effect of outlier pixels using harmonic mean instead of arithmetic mean of the image. The novelties of this proposed method are as follows

- We introduce a modified adaptive gamma correction technique namely Bias-Free Adaptive Gamma Correction (BFAGC).
- We have used harmonic mean instead of arithmetic mean.
- We have performed extensive experimental analysis on various images.

The rest of the paper is organized as follows. Section II provides our proposed method. Experimental analysis is discussed in Section III. Finally, Section IV concludes the paper with future work.

II. Proposed Method

In this paper, we propose an improved version of Adaptive Gamma Correction (AGC) method. Here, we use harmonic mean instead of arithmetic mean to mitigate the problem faced by AGC method. The overall procedure of BFAGC is shown in Algorithm 1.

The proposed BFAGC enhances both gray and color images. In case of color images, direct RGB color processing cannot keep up the actual color. Therefore, it creates an unwanted effect on the output image. Due to this problem, we need to color space conversion. In image processing, available color spaces are HSV, HSI, Lab, YUV, Luma etc. In our method, we use HSV color space. It divides the color and brightness information of an image into hue (H), saturation (S), and value (V) channels. We use this model because it has the ability to separate color information from the brightness (or lightness) information and it has a good capability to represent a color for human perception. After transforming RGB image into HSV, V channel is considered to enhance the image as it contains the value of brightness.

At the beginning of the Bias-Free AGC (BFAGC) method images are classified into low contrast and high contrast class using Eq. (1).

$$g(I) = \begin{cases} \varrho_1, & \text{D}\leq 1/\tau \\ \varrho_2, & \text{otherwise} \end{cases} \tag{1}$$

where, D = diff((μ_h+2σ), (μ_h-2σ)) and τ is a parameter used for defining the contrast of an image. σ and μ_h are the standard deviation and harmonic mean of the image intensity respectively. The harmonic mean is calculated by Eq. (2).

$$\mu_h = \frac{n}{\sum_i \frac{1}{x_i}} \tag{2}$$

where, n and x_i represent the total number of pixels and each pixel intensity respectively. μ_h is the harmonic mean of the image.

The aforementioned two groups of images are enhanced using different transformation functions. These two transformation functions are illustrated in the following subsections.

Algorithm 1 : BFAGC Algorithm

Data: Input image, I_{in}
Result: Output image, I_{out}
1: **if** I_{in} is colored **then**
2: $I_{hsv} \leftarrow$ Convert I_{in} from RGB to HSV
$R \leftarrow$ V channel of I_{hsv}
3: **else**
4: $R \leftarrow I_{in}$
5: **end if**
6: **if** R_{in} is low contrast image **then**
7: Set γ using Eq. (4).
8: **else**
9: Set γ using Eq. (9).
10: **end if**
11: Define the value of c using Eq. (5).
12: Compute R_{out} using Eq. (3).
13: **if** I_{in} is colored **then**
14: $R_{out} \leftarrow$ Convert HSV to RGB image
15: **else**
16: $I_{out} \leftarrow R_{out}$
17: **end if**

A. Low-contrast image enhancement (ϱ_1)

For the transformation of image, we use traditional gamma correction where γ controls the slope of the transformation function. The equation of traditional gamma correction is given by Eq. (3).

$$I_{out} = cI_{in}^{\gamma} \tag{3}$$

In our method, for low-contrast images we choose the value of γ by using Eq. (4) which is same as AGC

$$\gamma = -\log_2(\sigma) \tag{4}$$

The value of σ is small for low contrast class that causes large γ value that result in better contrast. Generally, for illuminating or darkening the output image intensities c is used. But, on the basis of the qualities of the respective image this method uses several values of c for several images using Eq. (5).

$$c = \frac{1}{1 + Heaviside(0.5 - \mu_h) \times (k-1)} \tag{5}$$

Where, k is defined by Eq. (6).

$$k = I_{in}^{\gamma} + (1 - I_{in}^{\gamma}) \times \mu_h^{\gamma} \tag{6}$$

and the Heaviside function is defined by Eq. (7).

$$Heaviside(x) = \begin{cases} 0, & \text{x} \leq 0 \\ 1, & \text{x} > 0 \end{cases} \tag{7}$$

The reason behind using harmonic mean instead of arithmetic mean, arithmetic mean is always greater than harmonic mean [23]. So, in Eq. (6), arithmetic mean will generate large value of k which will decrease the value of c, causing lower brightness of an image. On the other side, the value of harmonic mean generates lower value of k and increase the value of c which causes higher brightness of an image.

Then, BFAGC method divides low contrast class into two parts dark images and bright images. The transformation function should exhibit the bright intensities over a wider range of darker intensities in order to increase the contrast of this type of images. BFAGC method achieved this by choosing the γ value that is greater than 1. For transformation of bright images in Eq. (3), c becomes 1 and the transformation function becomes

$$I_{out} = I_{in}^{\gamma} \tag{8}$$

For dark images the transformation function needs to be exhibited the dark intensities to the higher intensities in order to increase the contrast and brightness of this type of images. For enhancement of this type of images this method uses Eq. (3).
By the aforementioned way, our method is able to handle bright and dark images by choosing the values of γ and c in low contrast class.

B. High or moderate-contrast image enhancement (ϱ_2)

When the intensities of an image are scattered over the available dynamic range then the image falls into high contrast class. The more important concern of this type of images is brightness adjustment than contrast enhancement. In the case of high contrast class, I_{out} and c are calculated like Eqs. (3) and (5). γ is calculated using Eq. (9).

$$\gamma = e^{(1-(\mu_h-\sigma))/2} \tag{9}$$

For this class of images γ falls within a small range around 1. This method also divides the high contrast class into two types bright images and dark images. For dark images there is comparatively greater harmonic mean and not much changes in the intensities. And for enhancement of bright images I_{out}, c and γ are calculated using Eqs. (3), (5) and (9) respectively. Here, the main target is to preserve the image quality.

C. Why BFAGC is Bias Free?

An image is biased when there is inhomogeneous classification of the pixels. This make an image to show certain character and fails to provide better contrast enhanced image. Hence, a proper classification of images is necessary to enhance contrast without incurring artifacts. The main purpose of image classification in BFAGC is to categorize the image in such a way so that half of the pixels fall in one side and the remaining half is in another side for providing maximum contrast for an image [23]. AGC method fails to categorize the images effectively in some cases due to using arithmetic mean because it can be affected by outlier pixels. Therefore, BFAGC uses harmonic mean to classify images properly. The reason for choosing harmonic mean instead of arithmetic mean can be described based on TABLE I.

TABLE I: Pixel values based on arithmetic mean and harmonic mean

Image Name	Hill	Advertise	Boat	Girl
Arithmetic Mean	32.31	59.46	43.92	24.75
Pixels Above Arithmetic Mean	126646	87528	81251	84731
Pixels Below Arithmetic Mean	556354	295712	145899	128469
Harmonic Mean	0.69	5.84	23.46	6.96
Pixels Above Harmonic Mean	315540	101017	91422	88400
Pixels Below Harmonic Mean	367460	282223	135728	124800

In TABLE I, we have seen that Hill image has arithmetic mean 32.31. The pixels above arithmetic mean and the pixels below arithmetic mean are 126646 and 556354 respectively that means 18% pixels are stayed in one side and other 82% pixels fall in another side. For maximum contrast enhancement we needed 50% pixels in one side and the remaining is another side. But, using arithmetic mean it fails to provide the desired results and in some of the cases it is affected by outlier pixels. So, we use harmonic mean instead of arithmetic mean in order to mitigate the problem. As, for example in TABLE I for Hill image harmonic mean is 0.69. Hence, the pixels above harmonic mean and the pixels below harmonic mean are 315540, 367460 respectively that means 46% pixels are stayed in one side and 54% pixels in another side. Thus, BFAGC method able to classify images properly. Hence, this method becomes bias- free and gives a balanced output.

III. Experimental Analysis

In this section, we have chosen some standard images to appreciate the effectiveness of proposed method. The performance of this method compare with some of the other existing techniques such as HE [15], CVC [24], LDR [19], RSWHE [18], AGCWD [20] and AGC [21]. Here, the experimental section is divided into two parts qualitative and quantitative measurement. For the purpose of experiment, we have used MATLAB 2012, Windows 8.1 operating system.

A. Qualitative Measurement

In this section, for qualitative measurement we took some images that were used in AGC method and some images from open source. After analyzing qualitative measurement, we found that BFAGC performs better result in most of the cases than other methods.

For "Advertise" image in Fig. 1, HE method over enhances the image by reducing natural look. HE transfers the original mean value from 59.46 to 130.08 which shows over enhancement problem. CVC and LDR are not able to enhance the images which can be seen from the Fig. 1. The mean value of CVC and LDR are 72.56 and 68.16.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 1: "Advertise" image enhancement (a) Original (b) HE (c) CVC (d) LDR (e) RSWHE (f) AGCWD (g) AGC (h) Proposed



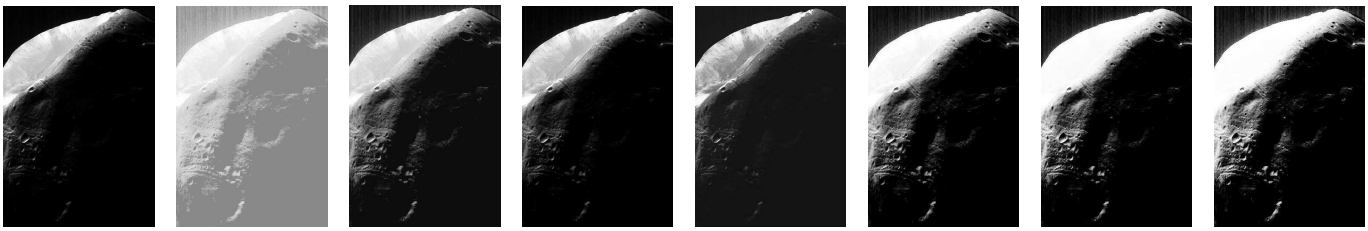
(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 2: "Girl" image enhancement (a) Original (b) HE (c) CVC (d) LDR (e) RSWHE (f) AGCWD (g) AGC (h) Proposed



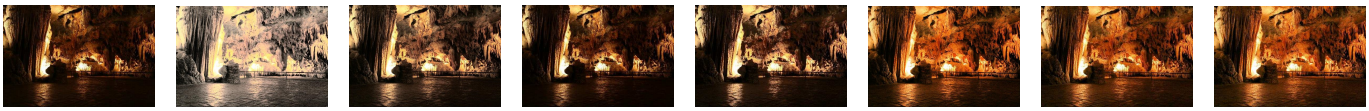
(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 3: "Fountain" image enhancement (a) Original (b) HE (c) CVC (d) LDR (e) RSWHE (f) AGCWD (g) AGC (h) Proposed



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 4: "Hill" image enhancement (a) Original (b) HE (c) CVC (d) LDR (e) RSWHE (f) AGCWD (g) AGC (h) Proposed



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 5: "Cave" image enhancement (a) Original (b) HE (c) CVC (d) LDR (e) RSWHE (f) AGCWD (g) AGC (h) Proposed



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 6: "Nature" image enhancement (a) Original (b) HE (c) CVC (d) LDR (e) RSWHE (f) AGCWD (g) AGC (h) Proposed

In RSWHE method the output is not satisfiable and the sky looks cloudy because the mean value reduced to 57.59. AGCWD unable to increase the overall brightness of this image. It transfers moderate bright pixels to the highest bright one. Although AGC method gives a better result than the aforementioned methods, our method gives brighter output with maximum identification of local details compared to other methods including AGC. Mean brightness of our proposed method for this image is 87.86 which shows overall brightness to the whole image.

For "Girl" and "Fountain" images in Figs. 2 and 3, we can see that HE method over enhances both images. The result produced by HE method does not look natural. It transfers most of the bright pixels to the highest bright region (original mean=24.73, 29.32 and transformed mean=136.30, 128.90) and causes over enhancement problem. CVC, LDR and RSWHE methods fail to improve the visual quality of these two images. We can see that AGCWD gives a comparatively better result than other existing methods but there is a lack of sufficient brightness. In the case of "Girl" image which is low contrast, it is needed to increase the brightness in such a way that the necklace can be easily identified. AGC method makes it possible to visualize the thin necklace but our method increases the brightness than AGC in such a way the necklace can be clearly identified and keeps the balance of the luminance of the hair region. The same thing is true for the Fountain image. Our proposed method brings out details of all object by increasing the brightness of the image (original mean=24.75, 29.32 and transformed mean= 96.55, 79.06).

"Hill" image in Fig. 4, is a dark low contrast image with mean brightness of 32.31. We can see that HE method over enhances the image by transferring the mean to 164.96. LDR fails to improve the result as there is not too much different between the original mean and the transformed mean (34.43). CVC method partially enhances the image. Unfortunately, RSWHE and AGCWD fail to improve the image quality properly and the output images are still dark because the enhancement of AGCWD depends on the maximum available intensity of the image and in this case the maximum intensity of the original image is small. AGC method enhances the image but there is a lack of proper brightness. The lower the value of k in Eq. (6), the higher the value of c in Eq. (5) is, causing an increase of brightness in Eq. (3). As the value of the harmonic mean is always less than the arithmetic mean. So, BFAGC increases the value of I_{out} from Eq. (3) giving overall brightness of the image. It gives mean brightness 58.76 which is balanced between dark and bright regions whereas the mean brightness value of AGC is 52.42.

"Cave" and "Nature" images in Figs. 5 and 6 are dark high contrast images where minimum pixels are 0, 2 and maximum pixels 255, 255 respectively. The mean value of these two images produced by HE is 125.48 and 125.69 which shows that most of the pixels are transferred to highest bright pixels incurring artifacts. CVC, LDR, and RSWHE methods are unable to improve the quality of the images. The result produced by AGCWD method provides limited brightness. Though AGC method enhances the image but our method gives a better result than other existing methods. The proposed method gives average brightness to all pixels (original mean=36.73, 122.42 and transformed mean of proposed method= 63.42 and 137.86).

TABLE II: Discrete Entropy (DE)

Image Name	Fountain	Nature	Advertise	Hill	Cave	Girl
Original	6.16	7.51	6.05	4.03	6.26	5.86
HE	6.85	7.84	7.15	3.66	6.88	6.65
CVC	6.92	7.48	6.82	3.92	0.82	6.86
LDR	6.94	7.59	6.64	4.01	7.06	6.78
RSWHE	5.93	7.31	5.67	3.64	6.20	5.60
AGCWD	7.15	7.39	6.76	3.74	6.79	6.64
AGC	7.23	7.49	6.43	3.64	6.74	7.06
Proposed	7.30	7.52	6.63	3.51	7.06	7.00

TABLE III: Root-Mean-Square (RMS)

Image Name	Fountain	Nature	Advertise	Hill	Cave	Girl
HE	0.30	0.30	0.29	0.15	0.32	0.28
CVC	0.25	0.33	0.34	0.28	0.001	0.31
LDR	0.24	0.33	0.34	0.27	0.27	0.27
RSWHE	0.27	0.28	0.23	0.22	0.26	0.17
AGCWD	0.27	0.35	0.36	0.34	0.28	0.20
AGC	0.25	0.34	0.35	0.35	0.26	0.32
Proposed	0.26	0.35	0.35	0.36	0.27	0.33

B. Quantitative Measurement

We can not predict which method works well only considering the visual quality of an image. Hence, we need a mathematical evaluation to identify weather the quality of an image is enhanced or not. Here, we give a concrete evaluation of our method using three quality metrics Mean brightness, Discrete entropy and Root-Mean-Square (RMS). They are described in the followings.

1) **Discrete Entropy (DE):** Entropy can be defined by the measurement of the uncertainty of a random variable. If the random variable is more, an image has more entropy [21]. In image processing, large entropy means better-contrast of the image. Table II provides the Discrete entropy values performed by different methods. From the table, We observe that in most of the cases BFAGC creates larger entropy than other methods. For example, if we take a look at "Fountain",

TABLE IV: Mean Brightness Value

Image Name	Fountain	Nature	Advertise	Hill	Cave	Girl
Original	29.32	122.42	59.46	32.31	36.73	24.75
HE	128.90	125.69	131.08	164.96	125.48	136.30
CVC	50.00	142.11	72.56	55.53	61.42	73.17
LDR	50.14	126.42	68.16	34.43	62.69	65.49
RSWHE	51.93	123.21	57.59	44.05	51.23	40.25
AGCWD	86.39	159.49	92.45	52.60	61.26	53.82
AGC	71.64	114.48	78.87	52.42	58.84	93.65
Proposed	79.06	137.86	87.86	58.76	63.42	96.55

"Cave" and "Girl" images, we can see that our proposed method gives highest value such as 7.30, 7.06 and 7.00 than other methods. Hence, it can be concluded that BFAGC provides better result than other existing image enhancement methods.

2) **Root-Mean-Square (RMS):** RMS measure the standard deviation of pixel luminance of an image which is divided by the mean luminance of the whole image. The higher the value of the RMS the better the contrast of the image [19]. Table III provides the RMS values performed by different methods. From the table, We have seen that in some cases BFAGC creates same RMS value as AGC and in some cases it provides larger RMS value than other methods. Hence, it can be concluded that some of the cases BFAGC gives the same contrast as AGC and some of the cases it provides a better contrast than other methods.

3) **Mean Brightness Value:** Mean is the most basic of all statistical measure [25]. In image processing, a higher value of mean identifies, it improves the brightness of the image. Table IV provides the mean values of different methods. From the table, We also observe that in most of the cases BFAGC creates higher mean value than other methods. Finally, it concludes that BFAGC increases the brightness than other existing image enhancement methods.

IV. Conclusion

In this paper, an improved version of adaptive gamma correction technique has been proposed to enhance the contrast and brightness of the image. Here, we use harmonic mean instead of arithmetic mean. This method is able to classify the images appropriately and not incurred by outliers due to using harmonic mean. This method is also called BFAGC for this reason. From the experimental section, we have shown that our method provides better contrast and brightness than other existing methods. In the future, our method will be applied in various fields such as medical image processing, remote sensing, digital photography, etc.

References

[1] H. K. Ranota and P. Kaur, "Review and analysis of image enhancement techniques," *International Journal of Information & Computation Technology*, vol. 4, no. 6, pp. 583–590, 2014.

[2] A. Piva, "An overview on image forensics," *ISRN Signal Processing*, vol. 2013, 2013.

[3] G. P. Ellrod, "Advances in the detection and analysis of fog at night using goes multispectral infrared imagery," *Weather and Forecasting*, vol. 10, no. 3, pp. 606–619, 1995.

[4] S. Bedi and R. Khandelwal, "Various image enhancement techniques-a critical review," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 3, 2013.

[5] K. Somasundaram and P. Kalavathi, "Medical image contrast enhancement based on gamma correction," *Int J Knowl Manag e-learning*, vol. 3, no. 1, pp. 15–18, 2011.

[6] Y. J. Huang, X. Wang, X. Li, and B. Moran, "Inverse synthetic aperture radar imaging using frame theory," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5191–5200, 2012.

[7] J. R. Jensen and K. Lulla, "Introductory digital image processing: a remote sensing perspective," 1987.

[8] D. Kouno, K. Shimada, and T. Endo, "Person identification using top-view image with depth information," in *Software Engineering, Artificial Intelligence, Networking and Parallel & Distributed Computing (SNPD), 2012 13th ACIS International Conference on*. IEEE, 2012, pp. 140–145.

[9] M. M. Rahman, S. Rahman, E. K. Dey, and M. Shoyaib, "A gender recognition approach with an embedded preprocessing," *International Journal of Information Technology and Computer Science (IJITCS)*, vol. 7, no. 7, p. 19, 2015.

[10] M. M. Rahman, S. Rahman, R. Rahman, B. M. Hossain, and M. Shoyaib, "Dtcth: a discriminative local pattern descriptor for image classification," *EURASIP Journal on Image and Video Processing*, vol. 2017, no. 1, p. 30, 2017.

[11] M. M. Rahman, S. Rahman, and M. Shoyaib, "Mcct: a multi-channel complementary census transform for image classification," *Signal, Image and Video Processing*, vol. 12, no. 2, pp. 281–289, 2018.

[12] M. M. Rahman, S. Rahman, M. Kamal, M. Abdullah-Al-Wadud, E. K. Dey, and M. Shoyaib, "Noise adaptive binary pattern for face image analysis," in *2015 18th International Conference on Computer and Information Technology (ICCIT)*. IEEE, 2015, pp. 390–395.

[13] R. N. Strickland, C.-S. Kim, and W. F. McDonnell, "Digital color image enhancement based on the saturation component," *Optical Engineering*, vol. 26, no. 7, p. 267609, 1987.

[14] S. Rahman, M. M. Rahman, K. Hussain, S. M. Khaled, and M. Shoyaib, "Image enhancement in spatial domain: A comprehensive study," in *Computer and Information Technology (ICCIT), 2014 17th International Conference On*. IEEE, 2014, pp. 368–373.

[15] M. Abdullah-Al-Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae, "A dynamic histogram equalization for image contrast enhancement," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, 2007.

[16] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE transactions on Consumer Electronics*, vol. 43, no. 1, pp. 1–8, 1997.

[17] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," *IEEE Transactions on Consumer Electronics*, vol. 45, no. 1, pp. 68–75, 1999.

[18] M. Kim and M. G. Chung, "Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 3, 2008.

[19] C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered difference representation of 2d histograms," *IEEE transactions on image processing*, vol. 22, no. 12, pp. 5372–5384, 2013.

[20] S.-C. Huang, F.-C. Cheng, and Y.-S. Chiu, "Efficient contrast enhancement using adaptive gamma correction with weighting distribution," *IEEE Transactions on Image Processing*, vol. 22, no. 3, pp. 1032–1041, 2013.

[21] S. Rahman, M. M. Rahman, M. Abdullah-Al-Wadud, G. D. Al-Quaderi, and M. Shoyaib, "An adaptive gamma correction for image enhancement," *EURASIP Journal on Image and Video Processing*, vol. 2016, no. 1, p. 35, 2016.

[22] L. Lu, Y. Zhou, K. Panetta, and S. Agaian, "Comparative study of histogram equalization algorithms for image enhancement," in *Mobile Multimedia/Image Processing, Security, and Applications 2010*, vol. 7708. International Society for Optics and Photonics, 2010, p. 770811.

[23] F. M. Amil, M. M. Rahman, S. Rahman, E. K. Dey, and M. Shoyaib, "Bilateral histogram equalization with pre-processing for contrast enhancement," in *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 2016 17th IEEE/ACIS International Conference on*. IEEE, 2016, pp. 231–236.

[24] T. Celik and T. Tjahjadi, "Contextual and variational contrast enhancement," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3431–3441, 2011.

[25] V. Kumar and P. Gupta, "Importance of statistical measures in digital image processing," *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 8, pp. 56–62, 2012.

Automatic Attendance System: A Face Recognition Approach

[1]Manukumar Rudresh, [2]C. Lalnunchama, [3]Arun Kumar, [4]Vijay Bhaskar Semwal & [5]Vinay Anand Tikkiwal
[1,2,3]Department of Computer Science Engineering , NIT Rourkela, Odisha, India
[4]Department of Computer Science Engineering , NIT Bhopal, MP, India
[5]Jaypee Institute of Information Technology, Noida, India
[1,2,3]{116cs0157, 116cs0197, kumararun}@nitrkl.ac.in, [4]vsemal@gmail.com, [5]vinay.anand@jiit.ac.in

Abstract— **Attendance taking is a tedious process and becomes much tougher with the increase in the number of students registered for the class. Many times, this may lead to the teacher skipping the whole process of attendance or they end up calling out only a few random names from the list to save their time. This is a bad system, as school/institute won't be able to maintain a proper student's record. But, marking attendance for a class may also waste a few minutes of the precious class time. To solve the above-mentioned problem, this paper presents an implementation of automatic attendance management system using a face recognition technique. The GUI is developed using MATLAB functionality. To develop this system, we have first developed a database consisting of the student's details of different age group, different texture like with a beard, without bread, veil etc. and then store the images of each student who has registered for that particular subject. By using the facial features of a person, the proposed system will be able to automatically mark the attendance. Hence, the proposed system is able to maintain a proper database of students and reduces the time wastage during the attendance marking. This research work has explored HOG (Histogram of Oriented Gradient) and PCA (Principal Component Analysis) techniques for face recognition and found HOG better.**

Keywords-Facial Recognition; Automatic Attendance; Image Recognition; Face Detection; Feature Extraction

I. INTRODUCTION

The recent years have witnessed a significant progress in various fields of technology. One such field is Computer Vision & facing Recognition [1]. The domain of computer vision has given numerous real time applications in the recent years as the ability to store large amount of data has boomed along with the increase in computational power of the machines especially with the use of GPUs. The huge amount of storage enabled us to learn the features more accurately and then use it for a variety of purposes including image classification or facial recognition. It can also be applied to object detection which too involves similar idea of computer vision and image processing. Computer vision provides the various techniques to detect the any object present in a digital frame. It will also lead the way to detect the human faces. Face detection, in which human faces are used to detect is the potential application of object class detection [2][3].

Earlier, University's professor used to take attendance using register given to them. But, it is a time-consuming process. It becomes much tougher with the increase in the number of students registered for the class. Many times this may lead to the teacher skipping the whole process of attendance or they end up calling out only a few random names from the list to save their time [4]. This is bad practice as the school/institute would not be able to maintain a proper record of the student's attendance. But, marking attendance for a class may also waste a few precious minutes of the class. It is also very difficult to verify students one by one in a large classroom, whether they are present or not [5].

The primary purpose of developing an attendance management system is to computerize the traditional way of marking attendance. Automated Attendance Management System performs the daily activity of attendance marking and analysis with reduced human intervention. During course registration, individual students face images are captured and stored in the database. The system consists of uploading of images of the students sitting in the classroom. This is followed by sending the image into the Face detection module where individual faces are detected in the image [6], and this is followed by recognition modules [7], which will recognize the student and then the attendance is marked on the database server. Here all the faces are detected from the input image and the algorithm compares them one by one with the face database. If any face is recognized, the attendance is marked on the server and stored in the database [8][9]. To ensure the presence of student, in class the image will be captured twice, once at the beginning and once at the end. Face database is the collection of images of student's faces. The attendance database contains the details about the students and it is also used to mark the attendance. In case, a student is present whose face is not recognized, the class teacher can update the system manually [10][11].

The remaining work is organized as follows: section II presents a description of face detection and extraction. Section III describes the database used and the attendance system. Further, section IV presents the results and discussion. At last, Section V concludes the paper.

II. FACE DETECTION AND EXTRACTION

Face Detection is the most important part in automatic attendance system. Face Detection is used to detect the faces of the students by detecting their faces automatically. And, its searches the student faces in the database. Thus it helps to reduce the burden of a teacher to mark manual attendance of the students. This section presents the process of face

detection and feature extraction of the stored images. Fig 1 shows the flowchart of the proposed system.

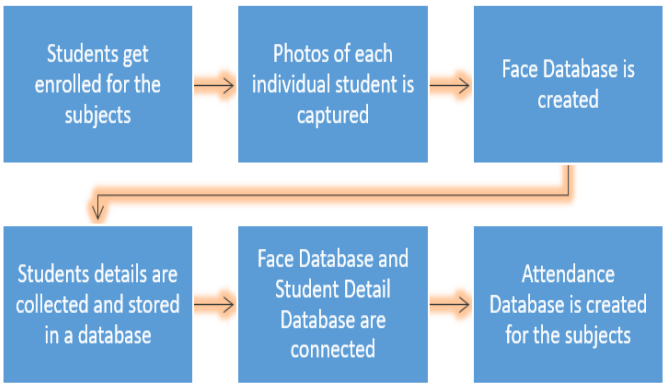


Fig.1. Flowchart of the proposed system.

Design of GUI Interface:

A. Making of GUI

A user interface is made with MATLAB built-in tools. It consists of two axes, one for the input image of the classroom and the other for output after it detects faces. Two push button was made for opening image and detect button. Table I represents the basic information about the functionality of developed GUI.

TABLE I. INFORMATION FOR GUI

Item	Type	Size	Colour
Load	Push Button	1 x 3	Grey
Detect	Push Button	1 x 3	Grey
Title	Static Text	1 x 10	Grey
Image Viewer	Axes	8 x 8	White
Face Viewer	Axes	8 x 8	White

B. Image Insertion

Images are inserted in the GUI file using the select button. The common file format of images which can be used is JPEG, BMP etc. In the case of unsupported file format insertion, the system will through the error/invalid selection messages. Fig. 2 shows the snapshot of the image acquisition interface of the developed GUI.

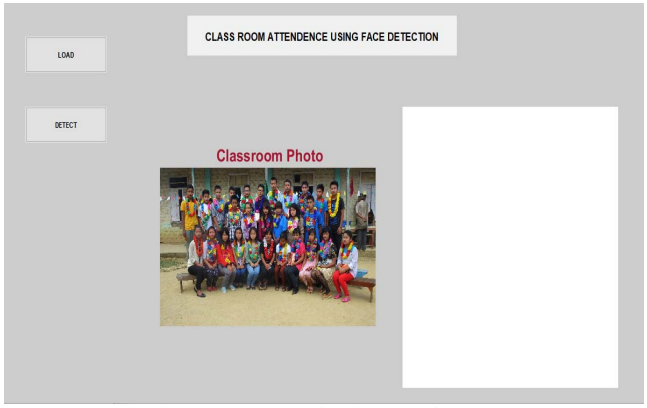


Fig.2. Image Insertion in a sample GUI.

Methodology

A. Face Detection & Recognition

Face Recognition is done using the Viola-Jones algorithm [12]. In this research work, the Viola-Jones algorithm is used for object detection. This algorithm works fast in face detection phase but the training phase is long. The algorithm is robust with a very high detection rate (true-positive), and a very low false-positive. In practical applications, at least two frames per second must be processed. The goal is to distinguish faces of familiar faces [13]. Fig. 3 shows the GUI interface of face detection.

B. Database

A database consists of the images of students faces taken during the class registration. The main folder consists of different folders for each of the students. In the subfolder of each student, eight photos of different shapes and angles are stored. So that a student face can be recognized from different angles and directions.

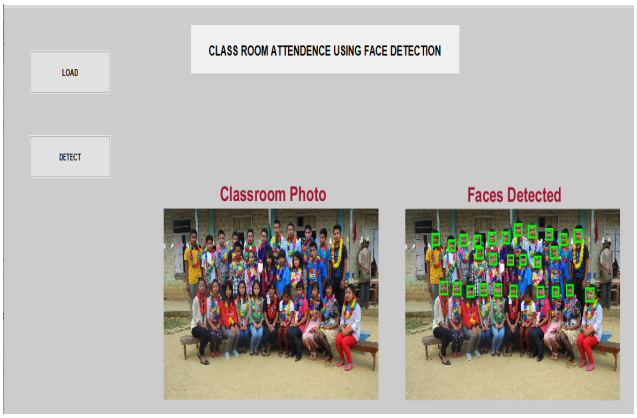


Fig.3. Detection of faces.

C. Face Extraction

The faces in the database are extracted using Histogram of Oriented Gradient algorithm (HOG). HOG is a popular technique to discover shapes within an image with a high detection rate. It works by sampling subregions of the image using a kernel. Then it checks the slope (Aka orientation gradient) and puts it in a bin. Then it samples another region until an entire image is been sampled. After the sampling, the algorithm searches the bin and tries to determine a consensus of slopes to determine the edge.

D. Comparison of Faces

During the comparison of faces, the feature of the individual faces is compared with 'person index' array. This is done to detect the faces. As the same faces will have a similar person index. This can be used to get the similar face from the database for the target face. Fig. 4 shows the query and recognized face.

III. FACE DATABASE AND ATTENDANCE SYSTEM

In this research work, a database of different texture, age group and orientation of faces has been created. So, each subject consisting of 8 different faces. The automatic attendance system will take an image as input from the classroom and detect the faces automatically and mark the

attendance. Fig. 5 presents the procedure to detect the face from a given image.

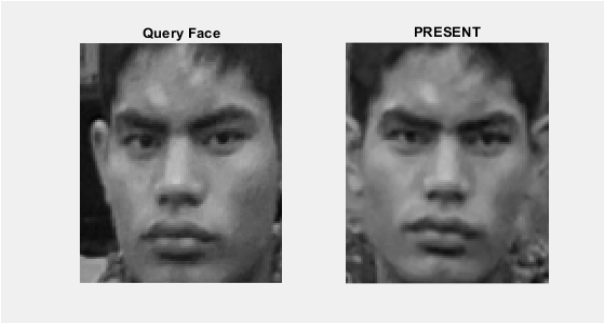


Fig.4. Attendance Procedure

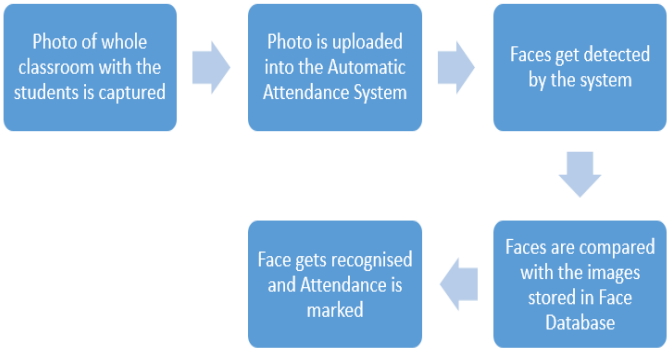


Fig.5. Recognition of Faces from Database

Fig. 6 shows the automatic attendance marked using automatic face recognition based system. An excel sheet is created from the developed GUI. The excel sheet consists of the name of students and roll numbers. Once, a student's face is detected she/he will be marked "Present" in the classroom.

	A	B	C	D
1	No.	RollNumber	29-Nov-18	
2	1	CS008	Present	
3	2	CS009	Present	
4	3	CS005	Present	
5	4	CS018	Present	
6	5	CS000	Present	
7	6	CS004	Present	
8	7	CS017	Present	
9	8	CS003	Present	
10	9	CS014	Present	
11	10	CS013	Present	
12	11	CS012	Present	
13	12	CS027	Present	
14	13	CS010	Present	
15	14	CS016	Present	
16	15	CS020	Present	
17	16	CS019	Present	
18	17	CS007	Present	
19	18	CS006	Present	
20	19	CS021	Present	
21	20	CS022	Present	
22	21	CS025	Present	
23	22	CS024	Present	

Fig.6. Attendance Sheet in MS Excel

IV. RESULTS & DISCUSSIONS

This section describes the results obtained from the proposed system. Fig. 7 shows the face detection accuracy curve using the proposed system up to 25 students. The proposed system achieves high accuracy except in case of low light and bad quality images. Table II shows the face detection rate under different sample size of students.

A. Accuracy of Face Detection Algorithm

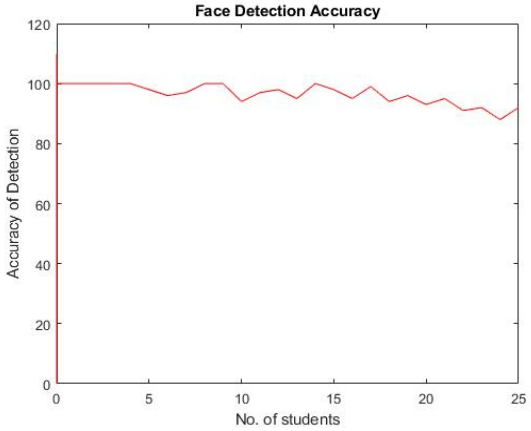


Fig.7. Accuracy of Face Detection.

TABLE II: FACE DETECTION RATE

No.	Students	Detected	Percentage
1.	4	4	100 %
2.	9	9	100 %
3.	14	13	93 %
4.	19	16	89 %
5.	25	22	88 %

B. Accuracy of Face Recognition

The proposed algorithm uses HOG and PCA for face recognition and it has been observed that HOG performing better than PCA. Therefore, this research work HOG has been applied. Table III shows the face recognition rate using the proposed system. Fig. 8 shows the face recognition accuracy curve using the proposed system up to 25 students.

TABLE III: FACE RECOGNITION RATE

No.	Faces	Recognized	Percentage
1.	4	4	100 %
2.	9	8	90 %
3.	14	12	86 %
4.	19	14	74 %
5.	25	19	78 %

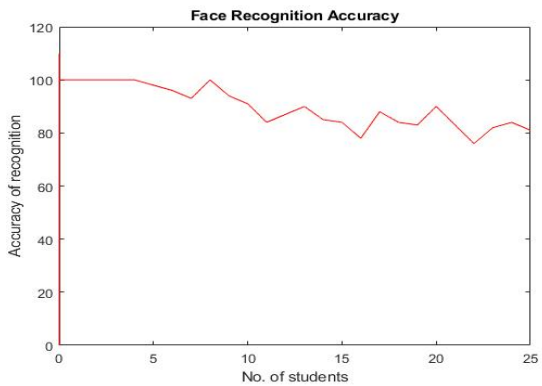


Fig.8. Accuracy of face recognition

C. Comparison with existing work

The proposed system is compared with the existing system [5] and the results show that the proposed system has better accuracy and detection rate. The accuracy table and plot are

represented in Table IV and Fig. 9 for existing and the proposed system.

TABLE IV: COMPARISON WITH THE EXISTING SYSTEM.

No.	Students	Recognized Faces		Accuracy	
		Existing	Current	Existing	Current
1.	4	4	4	100 %	100 %
2.	10	10	9	100 %	90 %
3.	14	11	12	79 %	86 %
4.	20	14	17	70 %	85 %

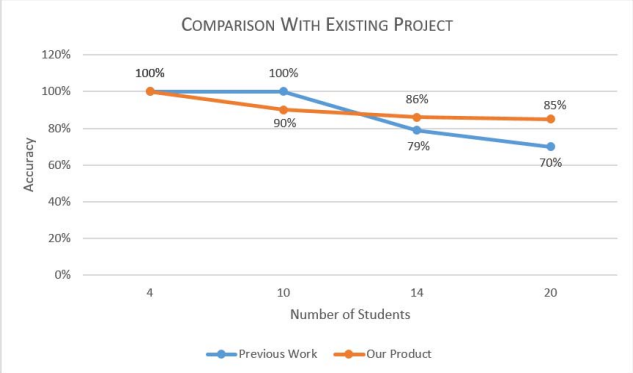


Fig.9. Accuracy comparison with the existing system.

V. CONCLUSION

This research work has presented an automatic attendance system based on face recognition. The proposed system has used the HOG and PCA based techniques to identify the student faces present in the class. A dataset has been created for the different texture, age group and orientation of faces. The results show an accuracy rate of 98% face detection using the proposed system. The proposed system is robust for different view and texture. In future, we plan to implement a gait based student identification system, which will detect the student walking through the front door of the class.

REFERENCES

[1]. Turk, M. A., and A. P. Pentland, D., 1991. "Face recognition using Eigen faces", in the *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,* 60(2), pp. 586 – 591.

[2]. M. H. Yang, N. A., and Kriegmao, D., "Face recognition using kernel Eigen faces", in the *Proceedings of IEEE International Conference on Image Processing,* 1(3), Sept, 2000, pp. 10 – 13.

[3]. Semwal, V. B., Raj, M., & Nandi, G. C., "Biometric gait identification based on a multilayer perceptron" *Robotics and Autonomous Systems*, *65*, 2015, pp. 65-75.

[4]. Slavković, M., & Jevtić, D., "Face recognition using eigenface approach", Serbian Journal of electrical engineering, 9(1), 2012, pp. 121-130.

[5]. Visar Shehu1, A. D., "Using real time computer vision algorithms in automatic attendance management system". In Contemporary Sciences and Technologies, South East European University, Macedonia2, 2008, pp. 1008 – 1014.

[6]. Viola, J. J., "Robust real time face detection". International Journal of Computer Vision, 57(2), 2004.

[7]. Semwal, V. B., Gaud, N., & Nandi, G. C., "Human gait state prediction using cellular automata and classification using ELM", In Machine Intelligence and Signal Analysis (MISP), 2019, pp. 135-145.

[8]. Arulogun O. T., Olatunbosun, A. F. O. A., and Olaniyi, O. M., "Rfid based student's attendance management system", International Journal of Scientific and Engineering Research, 4(5), 2013.

[1] Semwal, Vijay Bhaskar, Neha Gaud, and G. C. Nandi. "Human gait state prediction using cellular automata and classification using ELM." Machine Intelligence and Signal Analysis. Springer, Singapore, 2019, pp. 135-145.

[2] Vijay Bhaskar Semwal and G.C. Nandi, Robust and more accurate feature and classification using deep neural network, Neural Computing and Application, Springer.

[9]. Semwal, Vijay Bhaskar and Gora Chand Nandi. "Biometric gait identification based on a multilayer perceptron" Robotics and Autonomous Systems, 65, 2015, pp. 65-75.

[10]. Semwal, Vijay Bhaskar, et al. "An optimized feature selection technique based on incremental feature analysis for bio-metric gait data classification." Multimedia Tools and Applications :1-19.

[11]. Xi Zhao, Emmanuel Dellandrea and Liming Chen, "A People-Counting System based on Face-Detection and Tracking in a Video", Advanced Video and Signal Based Surveillance, (2009).

[12]. Tsong-Yi Chen, Chao-Ho Chen, Da-Jinn Wang, and Yi-Li Kuo, "A People-Counting System Based on Face Detection", Fourth International Conference on Genetic and Evolutionary Computing, 2010.

[13]. B. Fasel and J. Luettin, "Automatic Facial Expression Analysis: A Survey," Pattern Recognition, Vol. 36, 2003, pp. 259-275,.

[14]. Basbrain, Arwa M., Gan, John Q & and Clark, Adrian", "Accuracy Enhancement of the Viola-Jones Algorithm for Thermal Face Detection", Intelligent Computing Methodologies, Springer, 2017,pp. 71-82".

A Dynamic Face Recognition using HAAR & CNN: A Transfer Learning Approach

Amaan Iqbal[1], Kunal Kumar[2], Arun Kumar[3], Vijay Bhaskar Semwal[4], and Vinay Anand Tikkiwal[5]
[1,2,3]*Computer Science and Engineering, NIT Rourkela, Odisha, India*
[4]*Computer Science and Engineering, NIT Bhopal, MP, India*
[5]*Jaypee Institute of Information Technology, Noida, India*
[1,2,3]{*116cs0157, 116cs0197, kumararun*}*@nitrkl.ac.in,* [4]*vsemwal@gmail.com,* [5]*vinay.anand@jiit.ac.in*

***Abstract*—Face Recognition refers to the identification of a human face from an image, first by finding the region of interest and then by processing the extracted facial part in the region of interest to get the most distinguishing facial features. Face recognition is considered to be the less intrusive and very fast biometric technology which does not require any physical interaction with the person being identified. Further, with the increase in security needs and the advancement in technology, extraction, and processing of information has become much simpler. In this paper, we have implemented and compared two different methods for face recognition. First, we used traditional Haar Cascade Classifier and second, we have used a pre-trained VGGNet to identify human-faces from the input video stream in real-time. Finally, the accuracy of the two methods is compared. In both the methods, the classifier is first trained on a suitable number of images and then it is tested with the input video stream in real-time. If a person is found, its name is displayed beside the face using OpenCV, otherwise, it is characterized as Unknown.**

***Index Terms*—Face Recognition, Face Detection, Transfer Learning, Haar Cascade, Biometric, CNN.**

I. Introduction

These days identification of a person has its application in a wide variety of fields ranging from social media tagging, phone unlocking, attendance in schools, to security purposes where access to a service is granted only to authorized people. Thus suitable measures can be taken in all cases if a person is recognized or some safety measures could be taken when the identity of a person is unknown i.e., the face labels with which the model is not trained. So there is a need to recognize a human face not only with good accuracy but also with the least computation time.

The domain of computer vision has given many real-time applications in the recent years as the ability to store large amount of data has boomed along with the increase in computational power of the machines especially with the use of GPUs. The huge amount of storage enabled us to learn the features more accurately and then use it for a variety of purposes including image classification or facial recognition. It can also be applied to object detection which too involves similar idea of computer vision and image processing. To detect an object in a digital frame, computer vision has various solutions and techniques. Computer vision is also leading the way in which the human faces can be detected. A potential application of object class detection is face detection. Face detection, is done first detecting the eye region and then used the HAAR algorithm to generates face regions for eyebrows, mouth, and other parts. As a final step, different features matching is performed which detects the face.

Face detection involves finding the exact locations of faces from the input frame captured i.e., location, size, position, orientation etc of different faces is determined. Face detection is the first step for facial recognition, whose results directly affect the processing of the human faces and as a result, the accuracy of face recognition model. There are various face detection approach such as Statistics-based, knowledge-based and integration approach. In statistics-based approach a face is detected by judging all possible areas of images by classifier. It means, to look for face region as a class of models, and use a large number of face and non-face training samples to construct the classifier. On the other hand, the knowledge-based approach is known to detect faces for complex background images to some extent. It also has high detection speed. But, in but knowledge-based approach more integration features are required to further enhance the adaptability [18].

The methods compared in this paper are robust with high classification accuracy. The first model i.e., the Haar Cascade Classifier relatively requires more images for training but it is computationally efficient and gives high classification accuracy, whereas the other model i.e, the pre-trained VGGNet requires very less images for training and gives far better accuracy than Haar Cascade Classifier on the test input video stream but it is computationally in-efficient. Though, the pre-trained VGGNet needs a lot of training and testing time, it has real-time detection capability along with high detection accuracy. The generated human face image is a representation of a two-dimensional image as a finite set of digital values, called pixels. Pixel values typically represent gray levels, colours, heights, opacitys etc. The digitization means that a digital image is an approximation of a real scene. The reason for wide use of VGGNet although being less computationally efficient is due to the tremendous growth in the field of computer vision and technology to process large amount of data efficiently is being utilized. An example of this could be the use of GPU for Computer Vision tasks.

Transfer learning is a machine learning method where a model developed for a task is reused as a starting point for a model on a second task. In other words, we use the knowledge acquired by learning from related tasks to do well on the target.

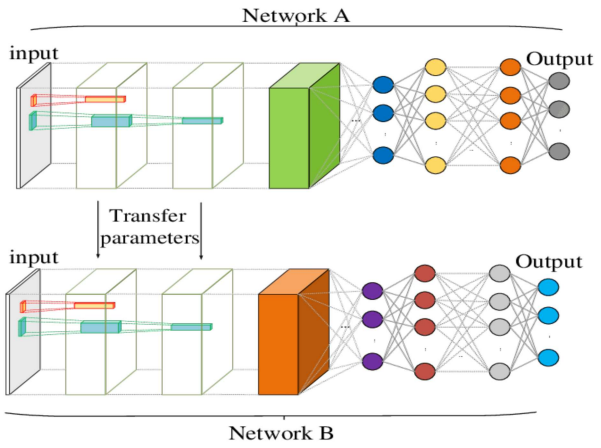


Fig. 1. Illustration of Transfer Learning Concept [19].

In deep learning, as the starting point on computer vision and natural language processing tasks pre-trained models are used. It is due to the vast compute and time resources required to develop neural network models on these problems and the huge jump in skill that they provide on related problems.

It is a popular approach in deep learning where pre-trained models are used as the starting point on computer vision and natural language processing tasks given the vast compute and time resources required to develop neural network models on these problems and from the huge jumps in skill that they provide on related problems.

The remainder of this paper is organized as follows: section II explains the experimental setup and tools used in the research work. Section III describes the implementation of the face recognition approaches. Section IV presents the results and discussion. At last, Section V concludes the paper.

II. Experimental Setup and Tools

This section explains the tools used to detect and track the face and the complete experimental setup.

A Laptop webcam is used to track images in real-time using OpenCV 3.4.1 for windows. the OpenCV module process it and captures the detected face parts. Then, the stored image is given to each of the model to recognizes the faces accordingly. The codes have been written using the Spyder IDE for windows.

A. OpenCV

OpenCV is an Open Source Computer Vision Library which consists of functions mainly useful for computer vision applications. It is free for use and is currently available in C, C++, Ruby, Python and Java. The use of OpenCV library for computer vision tasks reduces the lines of code significantly to perform all the major image processing tasks. It also takes care of small issues like memory leaks with certain programming languages like C and C++.

B. Spyder IDE

Spyder IDE is an environment used for writing and debugging Python codes. It offers a unique combination of the advanced editing, analysis, debugging, and profiling functionality of a comprehensive development tool with the data exploration, interactive execution, deep inspection, and beautiful visualization capabilities of a scientific package.

III. Implementation

The main objective of this research work is to perform detection and recognition of faces from real-time input video. The flow chart given below represents the proposed model:

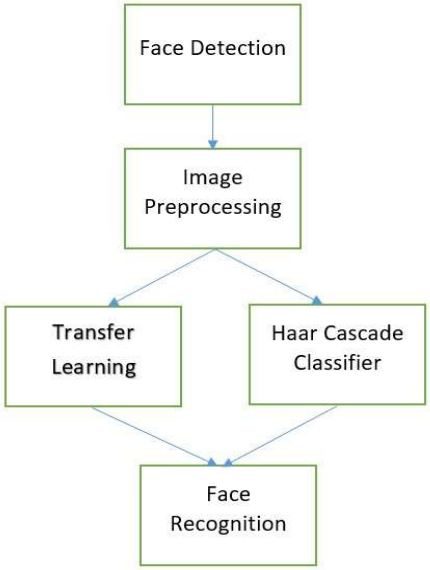


Fig. 2. A Flow Chart of Face Recognition Steps.

For training, 15-20 images have been taken and stored locally for training. After that the images are trained individually using Haar Cascade Classifier which implements Viola-Jones algorithm in its core and pre-trained VGGNet 16 which is a powerful Convolutional Neural Network technique for facial recognition. Although Haar Cascade requires sufficient amount of images for proper training, VGGNet gives really good accuracy with only one image due to the fact that its using pre-trained parameters. VGGNet utilizes 22 layers and 37 deep units. It takes input image of size 224 * 224 * 3 (RGB image). It uses 3*3 size Convolutional Layers, 2*2 size Max Pooling layers and then a fully connected layer at the end. The size of the pre-trained parameters which we used is 553MB. The image of VGGNet architecture is shown below :
Total Parameters Used : 138,357,544
Trainable Parameters : 138,357,544
Non-trainable Parameters : 0

$$softmax(z_i) = \frac{exp(z_i)}{\sum_{j=1}^{n} exp(z_j)} \quad (1)$$

TABLE I
PRE-TRAINED VGG16 ARCHITECTURE

Layers	Output Shape	Parameters #
Input Layer	(None, 224, 224, 3)	0
Block 1 : Conv2D	(None, 224, 224, 64)	1792
Block 1 : Conv2D	(None, 224, 224, 64)	36928
Block 1 : MaxPooling2D	(None, 112, 112, 64)	0
Block 2 : Conv2D	(None, 112, 112, 128)	73856
Block 2 : Conv2D	(None, 112, 112, 128)	147584
Block 2 : MaxPooling2D	(None, 56, 56, 128)	0
Block 3 : Conv2D	(None, 56, 56, 256)	295168
Block 3 : Conv2D	(None, 56, 56, 256)	590080
Block 3 : Conv2D	(None, 56, 56, 256)	590080
Block 3 : MaxPooling2D	(None, 28, 28, 256)	0
Block 4 : Conv2D	(None, 28, 28, 512)	1180160
Block 4 : Conv2D	(None, 28, 28, 512)	2359808
Block 4 : Conv2D	(None, 28, 28, 512)	2359808
Block 4 : MaxPooling2D	(None, 14, 14, 512)	0
Block 5 : Conv2D	(None, 14, 14, 512)	2359808
Block 5 : Conv2D	(None, 14, 14, 512)	2359808
Block 5 : Conv2D	(None, 14, 14, 512)	2359808
Block 5 : MaxPooling2D	(None, 7, 7, 512)	0
Flatten	(None, 25088)	0
Dense	(None, 4096)	102764544
Dense	(None, 4096)	16781312
Predictions(Dense)	(None, 10)	4097000

The webcam is used for real-time video monitoring and face part is detected from the video continuously. Whenever a face is detected, its cropped and stored locally from where the classifier extracts it for recognizing. The stored image is accessed using OpenCV and bounding box is made around the face using the facial coordinates. The Image inside the bounding box is then cropped and processed. It is then compared with the images of the person with which the model is trained. The model is initially trained using 15 images of all three of the team members using Haar Cascade Classifier and with only 1 image using VGGNet.

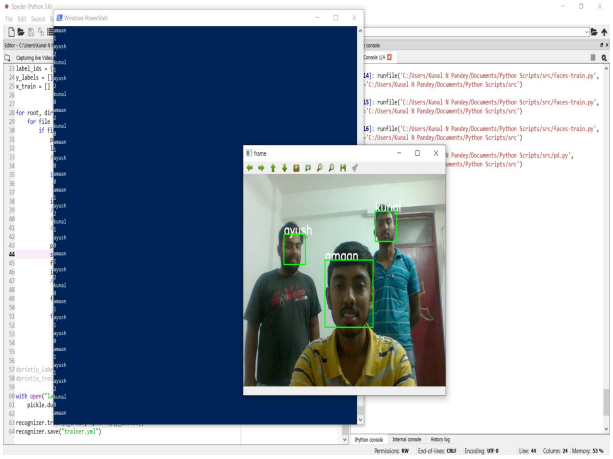


Fig. 3. Real-life Face Recognition.

VGGNet utilizes transfer learning which is widely used these days in Convolutional Neural Networks while dealing with images. We used VGGNet with 16 convolutional layers because of its uniform architecture for image classification and is found that it outperforms the traditional Haar cascade classifier significantly. Model is initially trained by feed forwarding the sensitivities with only one image using VGGNet and was able to classify the person with a very good accuracy.

A web interface has been developed where a user can signup with their name, email and password and can upload images directly on the website. These images are stored in the database and corresponding folder is created with the user's name. The number of images per user has been limited to 30 so that unnecessary overfitting is avoided for specific user. There is a also a button for a user to train the model after uploading their images so that the model can update itself. After training, a button has been provided in the website on clicking which a live window frame opens which captures the video in real-time and face is extracted continuously using infinite while loop and recognized in the same way as specified above. The developed website also has a documentation on how it can be used and a basic chat bot API to know users experience with the website. The is live and can be found on web [20].

We also implemented a module which finds the cosine similarity and euclidean distance between stored image and captures image.

$$cos\theta = \frac{A.B}{|A|.|B|} = \sum_{i=1}^{n} \frac{A_i.B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \quad (2)$$

If the cosine value is lesser than 0.45, then we say that the images are same. In case of euclidean distance comparison, if the distance between the images is less than 120, then we say that the images are same. In both the cases if the above conditions are not satisfied, then the images does not belong to the same person. Here we have represented images as vectors and found the similarity measures of two vectors. If both images are of same person, then measurement should be small. Otherwise, the measurement should be large if two images are of different person.



Fig. 4. Cosine & Euclidean based Comparison.

IV. RESULTS AND DISCUSSION

The deep learning model is trained using Haar Cascade Classifier and pre-trained VGGNet. In the training phase, a

certain number of images of 10 different person have been used. In the testing phase, the image is monitored using a real-time window frame continuously and the face part is detected and extracted using OpenCV by making a bounding box around the face which is then stored locally. In the same iteration, we take the stored image as input and gave it to the two face recognition models which we have used namely Haar Cascade Classifier and pre-trained VGGNet which utilizes transfer learning. This happens infinitely unless we terminate the browsing window. The captured image is fed to both of these networks and was found that VGGNet automatically extracts the best features for classification and outperforms traditional Haar Cascade classifier significantly.

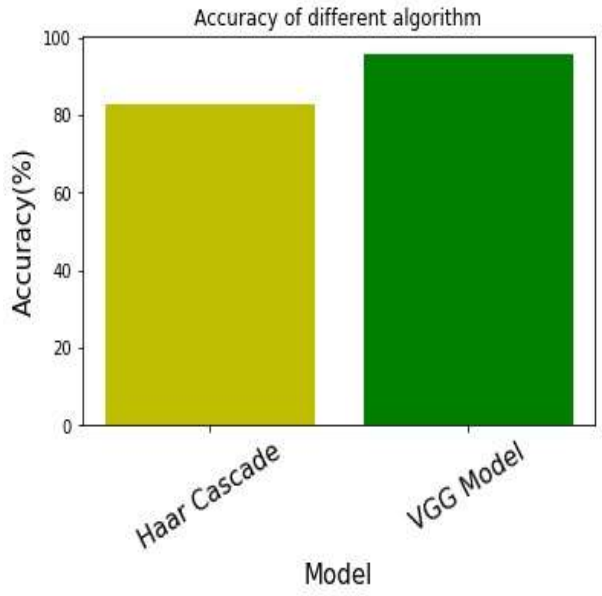


Fig. 5. Accuracy Comparison of Algorithms.

It has been found that the traditional Haar Cascade Classifier has an accuracy of approximately 83% whereas pre-trained VGGNet with transfer learning has an accuracy of approximately 92%. The only problem with VGGNet is that it is more computationally expensive than Haar Cascade Classifier because of its use of deep neural networks in its architecture as defined previously.

V. Conclusion

In this paper, We have implemented two deep learning techniques for facial recognition using Haar Cascade Classifier and pre-trained VGGNet with transfer learning. It has been found that VGGNet outperforms the traditional Haar Cascade Classifier significantly. Simultaneously, while capturing the images, our model displays the name of the person in the Video frame. Initially, we started with Haar Cascade Classifier but then we extended it to use more robust pre-trained VGGNet with transfer learning, a Deep Learning technique to extract best features automatically, resulting in the classification of images with even better accuracy at the cost of some extra computation. In this paper, only the detection of frontal human faces was considered. In the future we plan to extend this work to minimize the effect of rotation, side view, poses, expressions, disguises, and illumination difference while classifying an image. Also, we plan to incorporate the liveliness of image, or if it is a real image or a static picture.

References

[1] Taigman, Yaniv, et al. "Deepface: Closing the gap to human-level performance in face verification." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.

[2] Sun, Yi, Xiaogang Wang, and Xiaoou Tang. "Deep learning face representation from predicting 10,000 classes." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.

[3] Le, Quoc V. "Building high-level features using large scale unsupervised learning." Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on. IEEE, 2013.

[4] Sun, Yi, et al. "Deepid3: Face recognition with very deep neural networks." arXiv preprint arXiv:1502.00873 (2015).

[5] Sun, Yi, et al. "Deep learning face representation by joint identification-verification." Advances in neural information processing systems. 2014.

[6] Cao, Xudong, et al. "A practical transfer learning algorithm for face verification." Proceedings of the IEEE International Conference on Computer Vision. 2013.

[7] Semwal, Vijay Bhaskar, Manish Raj, and Gora Chand Nandi. "Biometric gait identification based on a multilayer perceptron." Robotics and Autonomous Systems 65 (2015): 65-75.

[8] Wen, Yandong, et al. "A discriminative feature learning approach for deep face recognition." European Conference on Computer Vision. Springer, Cham, 2016.

[9] Liu, Weiyang, et al. "Sphereface: Deep hypersphere embedding for face recognition." The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Vol. 1. 2017.

[10] Lu, Jiwen, Venice Erin Liong, and Jie Zhou. "Simultaneous local binary feature learning and encoding for face recognition." Proceedings of the IEEE International Conference on Computer Vision. 2015.

[11] Lu, Jiwen, et al. "Learning compact binary face descriptor for face recognition." IEEE transactions on pattern analysis and machine intelligence 37.10 (2015): 2041-2056.

[12] Tan, Xiaoyang, and Bill Triggs. "Enhanced local texture feature sets for face recognition under difficult lighting conditions." IEEE transactions on image processing 19.6 (2010): 1635-1650.

[13] Semwal, Vijay Bhaskar, Neha Gaud, and G. C. Nandi. "Human gait state prediction using cellular automata and classification using ELM." Machine Intelligence and Signal Analysis. Springer, Singapore, 2019. 135-145.

[14] Pan, Sinno Jialin, and Qiang Yang. "A survey on transfer learning." IEEE Transactions on knowledge and data engineering 22.10 (2010): 1345-1359.

[15] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436.

[16] Semwal, Vijay Bhaskar, Kaushik Mondal, and Gora Chand Nandi. "Robust and accurate feature selection for humanoid push recovery and classification: deep learning approach." Neural Computing and Applications 28.3 (2017): 565-574.

[17] Semwal, Vijay Bhaskar, et al. "An optimized feature selection technique based on incremental feature analysis for bio-metric gait data classification." Multimedia tools and applications 76.22 (2017): 24457-24475.

[18] Viraktamath, S. V., et al. "Face detection and tracking using OpenCV." The SIJ Transactions on Computer Networks Communication Engineering (CNCE) 1.3 (2013): 2321-2403.

[19] Joseph, L., et al. "Transfer Learning of Temporal Information for Driver Action Classification." 28th Modern Artificial Intelligence and Cognitive Science Conference, Fort Wayne Indiana, April 2017.

[20] Semwal, Vijay Bhaskar, Neha Gaud, and G. C. Nandi. "Human gait state prediction using cellular automata and classification using ELM." Machine Intelligence and Signal Analysis. Springer, Singapore, 2019. 135-145.

Heart Failure Risk Prediction and Medicine Recommendation using Exploratory Data Analysis

Sumaya Habib
Department of Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
sumaya.official@gmail.com

Maisha Binte Moin
Department of Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
maisha.moin@gmail.com

Sujana Aziz
Department of Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
suzana6179@gmail.com

Kalyan Banik
Algorithm Developer
Codemen Solutions Inc.
Dhaka, Bangladesh
kalyan.b.aninda@gmail.com

Hossain Arif
Department of Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
hossain.arif@bracu.ac.bd

***Abstract*— With the ever increasing population of the world, diseases and their possibilities are also increasing at an alarming rate. As time passes by, diagnosing diseases and providing appropriate treatment at the right time has become quite a challenge. Heart diseases have been a major cause of death worldwide. Therefore, this research has been focused on finding an efficient way to predict the chances of a heart failure and accordingly, recommend appropriate medicines to aid cardiologists in quicker decision making. The research includes finding the correlations or associations between the various medical profiles of the patients by utilizing the standard techniques of exploratory analysis and hence using the attributes suitably to predict the chances of a heart failure, as well as the medicine recommendations. A comparative study has also been included which shows the various attained accuracy rates of different machine learning algorithms including - Logistic Regression, Naïve Bayes, Decision Tree, Linear Support Vector Classifier, Random Forest, and Gradient Boosting Classifier. The Apache Spark framework has been used in order to make the system capable of handling big data.**

Keywords—Machine Learning, Exploratory Analysis, Heart Failure Prediction, Medicine Recommendation

I. INTRODUCTION

According to the data published by World Health Organization (WHO), in the year 2017, 14.31% of deaths in Bangladesh were caused by Coronary Heart Diseases [11]. In fact, the latest (2018) report provided by WHO has marked Heart Disease as one of the biggest causes of death all around the world [12]. With the growing population in countries like Bangladesh, it has become very difficult to provide proper treatment at the right time. This is mainly due to the reason that the medical test reports are usually manually analyzed by cardiologists before they can reach a reasonable conclusion. It is, undoubtedly, a time consuming process. Therefore, we intended on designing a system that would analyze the patient's medical test reports and provide cardiologists with the chances of a heart failure, and even suggest a medicine list that would be most suitable for the patient. The idea was to create a system that would facilitate cardiologists in taking quicker decisions so that more patients can receive treatments within a shorter period of time.

We have used an Austrian data set which originally contains 43 attributes, including information such as the patient's age, gender, existing diseases, smoking habits, and the prescribed medicines for each individual. A significant portion of the research had been invested on finding the correlations or associations between these attributes and selecting the attributes which would be a stronger candidate for aiding the predictions. Various techniques of exploratory analysis such as basic statistical graphs, heatmaps, etc. had been used for a better understanding of the data set. Once done with the exploratory analysis, the selected attributes were fed into various classification algorithms, which included: Logistic Regression, Naïve Bayes, Decision Tree, Random Forest, Linear Support Vector Machine Classifier (SVC), and Gradient Boosting Classifier. Therefore, our research focused on prototyping a model that would be able to provide valuable insights on the huge amount of patient data that is rapidly being generated all around the world.

This paper has been organized as follows: Section-II discusses the research work which were conducted on similar problems, Section-III contains the data set details and the Exploratory Data Analysis (EDA) techniques used, Section-IV demonstrates the methodologies and results, and Section-V concludes the paper.

II. RELATED WORK

Several researches have already been conducted on similar topics which provide rich guidelines on the basic idea of fusing Computer Science with Medical Science. Results of various surveys showed the impact of big data in the field of healthcare [5]. The research also included a comparison section of various tools and techniques including Hadoop, Storm, Spark, and High Performance Computer Cluster (HPCC), which provide solutions to the challenges imposed by big data.

Naïve Bayesian Classification technique was used to develop decision support in Heart Disease Prediction

System by Priyanga and Naveen [7]. With unknown input sample, Naïve Bayes Weighted Approach system classified into various sections and traced the risk of heart disease. The system was developed with 13 attributes provided by the user. The hidden knowledge could be extracted from related data.

Thomas and Princy proposed a system which predicted heart diseases using two modules: the classifier module and the prediction module. The classifier module used K-Nearest Neighbor (KNN) algorithm for classification and training, after which the second module found the risk rate using Iterative Dichotomiser 3 (ID3) algorithm. The mentioned data set consisted of 3 types of attributes: input, key, and prediction attributes [9].

While studying recent publications of emerging technologies which play a role in the healthcare sector, Miotto et al. discussed the latest applications of deep learning in the said sector which focused on biomedical data obtained from clinical imaging, genomes, etc. A tabular summary of the work done in this field using various deep learning architectures was also provided. They discussed the challenges faced, such as - data volume, temporality of the data, and the domain complexity; and scopes, such as - feature enhancement, federated inference, model privacy, incorporation of expert knowledge, and so on [4].

Soni, Ansari, Sharma, and Soni did a comparative study of data mining techniques proposed in recent years for heart disease prediction. The comparison conducted on the Cleveland data set using their Intelligent Heart Disease Prediction System showed that Decision Tree (89%) out-performed the other classification techniques, whereas Bayesian Classifier (86.53%) and Artificial Neural Networks (85.53%) also provided satisfactory results. They implemented Multilayer Perceptron Neural Network for comparison. Finally, their research showed that the results improve significantly with the use of genetic algorithm by reducing the data set to an optimal one [8].

Application of Hidden Naïve Bayes (HNB) classifier for the diagnosis of heart disease was used for testing its performance on the Statlog data set. Experimental results showed that HNB ensured comparatively better performance than other approaches. Proposed model exhibited data discretization and interquartile range filtering to enhance efficiency of HNB. The model recorded highest accuracy of 100% compared to Naïve Bayes classification model. HNB model assisted trustworthy and more authentic decision support system for automatic prediction of heart diseases [3].

The design and implementation of a worldwide medicine recommendation system framework that applies data mining technologies to the recommendation system had been proposed by Bao and Jiang [1]. The medicine recommendation system consisted of database system module, data preparation module, recommendation model module, model evaluation, and data visualization module. They researched medicine recommendation with Support Vector Machine (SVM), Back Propagation (BP) Neural Network and ID3 Decision Tree algorithm on the diagnosis data. Lastly, on the open data set, SVM recommendation model was chosen for the medicine recommendation system to get a better trade among model accuracy, efficiency, and scalability and to suggest an error checking process to ensure the accuracy and service quality.

Another study was based on an instance-based algorithm which permitted to explore big data sets by making pairwise relations between patients with metric-free method. Here each data set had ranks. By clumping these ranking orders, it was possible to conceptualize data based on each set, holding subsets of similar and correlated patients. The paper also demonstrated a visualizing tool for utilizing a database of diabetic patients. This module of a recommendation system applied the denoted algorithm to enhance data, structure patients, create relations between individuals and give recommendations [2].

III. Exploratory Analysis and Data Set

We have used the Austrian medical data set available from [6] for our research. The data set originally contained 45 attributes and 2300 rows. Table I shows some of these attributes and their classification category.

TABLE I. Some of the data set attributes and their classifitcations

Attribute	Classification	Attribute	Classification
age	numerical	pacemaker	yes/no
gender	male/female	crt-defib	yes/no
CAD-fam-history	yes/no	platelet-inhibitors	yes/no
hyperlipidemia	yes/no	oral-anticoagulants	yes/no
hypertension	yes/no	ace-at2	yes/no
NIDDM	yes/no	lipid-lowering	yes/no
smoking	yes/no/previous	diuretics	yes/no
CAD-year	numerical	beta-blockers	yes/no
chron-lung-disease	yes/no	calciumantagonists	yes/no
previous-stroke	yes/no	ivabradin	yes/no
depression	yes/no	nicorandil	yes/no
any-angina	yes/no	molsidomin	yes/no
angina-episodes-week	numerical	antiarrhythm	yes/no
heart-failure	yes/no	ranolazin	yes/no
heart-rate	numerical	glycosides	yes/no
atrial-fibrillation	yes/no	long-acting-nitrates	yes/no

A. Understanding the Data Set

Exploratory Data Analysis (EDA) is a strategy for skimming data sets and seeing what information they contain. This is often done through data visualization techniques such as - statistical graphs and plots. It basically helps in summarizing the properties of the data set [10].

We plotted bar graphs to see the frequency distribution of each individual attribute with the other attributes which are supposed to have a high correlation with them. Figure 1 below shows one of these graphs which demonstrate the percentage of patients who have been prescribed glycoside. The blue bars represent the patients who have the said diseases or symptoms, while the green bars represent those who do not. From this figure we concluded that heart-failure and atrial-fibrillation seems to have a positive correlation with glycoside. According to our consultation with a cardiologist, glycosides are, in fact, used to treat congestive heart failure (CHF) and cardiac arrhythmias.

Thus we verified that our data set is consistent with actual medical definition of the said attributes.

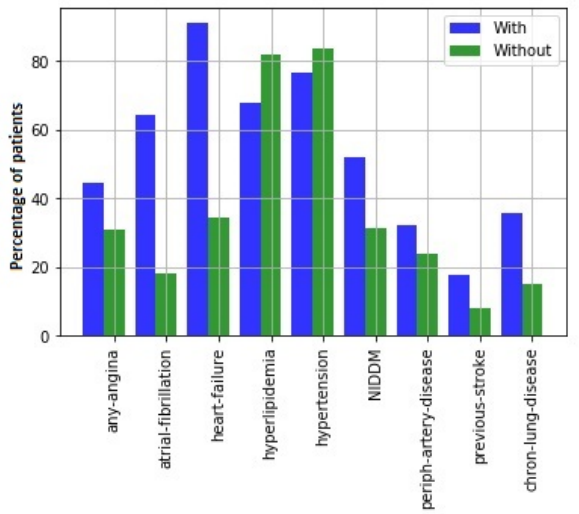


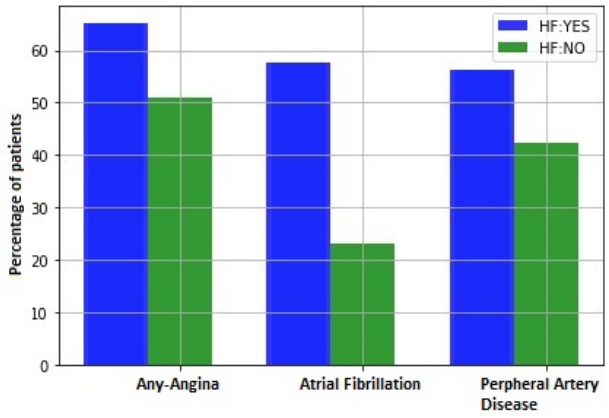
Fig. 1 Percentage of patients (who have been prescribed Glycoside) with and without the diseases and symptoms.

Some of the inferences made on the medicine columns on the basis of these bar graphs include:

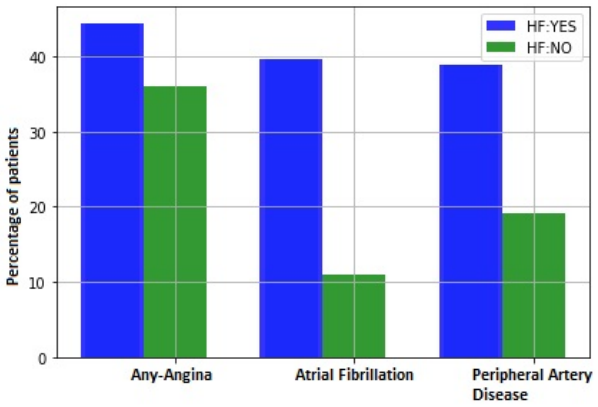
- **Ace-At2** having a positive correlation with Hypertension, Heart-Failure, NIDDM (non-insulin-dependent diabetes mellitus), and Periph-Artery-Disease.
- **Beta-Blockers** having a positive correlation with Hyperlipidemia and Hypertension; and a negative correlation with Periph-Artery-Disease.
- **Calciumantagonists** having a positive correlation with Any-Angina, Hypertension, and NIDDM.
- **Diuretics** having a positive correlation with Heart-Failure, Atrial-Fibrillation, Periph-Artery-Disease, Previous-Stroke, Any-Angina, and NIDDM.
- **Glycosides** having a positive correlation with Atrial-Fibrillation, Heart-Failure, NIDDM, and Any-Angina; and a negative correlation with Hyperlipidemia.
- **Long-Acting-Nitrates** having a positive correlation with Any-Angina and NIDDM; and a negative correlation with Atrial-Fibrillation, Hyperlipidemia, and Heart-Failure.

Since Heart-Failure was our prime target attribute, we have plotted a different set of bar graphs for visualizing the relationship between Heart-Failure (HF) and other diseases and symptoms, separately for males and females. Figure 2 shows two of these graphs. Here, the blue bars represent the percentage of patients who have heart-failure and also have the labeled diseases, while the green bars represent the percentage of patients who do not have heart-failure but have the labeled diseases. Some of the inferences made on these graphs include:

- In case of both males and females, most of the patients who had a record of Heart-Failure also had Atrial-Fibrillation.



(a) Report for Male Patients



(b) Report for Female Patients

Fig. 2 Heart-failure correlation with respect to the three main attributes related to the heart.

- Most patients (male and female) with and without Heart-Failure report had Angina. So the correlation between Angina and Heart-Failure can be considered as weak, from these graphs.
- Whereas almost 50% of females with a Heart-Failure did not have Peripheral Artery Disease, the distinction is lower for male patients. Therefore, although we might infer that females with Peripheral Artery Disease are likely to have a Heart Failure, this cannot be said for the male patients.
- As for the other diseases, both male and female patient reports showed similar trends.
- Female patients had a higher frequency of depression than male patients.

B. Association Between Attributes

Although the graphs provided us with a primary insight on the data set, we still needed stronger evidences to confirm these hypotheses. For that we have applied various techniques to figure out which attributes are strongly associated with each other. Since most of the columns were categorical, the techniques used includes Pearson Correlation Coefficient and Spearman's Rank Correlation Coefficient.

Both Pearson-R and Spearman-R coefficients range from -1 to 1, to show negative linear correlation to positive linear correlation and a 0 to indicate zero-correlation. The problem with these methods is that they consider the attribute values as ordinal data. Since most of the attributes that we were concerned with were nominal, correlation was not quite the correct term that we were looking for. What we really wanted to see was how much the nominal attributes were associated with each other. The technique used for that purpose is Cramér's V. Figure 3 demonstrates the calculated Cramér's V coefficients for some of the attributes. In consideration of the values obtained, we decided on a threshold value of 20%.

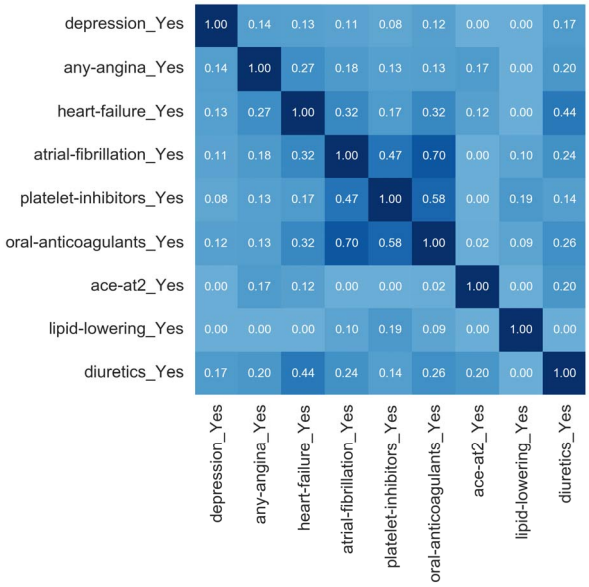


Fig. 3 Heatmap showing the Cramer's V coefficients of some of the attributes.

Some inferences based on the Cramér's V coefficients include:

- **Heart-Failure** has a strong association with Peripheral Artery Disease, Chronic Lung Disease, Angina, Atrial Fibrillation, Oral Anti-Coagulants, Diuretics and Glycosides.
- **Platelet Inhibitors** have a strong association with Atrial Fibrillation, Oral Anti-Coagulants and Glycosides.
- **Oral Anti-Coagulants** have a strong association with Gender, Heart-Failure, Atrial Fibrillation, Platelet Inhibitors, Diuretics and Glycosides.
- **Ace-At2** has a strong association with Hypertension and Diuretics.
- **Lipid Lowering** has a strong association with Hyperlipidemia

C. *Data Set Preprocessing*

From Table II, we can see that 23 of the selected attributes had at least 20% or more of the data missing. We could not directly eliminate these attributes since most of them were vital for the purpose of our research. The only columns dropped from this survey were: nitro-week, platelet-inhibitors-dual, CAD-year, beta-blockers-dosis, CAD-fam-history, angina-episodes-week, smoking, physical-activity, crt-defib, pacemaker and angina class.

For the rest of the attributes, we also could not just drop the missing rows since that would drastically decrease the size of our data. Therefore, we used Logistic Regression Classifier to predict the missing data using the 13 attributes with the least missing data rate. We used One-Hot Encoding before proceeding with the algorithm. Table III shows the iteration accuracy rates of each of these predicted attributes.

TABLE II. PERCENTAGE OF MISSING DATA FOR SOME ATTRIBUTES

Attribute	Percentage	Attribute	Percentage
calciumantagonists	0.000	hyperlipidemia	57.460
beta-blockers	0.000	crt-defib	57.503
age	0.695	ace-at2	57.721
heart-rate	1.304	physical-activity	57.721
hypertension	1.739	glycosides	57.808
NIDDM	2.001	molsidomin	58.025
nicorandil	3.392	oral-anticoagulants	58.199
chron-lung-disease	3.915	ranolazin	58.199
depression	4.393	previous-stroke	58.330
long-acting-nitrates	4.524	smoking	58.504
ivabradin	4.828	platelet-inhibitors-dual	58.504
periph-artery-disease	4.915	CAD-year	59.983
any-angina	14.702	heart-failure	60.244
angina-class	21.096	antiarrhythm	60.853
beta-blockers-dosis	30.187	diuretics	63.245
pacemaker	56.633	angina-episodes-week	70.987
lipid-lowering	56.633	CAD-fam-history	73.032
atrial-fibrillation	56.851	nitro-week	88.212
platelet-inhibitors	57.460		

TABLE III. ITERATION ACCURACIES OF PREDICTED ATTRIBUTES

Attribute	Iteration Accuracy
Heart-Failure	0.80
Lipid-Lowering	0.94
Atrial-Fibrillation	0.92
Platelet-Inhibitors	0.88
Ace-At2	0.94
Hyperlipidemia	0.94
Glycosides	0.88
Molsidomin	0.90
Oral-Anticoagulants	0.88
Ranolazin	1.00
Previous-Stroke	0.92
Diuretics	0.86
Antiarrhythm	1.00

After filling out the missing values, we dropped the columns - Antiarrhythm and Ranolazin because their unnaturally high accuracy rates were owing to the fact that most of their rows had a "No" classification.

After taking all these measures, our final processed dataset had a total of 25 attributes and 1614 rows (without any missing values). In order to check the reliability of this semi-synthesized dataset, we ran the Cramer's V on it. Comparing this new heatmap (figure 4) with the one obtained before, it was seen that although the values changed (due to an increase in data size), the relative association between the attributes remained very similar. Therefore, we concluded that this new dataset was in fact, consistent with the original dataset and thereafter, we proceeded with the machine learning algorithms.

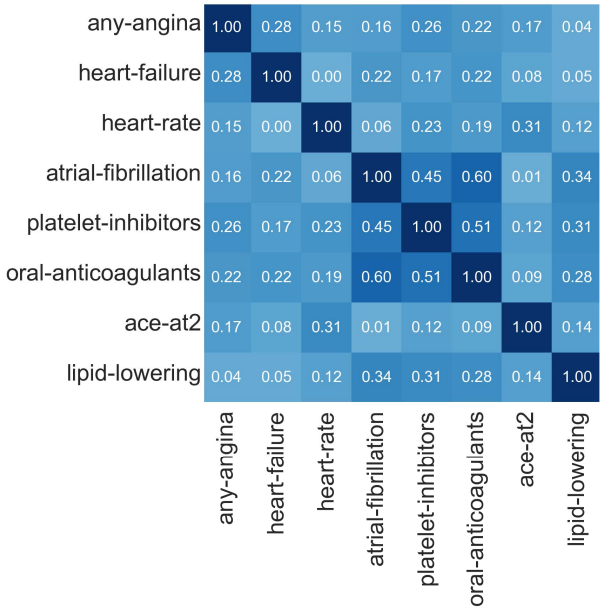


Fig. 4 The new Heatmap showing the Cramer's V coefficients of some of the attributes of the semi-synthesized dataset.

IV. Methodology and Results

A. Classifiers and Evaluation Techniques

We have used Logistic Regression (LR), Naïve Bayes (NB) (Bernoulli's Naïve Bayes to be more precise), Decision Tree (DT), Random Forest (RF), Linear Support Vector Classifier (SVC) and Gradient Boosting Classifier (GBC) for our comparative study.

The standard model evaluation criteria used includes: Accuracy, Precision, Recall, F1-Score, and Area under ROC (Receiver Operating Characteristic). Equations 1, 2 and 3 represent the equations for Precision, Recall and F1-score respectively, where TP, FP and FN stands for True-Positive, False-Positive and False-Negative, respectively.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

$$F1 = \frac{2(\text{Recall} * \text{Precision})}{\text{Recall}+\text{Precision}} \quad (3)$$

B. Hyperparameter Tuning

Our data set was randomly divided into two parts: 80% for training and validating the model, and 20% for testing. The data separated for training and validation were then used to tune the models, to figure out the optimum parameters for each classifier. The validation parameters were set to 10 fold cross-validation, with a maximum of 50 iterations. After the best model was selected for each classifier, they were once again tested on the entire data set, with a cross-validation of 10 folds.

C. Results

Tables III and figure 5 demonstrates the obtained evaluation readings during the modelling and testing phase. The model accuracy is highest for Logistic Regression and Decision Tree, followed by Gradient Boosting Classifier. The difference between accuracy rates of model performance and test performance is the least for Gradient Boosting Classifier, whereas it is the highest for Decision Tree. Despite the fact that Random Forest is an upgradation of Decision Tree, the overall model performance turned out to be better for Decision Tree, than for Random Forest. Naïve Bayes and SVC, on the other hand, showed similar model performance. Their test scores were also quite close. The test performance usually signifies how well the model performs on unexplored data. In our case, however, since the data size was not that big, a small portion of the test data was unexplored. In any case, the best performer among the six algorithms turned out to be Logistic Regression, closely

TABLE III. Model Performance

Algorithms	Accuracy	Precision	Recall	F1-Score	AUC
LR	0.91	0.92	0.91	0.91	0.88
NB	0.83	0.83	0.84	0.83	0.74
DT	0.91	0.92	0.92	0.92	0.87
RF	0.89	0.89	0.89	0.89	0.89
SVC	0.83	0.83	0.84	0.83	0.78
GBC	0.90	0.90	0.90	0.90	0.86

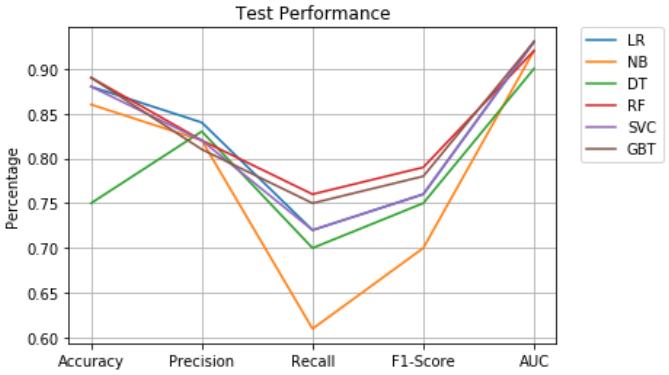


Fig. 5 Test Performance Comparison of the Algorithms.

followed by Gradient Boosting Classifier, Random Forest and SVC. Therefore, our system could predict the chances of a heart-failure and recommend medicines at the maximum model accuracy rate of 91% and test accuracy rate of 88% by using the Logistic Regression model.

D. Analysis

Logistic Regression applies logistic functions in order to map the predictions between the binary numbers 0 and 1. Since these models are still somewhat linear, they usually perform well when the data classes can be separated by a single decision barrier. Hence, it outperformed the other algorithms when run on this particular dataset. As for the classification trees, since they are designed to deal with non-linear decision boundaries, they force every variable in the tree to interact with every other variable. This can lead to poor predictions when the attributes are weakly correlated. The reason why SVC did not perform as well as Logistic Regression, could be that they are usually trickier to tune. In case of Naïve Bayes, on the other hand, it was no surprise that the other algorithms performed better. Since NB is far too "naive" or simple, other fine-tuned algorithms often get the upper hand.

V. CONCLUSION

With the aim to help Cardiologists take quicker decisions, we have tried to build a system that would predict the chances of a heart failure based on the patients' reports and also suggest a medicine list that would be most suitable for the patient. For our research we have used an Austrian medical data set which originally contained 45 attributes and 2300 rows. After completing Exploratory Data Analysis and data preprocessing, we were left with a total of 25 attributes, 1614 rows and a list of attributes that would be the best candidates for each predictive sub-system. Next, we used these features to predict our target attributes, namely - Heart Failure, Platelet Inhibitor, Oral Anticoagulant, Ace-At2, Lipid-Lowering, Diuretic, Beta-Blocker, Calciumantagonist, Ivabradin, Nicorandil, Molsidomin and Long Acting Nitrates. The algorithms used included Logistic Regression, Naïve Bayes, Decision Tree, Random Forest, Linear SVC and Gradient Boosting Classifier. Best model and test performance was attained by Logistic Regression, followed by Gradient Boosting Classifier and Decision Tree.

In order to make our model capable of handling big data, the final model was coded using the Apache Spark framework. It is an open source cluster computing framework which supports Hadoop Distributive File System (HDFS). The Python's API for Spark, known as PySpark, was used. It currently supports all the classifiers mentioned above.

REFERENCES

[1] Bao, Y. and Jiang, X. (2016). An intelligent medicine recommender system framework. In Industrial Electronics and Applications (ICIEA), 2016 IEEE 11th Conference on, pages 1383–1388. IEEE.

[2] Falip, J., Aït-Younes, A., Blanchard, F., Delemer, B., Diallo, A., and Herbin, M. (2017). Visual instance-based recommendation system for medical data mining. Procedia computer science, 112:1747–1754.

[3] Jabbar, M. and Samreen, S. (2016). Heart disease prediction system based on hidden naïve bayes classifier. In International Conference on Circuits, Controls, Communications and Computing (I4C), Bangalore.

[4] Miotto, R., Wang, F., Wang, S., Jiang, X., and Dudley, J. T. (2017). Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics.

[5] Patel, J. A. and Sharma, P. (2014). Big data for better health planning. In 2014 International Conference on Advances in Engineering and Technology Research (ICAETR), pages 1–5. IEEE Unnao.

[6] Pichlhöfer, O., Maier, M., Badr-Eslam, R., Ristl, R., Zebrowska, M., and Lang, I. M. (2017). Clinical presentation and management of stable coronary artery disease in austria. https://figshare.com/articles/Clinical_presentation_and_management_of_stable_ coronary_artery_disease_in_Austria/4925702. [Online; accessed 30-November-2018].

[7] Priyanga, P. and Naveen, N. (2017). Web analytics support system for prediction of heart disease using naive bayes weighted approach (nbwa). In 2017 Asia Modelling Symposium (AMS), pages 21–26. IEEE.

[8] Soni, J., Ansari, U., Sharma, D., and Soni, S. (2011). Predictive data mining for medical diagnosis: An overview of heart disease prediction. International Journal of Computer Applications, 17(8):43–48.

[9] Thomas, J. and Princy, R. T. (2016). Human heart disease prediction system using data mining techniques. In Circuit, Power and Computing Technologies (ICCPCT), 2016 International Conference on, pages 1–5. IEEE.

[10] Wikipedia contributors (2018). Exploratory data analysis — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Exploratory_data_analysis& oldid=871191702. [Online; accessed 30-November-2018].

[11] World Health Organization (2017). World health rankings. https://www. worldlifeexpectancy.com/bangladesh-coronary-heart-disease. [Online; accessed: 02December-2018].

[12] World Health Organization (2018). World heart day. https://www.who.int/ cardiovascular_diseases/world-heart-day/en/. [Online; accessed: 02-December-2018].

Chapman-Kolmogorov Relation Based Median String Algorithm for DNA Consensus Classification

Mohammad Shibli Kaysar
Deptartment of Computer Science and Engineering
Chittagong University of Engineering and Technology
Chittagong, Bangladesh
*shibli.kaysar@gmail.com

Mohammad Ibrahim Khan
Deptartment of Computer Science and Engineering
Chittagong University of Engineering and Technology
Chittagong, Bangladesh
muhammad_ikhancuet@yahoo.com

***Abstract*—Consensus string is the most frequent common pattern in a set of string. Consensus string is an important feature of DNA sequence. Many algorithm have been introduced to discover consensus string. Among them, median string algorithm is the most popular one. Basically, that is a brute force algorithm.DNA sequence is composed of a series of four letter alphabet ∑={a,c,g,t}. If the size of the consensus string is l, then the algorithm generates all the 4^l number of l length strings called motifs or l-mer. Then try to fit the motifs one by one with the sequence. In this paper we have discovered a way to reduce the search space using chapman kolmogorov relation. We found that, the proposed system can find the same consensus string within a shorter period of time than the time taken by the median string algorithm. As the l-mer size increases, the proposed system takes much less time than the median string algorithm. For l-mer size 7, we found the proposed system is 47 times faster than the median string algorithm.**

Keywords—DNA,consensus, motif,chapman-kolmogorov

I. INTRODUCTION

Gene regulatory binding motifs are short DNA sequences that control gene expression. The length of these motifs, are 6-15 base pare long. These motifs do not contain any specific starting and end point. Gene regulatory binding motifs can start from anywhere within the DNA sequence. Since the starting point is unknown, the end point also is unknown. The only difference between regulatory motif and random sequence of an equivalent length is that, regulatory motifs occur more frequently than random sequence. The location of regulatory motif may vary from sample to sample. Frequency of these motifs, also vary from sequence to sequence. To discover such regulatory motif from a DNA sequence involves finding such a continuation of pattern [1]. This can be illustrated in Fig. 1.

Any DNA sequence is a sequence of four symbol like a, c, g, and t. Searching for a particular motif in a DNA sequence involves searching a short pattern over the alphabet ∑={a,c,g,t}. The main idea of motif search can be divided into two parts a) exploitation of appropriate model for representation of motif sequence b) formulation of algorithm for motif search. The most widely used motif models are position weight metrics[2] and consensus sequence[3].Position weight metric typically use statistical method, report in a short time but there is no guarantee about a global optimum[4,5,6]. Exact algorithms use consensus model.

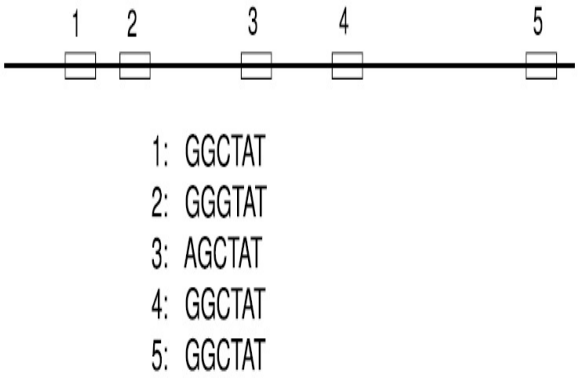


Fig. 1. Consensus string.

Most of the exact algorithms takes all string patterns of length l over alphabet ∑ as candidate motifs, and output the common pattern in all input sequences. In case exact algorithms, initially the search space is $O(|\Sigma|^l)$, that grows dramatically with the rise of $|\Sigma|$. As a result, most exact algorithms are developed for sequences where |∑|=4. But in case of protein data set where |∑|=20, these algorithms cannot find a low conserved motif within a reasonable time.

The idea of motif stem increased the potency of exact algorithms over large alphabet[7]. Consensus string (or closest string or center string) issues are to search out a representative string of a given set S of strings. The consensus issues are major issues in multiple string comparison and have been researched extensively [8,9,10,11,12,13,14,15] to unravel several issues arising in computational biology like motif finding, PCR primer style, and genetic probe style. Since most of the consensus issues are NP-complete, researchers even have designed fixed-parameter algorithms [10,16,17,18] approximation algorithms [13,17,19,20,21,22,23] and algorithms for a short range of strings [10,24,25].

II. MEDIAN STRING ALGORITHM FOR MOTIF SEARCH

Median string algorithm takes as input a set of t DNA sequences. It then generates all possible *l-mers* of length l. Since there are only 4 types of nucleotides in a DNA sequence, the number of all possible *l-mers* is 4^l. Then it starts from the start of each sequence, goes to the end by placing each *l-mer* and calculate distance. In this way for each *l-mer* ,

the minimum distance between the *l-mer* and each row is calculated. The sum of minimum distance of all the rows for a particular *l-mer* is the minimum distance of the *l-mer* with the DNA sample. In this way, the minimum distance for all the *l-mers* is calculated. The *l-mer* with lowest distance is the consensus string. The algorithm is listed below:

Median_string_search()

```
{inputs: DNA,t,n,l
 output: bestword
 procedure:
 MedianStringSearch (DNA, t, n, l)
        bestWord ←AAA…A
        bestDistance ←∞
           for each l-mer s from AAA…A to TTT…T
                 if TotalDistance(s, DNA) < bestDistance
                    bestDistance←TotalDistance(s, DNA)
                    bestWord ← s
          return bestWord
}
```

In median string algorithm, all the 4^l combination of *l-mers* need to be examined. When the length of *l* increases, the number of candidate *l-mers* increase according to 4^l the execution time also increases. But all the 4^l *l-mers* are not significant. A significant amount of the generated 4^l *l-mers,* will not be the consensus. Even some candidate motifs can be found in no sequence. Those motifs got generated as a result of permutation and combination. For those motifs too, median string still perform calculation and contribute to the time complexity. In this paper, we introduced chapman-kolmogorov relation to exclude insignificant candidate motifs and reduce the time complexity of median string algorithm.

III. Chapman-Kolmogorov Relation

A time series of random variables is said to be a markov chain if it has the following property:

$$P(X_t = x_t | X_0 = x_0, X_1 = x_1, \ldots\ldots\ldots, X_{t-1} = x_{t-1}) \qquad (1)$$

It is said to be homogeneous or stationary if it also satisfies

$$P(X_t = z | X_{t-1} = y) = P(X_{t+k} = z | X_{t+z-1} = y) \forall k \geq 0 \quad (2)$$

In other words, the time series of random variable is markovian if the future value of the random variable depend only on the value of random variable has at the present time. The value of the random variable in the future must be independent of the values that it had in the past. We showed that one time series of random variables have these property their behavior can be represented by a graph like the one shown in Fig. 2.:

The possible values that the random variables can have are shown in circles, so in this example the random variables can have the values A, B and C. Instead of talking about the values that the random variables can have though, we state the circles represents the states the system can be in.

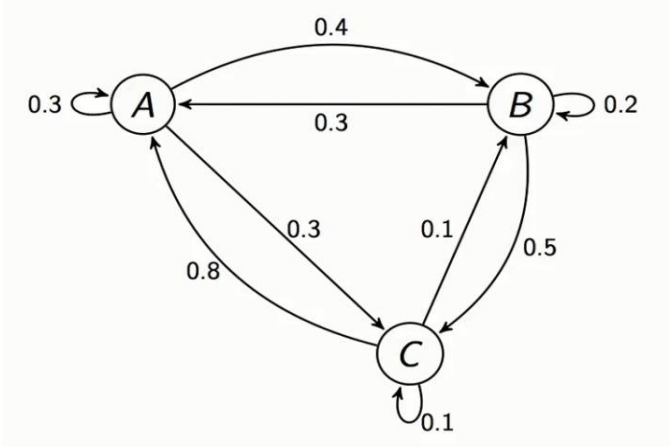


Fig. 2. Markov chain

The circles that represent states are connected by arrows and above these arrows numbers between 0 and 1 appear that represents the probability of moving between the states in a single time step. In this example the probability of moving from state c to state a is 0.8 in a single time step. Alternatively we might state the probability that a random variable has value c followed by a is equal to 0.8. From this markov chain, we can construct a transition probability matrix. We said that the rows in the matrix represent the initial states the system is in and the columns represents the states where the system moves at the next step. The position (a,b) in the matrix shows the probability of moving from state a to state b.

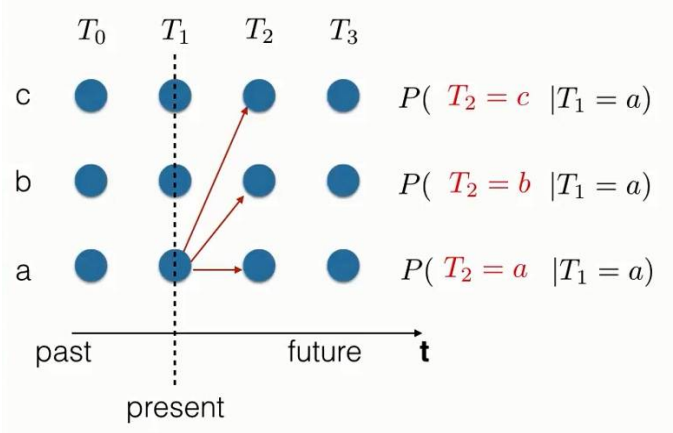


Fig. 3. One step transition

We now know how to obtain probability of T_2 being equal to a,b,c if T_1 =a. We can read this information either from the transition graph or from the one step transition probability matrix. What we would like now is to determine the conditional probability that T_3 takes on these three different values. In other words we would like to know, whether or not we can make prediction about the state of the system after two

or more time steps rather than after a single step. To see how, let us first consider the probability to the following happens in the future. We start with T_1=a, then move to T_2=c then again move to T_3=b. We know that we can get the probability of a single step from a to c from the transition matrix. Once we arrive in state c , i.e T_2=c, the system immediately forgets that T_1 was equal to a. The process is after all markovian and the state of the system was in the past do not affect the probability of the states in the future. Once we have moved from T_1=a to T_2=c, T_2=c becomes the present state of the system. As a consequence of this, we can also get the probability of the second transition T_2=c to T_3=b from the one step transition probability matrix. Because of the markov property, the event of moving from state T_1=a to T_2=c is independent of the event of moving from state T_2=c to T_3=b. The markov property ensures that the system immediately forget everything that happened in the past as soon as that is happened. Thus two adjacent transition are completely independent of each other. The only thing that affect the value of T_3 is the value of T_2. Now remember that we can obtain the probability that two independent event occur by multiplying the individual probability. Consequently for this particular problem, we can calculate the probability that the system will be moving from T_1=a to T_2=c then from T_2=c to T_3=b by multiplying the conditional probability T_3=b given T_2=c.

Let's now suppose, we want to calculate the probability for moving from state a to state b over two steps. There are three ways that can happen. We can go from state a to state c then on to state b. We can go from state a to state b then remain in state b. We can remain in state a for one step and then transfer to state b. The probability of each of the paths is shown in fig.

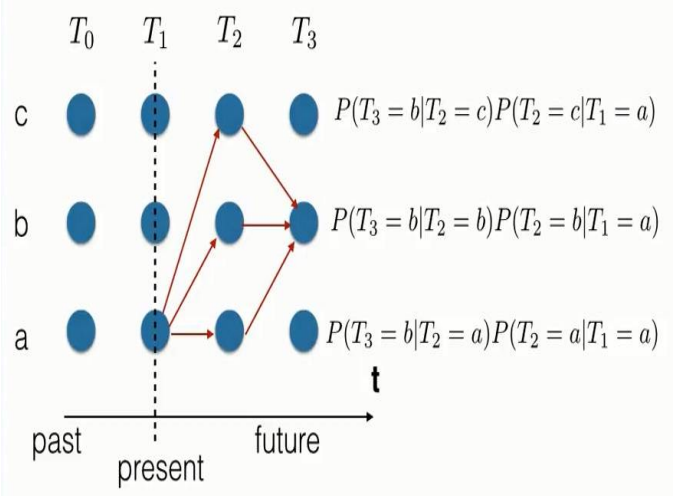


Fig. 4. Two step transition

Each of these pathway is mutually exclusive. As the random variable T_2 can only have one value, it cannot have two value simultaneously. We can thus obtain the total probability of moving from step a to step b over the course of two step by adding the probabilities of each of the paths. If we have only three steps, we can write everything explicitly. If we have more than three steps we can write that as a series notation like Eq

.

$$P(T_3 = b|T_1 = a) = \sum P(T_3 = b|T_2 = x_n)P(T_2 = x_n|T_1 = a) \quad (3)$$

This sum here rounds up all the possible value that the random variable T_2 can take. This sum of product is nothing more than the definition of the matrix product. Because the markov chain is homogeneous, the first matrix and second matrix are identical. We can thus obtain the two step probability matrix by multiplying the one step probability matrix by itself. In general,

$$A^{m+n} = A^m A^n \quad (4)$$

This is what is called chapman kolmogorov relation.

IV. Chapman Kolmogorov Relation Based Median String Algorithm

Consider a database with 10 data samples ($S1,S_2,\ldots\ldots,S_{10}$). Each sample is compose of 80 nucleotides.

TABLE I. DNA Data Set

SID	Sample
S_1	tagtggtcttttgagtgtagatctggagggaaagtatttccaccagttcggggtcacccagcaggcagggtgacttaat
S_2	cgcgactcggcgctcacagttatcgcacgtttagaccaaaacggagttggatccgaaactggagtttaatcggagtcctt
S_3	gttacttgtgagcctggttagacccgaaatataattgttggctgcatagcggagctgacatacgagtaggggaaatgcgt
S_4	aacatcaggctttgattaaacaatttaagcacgtaaatccgaattgacctggtgacaatacggaacatgccggctccggg
S_5	accaccggataggctggttattaggtccaaaaggtagtatcgtaataatggctcagccatgtcaatgtgcggcattccac
S_6	tagattcgaatcgatcgtgtttctccctctggtggttaacgaggggtccgaccttgctcgcatgtgccgaacttgtaccc
S_7	gaaatggttcggtgcgatatcaggccgttctcttaacttggcggtgcagatccgaacgtctctggaggggtcgtgcgcta
S_8	atgtatactagacattctaacgctcgcttattggcggagaccatttgctccactacaagaggctactggtgtgatccgta
S_9	ttcttacacccttctttagatccaaacctgttggcgccatcttcttttcgagtccttgtacctccatttgctctggtgac
S_{10}	ctacctatgtaaaacaacatctactaacgtagtccggtctttcctggtctgccctaacctacaggtcgatccgaaattcg

From this database we can construct a transition matrix like the following:

$$Transition = \begin{array}{c|cccc} & a & c & g & t \\ a & 49 & 39 & 46 & 50 \\ c & 48 & 46 & 39 & 56 \\ g & 40 & 48 & 58 & 53 \\ t & 48 & 55 & 56 & 59 \end{array}$$

Fig. 5. Transition matrix

The transition matrix contains the number of occurance of each of the sixteen possible digrams in the database. For

example Transistion(1,1)=49, that means there are 49 aa in the database. Similarly, Transition(4,4)=59, which means there are 59 tt in the data base. There are total 790 digrams in the database. Dividing all the elements of the matrix by sum of element, produces transition probability matrix like Fig. 6. A[i,j] represents the probability of transition from position *i* to position *j*.

```
A=      a           c           g           t
   a[0.06202532 0.06075949 0.05063291 0.06075949]
   c[0.04936709 0.05822785 0.06075949 0.06962025]
   g[0.05822785 0.04936709 0.07341772 0.07088608]
   t[0.06329114 0.07088608 0.06708861 0.07468354]
```

Fig. 6. Transition probability matrix

In this matrix A(1,1) indicates the probability of finding an a after a, A(2,3) indicates the probability of finding a g after c etc. If we multiply A by itself, we will have a matrix that indicates the probability of occurance of a particular nucleotide at the second step after a particular nucleotide. let us suppose, we are searching for 5 nucleotide sequence i.e our l-mer size is 5. In this case, if we multiply A by itself by 4 times, we will have a matrix like Fig. 7. This matrix contains the probability at fifth position.

```
A⁵=     a           c           g           t
   a[9.18004782e-07 8.28077016e-07 3.32785581e-07 8.28077016e-07]
   c[2.93215942e-07 6.69350669e-07 8.28077016e-07 1.63560335e-06]
   g[6.69350669e-07 2.93215942e-07 2.13306707e-06 1.78980072e-06]
   t[1.01558100e-06 1.78980072e-06 1.35907646e-06 2.32340330e-06]
```

Fig. 7. Transition probability matrix after five step.

This matrix indicates the probability of occurance of a particular nucleotide after 5 step of a particular nucleotide. For example, A(1,1) indicates the probability of occurance of an a after five step of an a. There are 16 elements in the matrix. Summing all the elements and then dividing that value by 16 we will get a threshold value. For the values that are greater than the threshold value, placing a 1 at that location and placing a 0 at other location we will have another matrix like Fig. 8. We can call this a significant matrix S. This matrix tells which transitions are significant and which transitions are not. The significant matrix contains 0's at some positions and 1's at some positions. 0's means insignificant, where as 1's mean significant.

```
S=     a  c  g  t
   a[0. 0. 0. 0.]
   c[0. 0. 0. 1.]
   g[0. 0. 1. 1.]
   t[0. 1. 1. 1.]
```

Fig. 8. Significant matrix

Having the significant matrix, we generate some rules acc ording to the significant matrix. For example, S(2,4)=1 means that the rule c→t is significant, S(3,4)=1 means rule g→t is sig nificant, S(2,3)=0 means rule c→c is insignificant etc. As per t he significant matrix, generated rules are listed in table II.

TABLE II. GENERATED RULES

First element (l_1)	Second element(l_2)	Count	Rule
c	t	56	c→t
g	g	58	g→g
g	t	53	g→t
t	c	55	t→c
t	g	56	t→g
t	t	59	t→t

From these rules, we can generate some motif of length 5. Any motif must start either from a or c or g or t. We will have four trees for motifs starting with four nucleotide. For exampl e one tree is depicted below in Fig. 9.

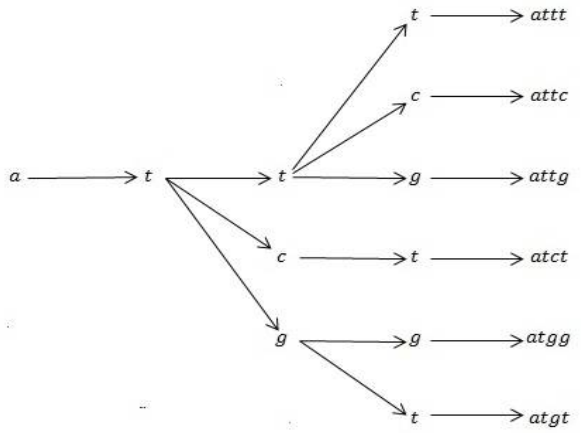


Fig. 9. Motif generation tree

From all the four trees, we will have a short list of motifs which is much shorter than 4^l. We will now use this short set of motifs as the search space of median string algorithm.

V. RESULT AND IMPLEMENTATION

To check the idea behind this concept, we have used Jupyter Notebook as Python programming language. The system configuration is:

- Processor: Intel core i7 CPU
- Clock rate:3.6 GHz
- HardDisk:1000GB
- RAM:8GB

Chapman Kolmogorov relation based median string algorithm accelerates the whole process easy and faster. We have compared the performance of Chapman Kolmogorov relation based median string algorithm with the performance of median string algorithm. In this demonstration we have compared the number of motifs generated by Chapman Kolmogorov relation based median string algorithm and median string algorithms for different *l-mer* size. Since our proposed system produces a only the significant motifs, for all l-mer size the number of motifs generated by the proposed system are found to be shorter than those produced by the

median string algorithms. The scenario is depicted in the table III.

TABLE III. NUMBER OF GENERATED MOTIFS

l-mer Size	Number of Motifs (Chapman Kolmogorov Relation Based)	Number of Motifs (Median String)
2	6	16
3	14	64
4	31	256
5	70	1024
6	157	4096
7	353	16784

The table narrates the results of both the methods used here. The results are the number of motifs generated by both the methods for different *l-mer* size. For *l-mer* size 2, Chapman Kolmogorov relation based system generates 6 motifs whereas Median string algorithm generates 16 motifs. So for *l-mer* size 2, Markov chain based system is (16-6)/6=1.66=166% better than Median string based system.

For *l-mer* size 3, Chapman Kolmogorov relation based system produces 14 motifs and Median string produces 64 motifs. Hence Chapman Kolmogorov relation based systems performance is (64-14)/14=3.57=357% better than median string performance. For *l-mer* size 5 , Chapman Kolmogorov relation produces 70 motifs and Median string produces 1024 motifs, as result Chapman Kolmogorov relation based system is (1024-70)/70=13.62=1362% better than Median string. For last data set the difference between two system reaches 4654%.So it is clear that, as the length of *l-mer* increases, Chapman Kolmogorov relation based system will produce comparatively more less number of motifs than the median string algorithm. The graphical analysis in Fig. 10. also reflects the same impact. The gray line is gradually in upward trend for Median String outcome and the blue line is the Chapman Kolmogorov relation based system outcome.

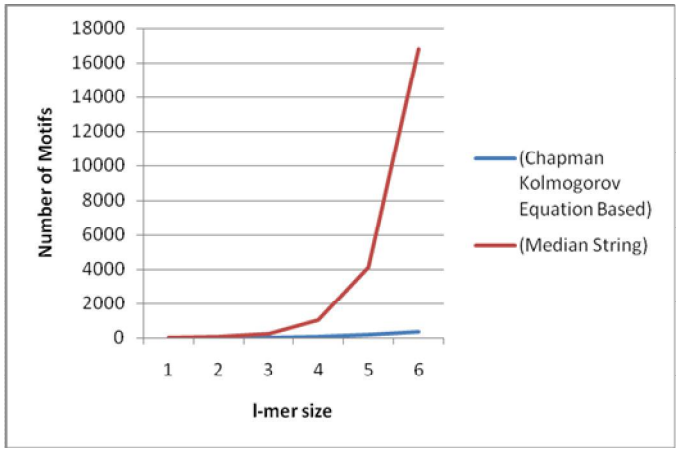


Fig. 10. Number of motifs generated

Same datasets have been imposed on core i7 computer with higher storage and RAM. In that case the processing time of Chapman Kolmogorov relation based system are also less than Median string based processing time.

We have checked every level time consumed by both methods. Median string method takes huge time for DNA consensus. Time consumed by these two methods convey their strengths

TABLE IV. TIME COMPARISION

l-mer Size	Chapman Kolmogorov relation Based Time(ms)	Median String Time(ms)
2	2.57	6.6
3	6.86	30.4
4	16.8	135
5	41.4	590
6	97.4	2590
7	236	11100

Every rows of the table IV demonstrates that Chapman Kolmogorov relation based Median string algorithm consumed less time than Median string orientation. For *l-mer* size 2, the time difference is (6.6-2.57)ms=4.03 ms. It is about 4.03/6.6=61% of the estimated time. That indicates that, Chapman Kolmogorov relation centric Median string finds DNA consensus faster than ordinary Median string algorithm. The time differences have been increased as the length of *l-mer* size increased. From the last row, Chapman Kolmogorov relation based analysis consumed 236 ms whereas ordinary median string algorithm consumed 11100 ms. There are 10864 ms difference for *l-mer* size 7. However it can be applicable for small *l-mer* size too. When *l-mer* size increases, we should consider Chapman Kolmogorov relation based system. In Fig. 11.we can see that, as the l-mer size increases, the time taken by the proposed system is much less than the time taken by the ordinary median string algorithm . Because the proposed system only produces significant motifs.

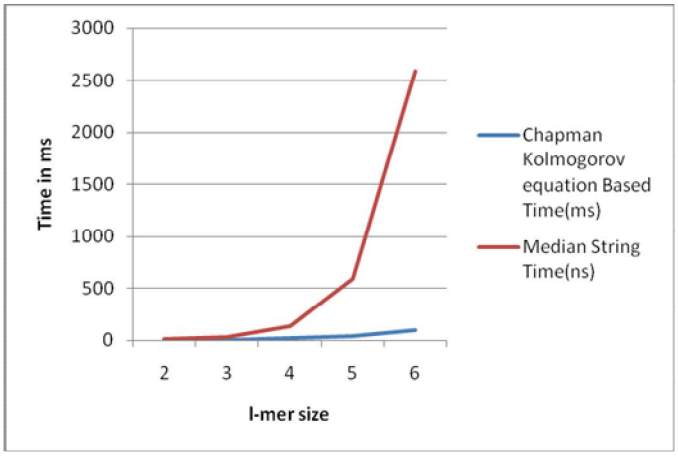


Fig. 11. Time comparision

VI. Conclusion

Median string algorithm generates all the 4^l number of l-mers, then fits each l-mer with the database and calculate the distance. In course of generating the l-mers, median string algorithm, generates significant and insignificant l-mers. Though this process takes much time but this is a certain process. Since we can generate a markov chain from the DNA data set, then also can find the chapman-kolmogorov relation. From this relation we generated a set of rules. Using that rules we produced a reduced set of significant motifs. Since the search space has been reduced, the proposed system takes much less time to find the consensus string. The way we have applied chapman-kolmogorov relation, there might be some other way of application. We believe that, further research may find some other way of application of chapman kolmogorov relation that may reduce the set of rules as well as number of generated motifs. We are working on that ground.

Reference

[1] M.Kellis, N.Patterson, B. Birren, B.Berger, and E. S. Lander, "Methods in comparative genomics: Genome correspondence, gene identification and regulatory motif discovery". *Journal of Computational Biology*,Vol.11,pp.319–355,2004.

[2] J. D.Thompson,D. G. Higgins, and T. J.Gibson,"CLUSTAL W:Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice". *Nucleic Acids Research,* Vol.22, pp. 4673–4680,1994.

[3] T. D. Schneider, "Consensus sequence Zen". *Applied Bioinformatics,* Vol. 1. pp. 111–119, 2002.

[4] C.Lawrence, S.Altschul, M.Boguski, J.Liu, A.Neuwald, and J.Wootton , "Detecting subtle sequence signals: A gibb"s sampling strategy for multiple alignment". *Science*, Vol. 262, pp. 208–214, 1993.

[5] T.Bailey, and C.Elkan, "Fitting a mixture model by expectation maximization to discover motifs in biopolymers". *2nd International Conference on Intelligent Systems for Molecular Biology,* 14-17 August,Stanford,CA, pp. 28–36, 1994.

[6] Y.Zhang, H.Huo, and Q.Yu, "A heuristic cluster-based em algorithm for the planted (l, d) problem". *Journal of Bioinformatics and Computational Biology*, Vol. 11, no. 4, pp. 1350009, 2013.

[7] P.Kuksa, and V.Pavlovic, "Efficient motif finding algorithms for large-alphabet inputs". *BMC Bioinformatics,* Vol. 11, no. Suppl 8, p. S1, 2010.

[8] S.Altschul, and D.Lipman, "Trees, stars, and multiple sequence alignment". *SIAM Journal on Applied Mathematics*, Vol.49, pp.197–209, 1989.

[9] J.Gramm, F.Hüffner, and R.Niedermeier, "Closest strings, primer design, and motif search". *Proceedings of the Sixth Annual International Conference on Computational Biology, RECOMB 2002,*18-21 April, Washington, DC, USA,pp. 74–75, 2002.

[10] J.Gramm, R.Niedermeier, and P.Rossmanith, "Exact solutions for closest string and related problems". *Proceedings of the 12th International Symposium on Algorithms and Computation,* 19-21 December,Christchurch,Newzealand, pp. 441–453, 2001.

[11] R.M.Karp, "Mapping the genome: some combinatorial problems arising in molecular biology".*Proceedings of the 25th Annual ACM Symposium on Theory of Computing,* 16-18 May, San Diego,CA,USA,pp. 278–285, 2001.

[12] M.Li, B.Ma, and L.Wang, "On the closest string and substring problems". *Journal of the ACM,* Vol.49, no.2, pp.157–171, 2002.

[13] H.Mauch, M.J.Melzer, and J.S.Hu, "Genetic algorithm approach for the closest string problem". *Proceedings of the 2nd IEEE Computer Society Bioinformatics Conference,*pp. 560–561, 2003.

[14] C.N.Meneses, Z.Lu, C.A.S.Oliveira, and P.M. Pardalos, "Optimal solutions for the closest-string problem via integer programming". *INFORMS Journal on Computing,* Vol.16, no.4, pp.419–429, 2004.

[15] F.Nicolas, and E.rivals, "Complexities of the centre and median string problems". *Proceedings of the 14th Symposium on Combinatorial Pattern Matching,* 25-27 June, Michoacan , Mexico, pp. 315–327, 2003.

[16] J.Gramm, R.Niedermeier, and P.Rossmanith, "Fixed-parameter algorithms for closest string and related problems". *Algorithmica*, Vol.37, no.1, pp.25–42, 2003.

[17] B.Ma, and X.Sun, "More efficient algorithms for closest string and substring problems". *Proceedings of the 12th Annual International Conference on Research in Computational Molecular Biology,*30 March-2 April, Singapore, pp. 396–409, 2008.

[18] N.Stojanovic, P.Berman, D.Gumucio, R.Hardison, and W.Miller, "A linear-time algorithm for the 1-mismatch problem". *Proceedings of the 5th International Workshop on Algorithms and Data Structures,* 6-8 August, Nova Scotia, Canada,pp. 126–135, 1997.

[19] A.Ben-Dor, G.Lancia, J.Perone, and R.Ravi, "Banishing bias from consensus sequences".*Proceedings of the 8th Symposium on Combinatorial Pattern Matching,* Aarhus,Denmark, pp. 247–261, 1997.

[20] L.Gasieniec, J.Jansson, and A.Lingas, "Efficient approximation algorithms for the Hamming center problem". *Proceedings of the 10th ACM-SIAM Symposium on Discrete Algorithms,* 17-19 January, Baltimore, MD, USA, pp. 905–906,1999.

[21] L.Gasieniec, J.Jansson, and A.Lingas, "Approximation algorithms for Hamming clustering problems". *Journal of Discrete Algorithms*, Vol.2, no.2, pp.289–301, (2004).

[22] K.Lanctot, M.Li, B.Ma, S.Wang, and L.Zhang, "Distinguishing string selection problems".*Proceedings of the 10th ACM-SIAM Symposium on Discrete Algorithms,* 17-19 January, Baltimore, MD, USA, pp. 633–642, 1999.

[23] M.Li, B.Ma, and L. Wang, "Finding similar regions in many strings". *Proceedings of the 31st Annual ACM Symposium on Theory of Computing,* pp. 473–482, 1999.

[24] C.Boucher, D.Brown, and S.Durocher, "On the structure of small motif recognition instances".*Proceedings of the 15th Symposium on String Processing and Information Retrieval,* pp. 269–281, 2008.

[25] S.Sze, S.Lu, and J.Chen, "Integrating sample-driven and pattern-driven approaches in motif finding". *Proceedings of the 4th Workshop on Algorithms in Bioinformatics,* pp. 438–449, 2004.

Performance Analysis and Benchmarking of Clustering Algorithms with gene datasets

Meskat Jahan
Computer Science and Engineering
Rangamati Science and Technology University
Rangamati, Bangladesh
meskat.jahan@gmail.com

Mahmudul Hasan
Computer Science and Engineering
Cumilla University
Cumilla, Bangladesh
mhasanraju@gmail.com

***Abstract*— Clustering is the identification of similar data from a rough or scaled or transformed data and grouping into clusters. Cluster shows symmetry and asymmetry of data and its relations. In this paper, comparisons of three fuzzy clustering algorithms and two conventional clustering algorithms are represented. The analysis is conducted on four datasets which include three gene expression datasets. Here, clustering performance is evaluated using both internal and external validation measurements and an attempt for searching the optimum number of the cluster has taken. This analysis provides an effective way of selecting a suitable algorithm for a particular dataset among different hardcore and soft-core clustering approaches.**

Keywords*— *Cluster analysis; Fuzzy C-Means; K-Means; K-Median; GG; GK;

I. INTRODUCTION

Clustering is the separation of the elements of a dataset in different groups by ensuring similarity in a group and dissimilarity between groups. For categorizing the data, firstly characterization and variation in their attributes are studied. This task is challenging for unknown datasets. Hard and soft clustering algorithms have been developed based on distance measurement techniques. Established algorithms for finding out the behaviour of the datasets, extracting meaningful information and to reveal the underlying hidden pattern are hierarchical clustering, partitioning clustering, clustering based on density and clustering based on grid etc. The challenge is in choosing a suitable algorithm for a given dataset for an optimal and feasible solution. Choosing the optimal number of clusters for a particular combination of algorithm and dataset is another challenge. The optimal number of cluster varies from the dataset to dataset and even a particular dataset may produce different results with a variety cluster numbers. So, there is a need to compare the performances of clustering techniques on various datasets to find out their behaviour.

In this paper, a comparative analysis of hard and soft clustering algorithms are presented. K-Means and K- Median are the representative of hardcore clustering and Fuzzy C-Means, Fuzzy K-Means, Gath- Geva clustering (GG) and Gustafson- Kessel algorithm (GK) are of soft or fuzzy clustering. Wisconsin Breast Cancer, Iris, Leukemia and Motor Cycle datasets have been used for performance evaluation. Structural nature of a dataset is revealed by accurate clustering. Cluster accuracy is ensured by minimizing the classification error rate, picking the optimum number of clusters and validated by validation measurement techniques. The following sections describe the comparative studies of selected clustering algorithms.

II. LITERATURE REVIEW

According to Daxin Jiang, Chun Tang, Aidong Zhang [1], the clustering process is useful for subtypes of cells, gene function, and gene regulations. The fuzzy clustering is used widely on the real-time dataset, multivariate dataset, gene expression dataset etc. [1]. For the analysis of Breast Cancer dataset, a comparative study between K- Means and Fuzzy C-Means was conducted by A Sathish and J Mohana Skundaram [2]. According to Anil K. Jain, clustering has a rich and long history in different scientific fields [3]. In spite of the fact that almost 50 years ago K-Means was proposed and since then, thousands of clustering algorithms have been published but K-Means is still in use [3]. This expresses the difficulty of designing a clustering algorithm whatever it is a general purpose or not and the ill-posed problem of clustering using Different distance measurements such as Mahalanobis distance based fuzzy clustering was established by Hsiang-Chuan Liu, Bai-Cheng Jeng, Jeng-Ming Yih, and Yen-Kuei Yu [4]. Ms Aparna K and Dr Mydhili K Nair [6] studied different partition clustering techniques which provide a glance at the recent problems. According to A. J. Patil, C. S. Patil, R. R. Karhe and M. A. Aher [7], K-Means and FCM, both are sensitive to the outliers and have higher noise ratio. Advancements in fuzzy clustering were done by introducing a mapping function in KFCM (Kernel Fuzzy C-Means). They compared the K-Means algorithm with the FCM algorithm. To overcome the limitations of FCM, they studied Kernel Fuzzy C-Means (KFCM) algorithm. Neha Soni and Amit Ganatra [8], also conducted a comparative study of clustering algorithms. V. Leela, K. Sakthi Priya and R. Manikandan [9], gives comparative studies based on Iris dataset. Shraddha K. Popat and Emmanuel M. [10], illustrated the various clustering and similarity measurement techniques and the types of clustering techniques. O. A. Mohamed Jafar and R. Sivakumar [11] also conducted a comparative study based on hard and fuzzy clustering algorithms with validity indices. The Fuzzy C-Means algorithm is a soft version of K-Means algorithms and Fuzzy K-Means is the modified version of K-Means. Besides, Gath-Geva clustering algorithm (GG) and Gustafson- Kessel clustering algorithm (GK) are also available. Jyotismita Goswami [12] found the suitability of the application of the K-Means algorithm. The K-Means algorithm falls in one of the partitioning methods [12]. The K-Medoids is another partitioning algorithm [12]. Others variations of K-Means and related algorithms are Minimum Spanning Tree Algorithm, Bond Energy Algorithm, GA based clustering etc. Few variations of the partitioning

algorithms are Genetic K-Means Algorithm (GKA), Fast Genetic K-Means Algorithm (FGKA), Incremental Genetic K-Means Algorithm (IGKA), Genetically Guided Algorithm (GGA), Hybrid Genetic-based clustering algorithm (HGA-clustering), Semi-Supervised clustering algorithm, K-Means Fast Learning Artificial Neural Network (KFLANN), Single Program Multiple Data (SPMD) and Genetic Weighted K-Means Algorithm (GWKMA). Nidhi Grover [13] illustrated various fuzzy clustering algorithms like Fuzzy C-Means Algorithms (FCM), Fuzzy Possibilistic C-Means Algorithms (FPCM) and Possibilistic C-Means Algorithm (PFCM) along with suitability of the methods. PCM enables noisy data clustering but it is very sensitive for initialization. To overcome the problems of Possibilistic C-Means, Fuzzy Possibilistic C-Means algorithms, Pal and Bezdek integrated features of PCM & FCM and FPCM, PFCM algorithms were proposed by pal [14]. Q. Huang, X. Huang, Z. Kong, X. Li and D. Tao[15], developed an evolutionary biclustering algorithm BP-EBA which consist of two phases where the first one is rows and columns evolution and the second one is bicluster evolution. Lu, H., Chen, J., Yan, K., Jin, Q., Xue, Y. and Gao, Z. [16], developed a hybrid algorithm by bringing together the Mutual Information Maximization (MIM) and the Adaptive Genetic Algorithm (AGA).

III. ALGORITHMS AND DATASETS

K- Means, K-Mediods, Fuzzy C-Means, Gustafson-Kessel and Gath-Geva clustering algorithms have been used for the analysis of Wisconsin Breast Cancer, Leukemia, Iris and Motor Cycle dataset[5]. The Principal Component Analysis is used for normalizations, visualization and reducing the dimensionality of the dataset. By applying PCA, the principal components of the dataset are extracted out.

IV. DATA PREPROCESSING

Data preprocessing procedure is cleaning the data from unwanted noise and transform it into a workable format. Normalization procedure is used for preprocessing data. Here, the real-time multivariate datasets are used, some of the data values may be missing and all of the datasets are not scaled well. To scale all the data, PCA is applied for normalization of the dataset.

V. TIME COMPARISON STUDY ON CLUSTERING ALGORITHMS

This section represents the comparison among selected clustering algorithms based on the relation of execution times and the number of clusters.

From Fig. 1, FCM requires less time than other algorithms. Other algorithms take much more times than FCM. So, the FCM clustering algorithm provides a better solution for Wisconsin Breast Cancer dataset.

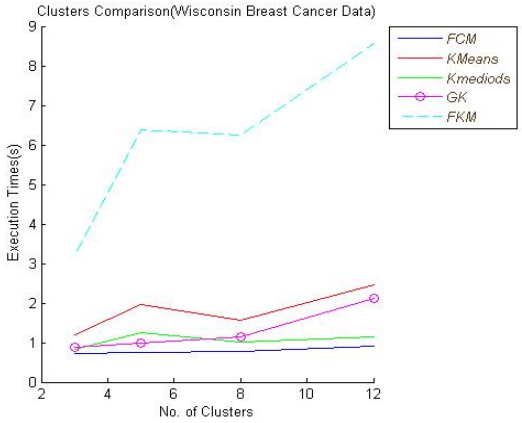


Fig. 1. Algorithms Execution Time comparison on the Wisconsin Breast Cancer dataset.

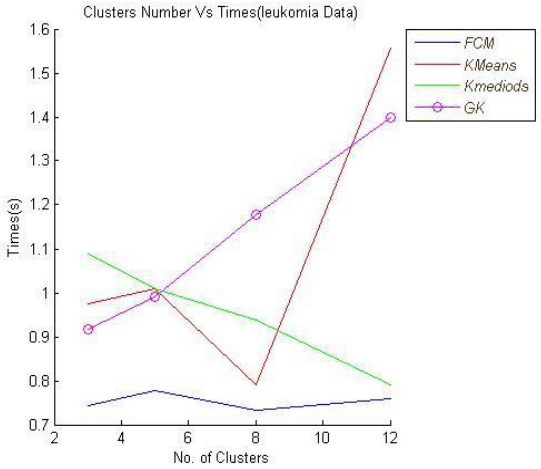


Fig. 2. Algorithms Execution Time comparison on the Leukemia dataset.

From the performance comparison graph of Leukemia dataset in Fig.2, the time requirement for K-Mediods algorithm decreases with the increase of the number of clusters and time requirement graph for GK algorithm is opposite. On the other hand, FCM requires shorter times for execution and does not change much with the increase of cluster number. So, the FCM algorithm suits for Leukemia dataset.

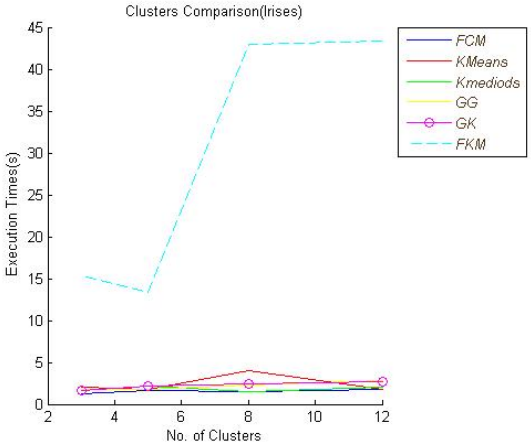


Fig. 3. Algorithms Execution Time comparison on the Iris dataset.

In Fig. 3, Except FKM (Fuzzy K-Means) other algorithms, require nearly the same time for execution with the change of the number of clusters.

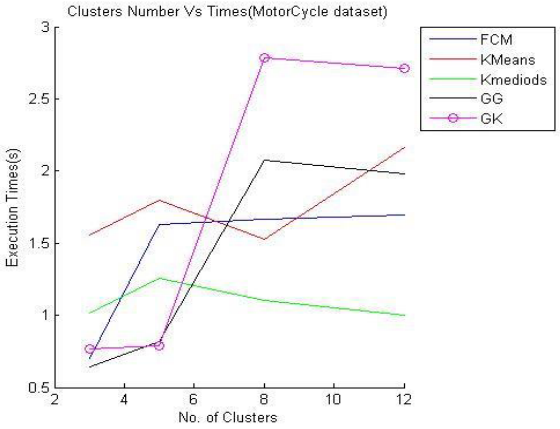


Fig. 4. Algorithms Execution Time comparison on the Motor Cycle dataset.

In Fig.4, FCM, GG and GK algorithm takes longer time for execution with the increased number of clusters. Time requirement varies less with the number of clusters for K-Means and K-Mediods algorithms and K-Mediods takes less time. So, for Motor Cycle dataset, K-Mediods algorithm provides an acceptable solution.

VI. Cluster Validation

Cluster performance has been evaluated based on Partition Coefficient (PC), Classification Entropy (CE), Separation Index(S), Partition Index (SC), Xie and Beni's Index (XB), Dunn's Index (DI), Alternative Dunn Index (ADI) techniques. For WBC and Motor Cycle dataset considered the number of clusters is three and, for Iris and Leukemia dataset it is two. Comparative study on validation indexes of selected algorithms for the four datasets is stated below:

TABLE I. Validation Indexes of Wisconsin Breast Cancer dataset.

Algorithm	PC	CE	SC	S	XB	DI	ADI
FCM	0.84	0.294	0.951	0.002	2.572	0.102	0.000
GK	0.88	0.208	0.579	0.001	2.217	0.071	0.000
K-Means	1.00	NA	0.698	0.002	2.655	0.093	0.000
K-Mediods	1.00	NA	0.729	0.002	INF	0.093	0.000

TABLE II. Validation Indexes of Motor Cycle dataset.

Algorithm	PC	CE	SC	S	XB	DI	ADI
FCM	0.736	0.489	1.059	0.009	3.010	0.011	0.005
GK	0.774	0.416	1.459	0.016	4.829	0.032	.0017
GG	0.961	NA	NA	NA	NA	0.015	0.004
K-Means	1.000	NA	0.879	0.008	5.397	0.037	0.003
K-Mediods	1.000	1.000	NA	1.365	0.015	INF	0.025

TABLE III. Validation Indexes of Iris dataset

Algorithm	PC	CE	SC	S	XB	DI	ADI
FCM	0.88	0.214	0.762	0.005	6.29	0.095	0.011
GK	0.92	0.163	0.606	0.004	16.65	0.188	0.006
GG	0.99	.0004	0.991	0.007	1.906	0.188	0.002
K-Means	1.00	NA	0.615	0.004	41.18	0.098	0.002
K-Mediods	1.00	NA	0.612	0.004	INF	0.098	2e-17

TABLE IV. VALIDATION INDEXES OF LEUKEMIA DATASET

Algorithm	PC	CE	SC	S	XB	DI	ADI
FCM	0.9	0.19	0.64	0.009	5.15	0.381	0.029
GK	0.96	0.09	0.54	0.007	4.05	0.381	0.115
GG	1.0	8 e-9	1.71	0.024	2.66	0.368	0.045
K-Means	1.0	NA	0.55	0.008	4.74	0.381	0.033
K-Mediods	1.0	NA	0.57	0.008	INF	0.368	0.022

Finding the optimal number of cluster is another study goal. Clustering is an unsupervised task and it is sometimes impossible to get the optimal number of cluster for the unknown datasets. In this paper, above-mentioned cluster validation techniques are applied for searching the optimal number of cluster. The higher values of PC and lower values of CE indicates the optimal number of cluster. The lower values of SC, S & XB indexing represents better clustering performance and the highest values of DI and ADI indicates the accurate number of clusters. So, upon considering all the mentioned validation criteria on the above-selected algorithms, attempts are taken to find the optimal number of cluster for the selected datasets. The outcomes are three clusters for WBC dataset, two clusters for Leukemia dataset, two clusters for Iris dataset and three clusters for Motor Cycle dataset.

VII. Cluster Validation Comparison

Study on validation results versus the number of the clusters on the applied clustering algorithm for the different dataset is as follow:

A. WBC dataset

Fig. 5 shows that the FCM algorithm provides a feasible solution and K-Mediods provides the highest accuracy for WBC dataset. Because FCM curves become steady after riching pick values and other curves are stepping downward.

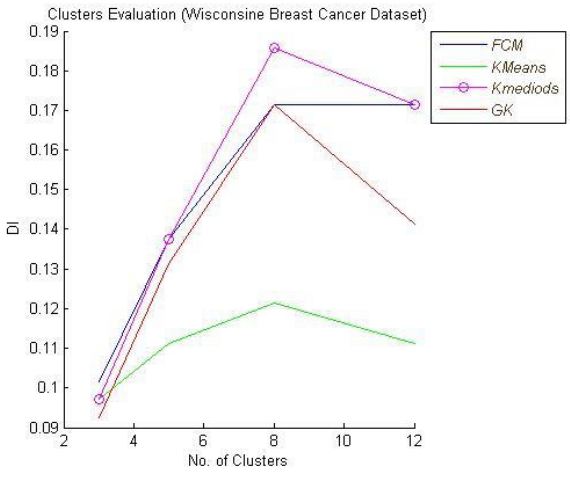


Fig. 5. Validation comparison using DI index on WBC dataset.

B. Leukemia dataset

Fig.6. reveals, the FCM algorithm provides a better result than other algorithms for the Leukemia dataset. Because FCM curves are increasing which shows better results and others are not following any pattern.

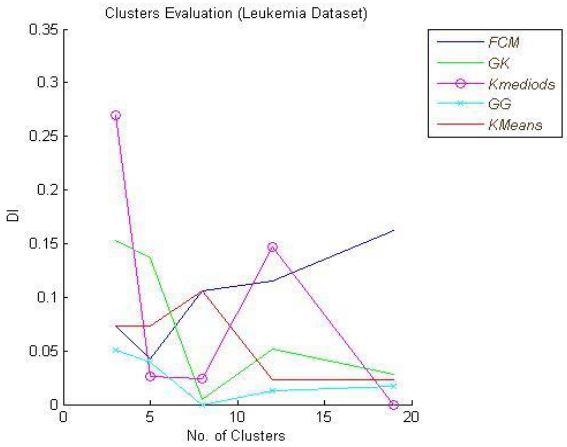


Fig. 6. Validation comparison using DI index on Leukemia dataset.

C. Iris dataset

From Fig.7, GG algorithm provides a better result than other algorithms for the clustering of the Iris dataset. Because only GG curve provides higher values than other curves which represent better clustering.

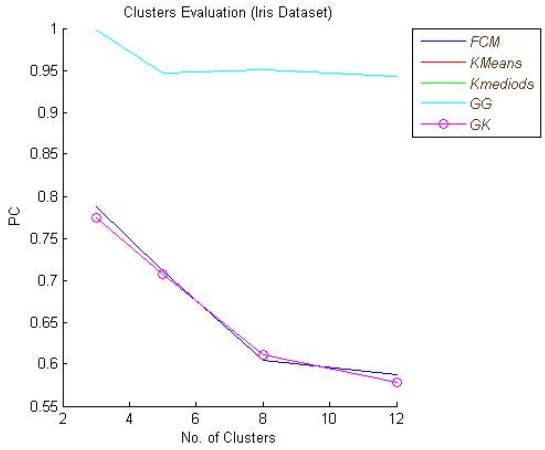


Fig. 7. Validity comparison using PC index on Iris dataset.

D. Motor Cycle dataset

From Fig.8, FCM algorithm provides a better result than other algorithms on the Motor Cycle dataset because only FCM curve shows increasing values of DI, which indicates good cluster. Rest of the algorithm curves does not follow any defined pattern.

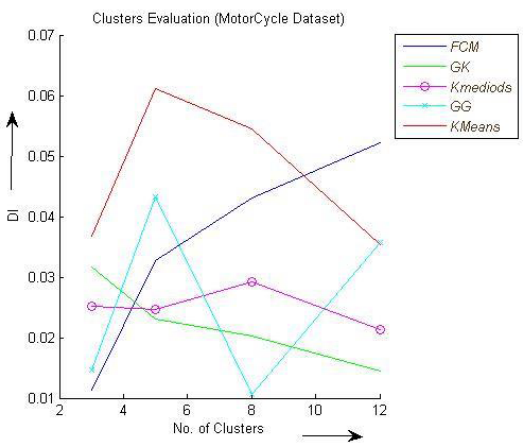


Fig. 8. Validity comparison using DI index on Motor Cycle dataset.

VIII. Analysis of Validation Results

Table I to Table IV shows validation indexing values for different clustering algorithm for selected datasets. Study on validation criteria on the selected datasets for finding their optimal values has been represented in Fig. 5 to Fig. 8. If the clustering is performed accurately, the output will be optimum. In the case of hard clustering, PC is always constant indicating the behaviour of hard partitions. Comparative studies are focused on the values of PC because of the selected datasets are overlapping in nature and PC provides the fuzzy partition for these datasets. Performance of algorithms on selected datasets based on execution time and validation indexes evaluated as - For WBC dataset, FCM clustering algorithm is best if we consider less execution time. But, if we consider the Dunn Indexing as the validation performance criteria, then K-Mediods algorithm produces the highest accuracy. For Leukemia dataset, the FCM algorithm is best by considering both criteria. For Iris dataset, GG algorithm is preferable by considering both criteria. For Motor Cycle dataset, K-Mediods is the suitable algorithm from time comparisons but Dunn Indexing validation measurements indicate the K-Means algorithm.

IX. Conclusion

In this paper, a comparative study among five algorithms which are widely categorized by hard and soft clustering schemes for four datasets have been covered based on executions times and cluster validation techniques. This reveals an effective way of selecting a suitable clustering algorithm for the particular dataset. For WBC dataset, fuzzy clustering gives more accurate result than hard clustering. When the number of clusters is increased, the PC index of FCM indicates fluctuation in results and for lower values, it increases. So for the higher number of clusters, the performance becomes poor. In GK algorithm, the result is constant after the four cluster is obtained. In case of Iris dataset, PC index of FCM shows justifiable results for two clusters. For Leukemia dataset, the highest and lowest values of PC indexing indicates good result for two clusters. From the analysis of Motor Cycle dataset using GK and FCM algorithms, PC shows the highest values near three and also the lowest value of CE near three indicates three number of cluster produce an acceptable solution.

References

[1] Daxin Jiang, Chun Tang, Aidong Zhang, "Cluster Analysis of Gene Expression Data: A Survey", IEEE Transactions on Knowledge and Data Engineering, Volume: 16, Issue: 11, Nov.2004, Page(s):1370 – 1386.

[2] A Sathish, J Mohana Sundaram, "A Comparative Study on K-Means and Fuzzy C- Means Algorithm for Breast Cancer Analysis", International Journal of Computational Intelligence and Informatics, Vol. 4, No.1, April - June 2014.

[3] Anil K. Jain, "Data clustering: 50 years beyond K-means", ELSEVIER, Pattern Recognition Letters 31 (2010), Pages: 651-666.

[4] Hsiang-Chuan Liu, Bai-Cheng Jeng, Jeng-Ming Yih, and Yen-Kuei Yu, "Fuzzy C- Means Algorithm Based on Standard Mahalanobis Distances", International Symposium on Information Processing (ISIP'09)Huangshan, P. R. China, August 21 - 23, 2009, Pages:422-427, ISBN 978-952-5726-02-2

[5] Lichman, M. (2013). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

[6] Ms Aparna K, Dr Mydhili K Nair, "A Detailed Study and Analysis of different Partitional Data Clustering Techniques", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, Issue 1, January 2014, Pages: 8353- 8359, ISSN: 2319-8753.

[7] A.J.Patil, C.S.Patil, R.R.Karhe, M.A.Aher, "Comparative Study of Different Clustering Algorithms", International Journal of Advanced Research in Electrical, Electronics & instrumentation engineering, ISSN (Print): 2320 – 3765, Vol. 3, Issue 7, July 2014, DOI: 10.15662/ijareeie.2014.0307015, Pages:10491-10497.

[8] Neha Soni, Amit Ganatra, "Comparative study of several Clustering Algorithms", InternationalJournal of Advanced Computer Research, ISSN (print): 2249-7277, ISSN (online): 2277-7970, Volume-2, Number- 4, Issue-6, December-2012, Pages: 37-41.

[9] V. Leela, K. Sakthi Priya and R. Manikandan, "Comparative Study of Clustering Techniques in Iris Data Sets", World Applied Sciences Journal 29 (Data Mining and Soft Computing Techniques): 24-29, 2014, ISSN 1818-4952, DOI:10.5829/idosi. Wasj.2014.29.dmsct.5

[10] Shraddha K.Popat, Emmanuel M., "Review and Comparative Study of Clustering Techniques", International Journal of Computer Science and Information Technologies, Volume: 5 (1), 2014, Pages: 805-812, ISSN: 0975-9646.

[11] O. A. Mohamed Jafar and R. Sivakumar, "A Comparative Study of Hard and Fuzzy Data Clustering Algorithms with Cluster Validity Indices ", Proceedings of International Conference on "Emerging Research in computing, Information, Communication and applications ", ERCICA 2013, ISBN: 9789351071020

[12] Jyotismita Goswami, "A Comparative Study on Clustering and Classification Algorithms", Journal of Scientific Engineering and Applied Science (IJSEAS) - Volume-1, Issue-3, June 2015 ISSN: 2395-3470

[13] Nidhi Grover, "A study of various Fuzzy Clustering Algorithms", International Journal of Engineering Research, ISSN:2319-6890) (online),2347-5013(print) Volume No.3, Issue No.3, pp: 177-181, 01 March 2014

[14] Nikhil R. Pal, Kuhu Pal, James M. Keller, and James C. Bezdek," A Possibilistic Fuzzy C-means clustering Algorithm", IEEE TRANSACTIONS ON FUZZY SYSTEMS, VOL. 13, NO. 4, AUGUST 2005

[15] Huang, Q., Huang, X., Kong, Z., Li, X. and Tao, D., 2018. Bi-Phase Evolutionary Searching for Biclusters in Gene Expression Data. IEEE Transactions on Evolutionary Computation.

[16] Lu, H., Chen, J., Yan, K., Jin, Q., Xue, Y. and Gao, Z., 2017. A hybrid feature selection algorithm for gene expression data classification. Neurocomputing, 256, pp.56-62.

Linear Regression Models for Total Body Water Measurement Based on Bioelectrical Impedance Analysis

Mst. Farzana Khatun, Md. Sohel Rana and Ajay Krishno Sarkar
Department of Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology, Bangladesh
Email: farzana0071@gmail.com

***Abstract*—This work proposes new linear regression models for total body water (TBW) measurements for both male and female human being based on bioelectrical impedance analysis. In this research total 2817 (1397 male and 1420 female) data have been used. Age, height, bioelectrical impedance at 100 kHz frequency, body mass index (BMI) have been used for the development of mathematical models. The proposed models have been analyzed statistically and the results show that the correlation (Pearson) coefficients are 0.998 ($p<0.001$) for male and 0.997 ($p<0.001$) for female individuals which denote excellent matching with actual data. Besides intervals of LOA is only -1.36 L to 1.59 L and -1.17 L to 1.12 L for male and female data respectively and most of the data follows such a small limit of agreement. The absolute errors (mean ± Standard Deviation) are (0.64 ± 0.41) L and (0.48 ± 0.32) L for male and female data respectively whereas the bias are 0.11 L for male and -0.02 L for female population only. The RMSE are also very low and which are 0.76 L for male and 0.58 L for female people. Comparing the results of this research with existing models it is seen that the proposed models can be more suitable for TBW measurement.**

***Index Terms*—Bioelectrical impedance analysis, body compositions, linear regression, total body water**

I. Introduction

Fluid content in blood, tissue, bones and elsewhere is defined as body water in physiology. Total body water (TBW) is the summation of the water content of various fluid compartments along with different ratios making significant portion of human body, both by weight and by volume. The average percentage of total body water with respect to body weight varies according to gender. In general, the amount of total body water in human adult male and female are approximately 60% and 55% respectively [1]. Different parameters like age, height, weight, sex, health condition etc. have great impact on the variation of the normal percentage of water in body. To set the normal range of total body water in different people, numerous statistical analysis were conducted over different community people considering physical parameters like age, weight, height, sex and the figure for water portion by weight was found to be 58±8% for males and 48±6% for females [2]. Health condition is also a prominent factor for the amount of total body water, for example, newborn infants have 93% water whereas some obese people have 15% water by weight [3]. These statistical analysis cannot be fixed for all population due to many reasons like number of people, variations of physical parameters, methodology and so on. Different types disease have effect on the deviation of the normal volume of the total water in the human body such as hydration status, anorexia nervosa (eating disorder), cancer, kidney related diseases etc. [4]–[7].

Bioelectrical impedance analysis (BIA) is the procedure which results the electrical impedance i.e. the opposition to the flow of electric current in the body and this procedure is commonly used for the analysis of body composition non-invasively. The procedures which are not involve in the break or damage of skin, organ and tissues for measurement purposes are called as non-invasive ways. There are several mathematical models for TBW measurement by non-invasive based on both anthropometry and bioelectrical impedance analysis and this models are based on different numbers of people along with comparison with respect to pathological tests. But there are different types of limitations of these models such as insufficient number of data, level of accuracy, disease dependency etc. In this paper, two linear regression models for TBW measurement based on bioelectrical impedance analysis have been developed (one for male and another for female) considering adequate number of data with better accuracy which can be implemented in any health condition.

II. Materials and Methods

A. Subjects

To develop TBW measurement models, 2817 (1397 male and 1420 female) participant's data have been used from National Health and Nutrition Examination Survey (NHANES), 2003-2004 (version 7, updated in July 2016) conducted over US citizen [8]. United States Department of Health and Human Services, Centers for Disease Control and Prevention and National Center of Health Statistics operated this survey in their Mobile Examination Center (MEC) and documented all the results jointly. This survey published different types of pathological tests results along with physical parameters data and electrical data. For model development and validation, deuterium oxide (D_2O) based pathological test for TBW measurement were considered. HYDRA Bio-Impedance Spectrum Analyzer (Model 4200) manufactured by Xitron Technologies, Inc., San Diego, California was used for bioelectrical impedance data collection. The examinations were performed

by trained health technicians maintaining protocol which was written in the Body Composition Procedures Manual on the NHANES website.

B. Pathological Test for TBW Measurement

To determine the amount of TBW, deuterium oxide (D_2O) is used as criterion for pathological test. Using D_2O, TBW can be determined in two ways. In one method describe in [9], a dose of D_2O (1 g/kg of lean body weight, estimated from anthropometric measurements) is given into the body to mix up with blood and then blood samples are taken when steady state is reached after greater than four hours. Urine is simultaneously collected to determine net loss of heavy water as well as plasma water was separated out from the blood samples using an ultrafiltration technique that is performed by centrifugation at 2110×g for two hours in centrifuge tubes with a 5000 D cut-off membrane. The volume of TBW is calculated by following formula:

$$TBW = \frac{1}{N}\sum \frac{M - Lt}{(D_2O)t} \tag{1}$$

where M equals the ingested mass of D_2O (g), $[D_2O]$ equals the serum concentration of D_2O at time t, L equals the sum of urine loss at time t, and N equals the number of samples taken after the four-hour equilibration period.

TBW can be determined using flowing after glow mass spectrometry measurement of deuterium abundance in breath samples from individuals [10]. A known dose of deuterated water (Heavy water, D_2O) is ingested and allowed to equilibrate within the body water. The FA-MS instrument then measures the deuterium-to-hydrogen (D:H) ratio in the exhaled breath water vapor. The total body water is then accurately measured from the increase in breath deuterium content in relation to the volume of D_2O ingested. These pathological tests are invasive which require more time and conditions are applicable.

C. Antropoemtry

Maintaining well-defined standard procedures and well trained persons are prime concern for anthropometric measurements [11]. Weight is measured to closest 0.01 kg by portable electronic scale and height is measured adjacent to 0.01 cm using stadiometer, a portable gauge (Seca, Germany). There are several anthropometric mathematical models for TBW models. Hume *et al*. [11] proposed two models for TBW measurement (one for male and another for female).

$$TBW(male) = -14.01 + 0.19h + 0.30w \tag{2}$$

$$TBW(female) = -35.27 + 0.34h + 0.18w \tag{3}$$

Watson *et al*. [12] also proposed two models for TBW measurement.

$$TBW(male) = 2.45\text{–}0.09a + 0.10h + 0.34w \tag{4}$$

$$TBW(female) = -2.10 + 0.11h + 0.25w \tag{5}$$

Chumlea *et al*. [13] showed another two anthropometric equations for TBW measurement.

$$TBW(male) = -18.37\text{–}0.09a + 0.34w + 0.25h \tag{6}$$

$$TBW(female) = -16.71\text{–}0.05a + 0.22w + 0.24h \tag{7}$$

Medoua *et al*. [14] showed following anthropometric equations for total body water (TBW) measurement considering 40 Cameroonian hemodialysis patients along previous related models, but the accuracy was not satisfactory.

$$TBW = 13.90 + 0.002a + 0.329w + 1.85s \tag{8}$$

Where, a = age in years, w = weight in kg, h = height in cm, s = sex (1 for male and 0 for female).

D. Bioelectrical Impedance Analysis (BIA)

Bioelectrical impedance analysis is a tetra electrodes based method in which two electrodes are connected on right hand and another two electrodes are connected in right foot [5]. In right hand, one electrode is connected on wrist and another is connected on the bone of little finger and in the right foot, one is connected on bone of big toe side of ankle and another is connected on base of the second toe. Two electrodes are for sending constant current to human body and another two electrodes are for measuring voltage drop. Using voltage drop and sending current, bioelectrical impedance analyzer computed different types of impedance related values which are available in the databases.

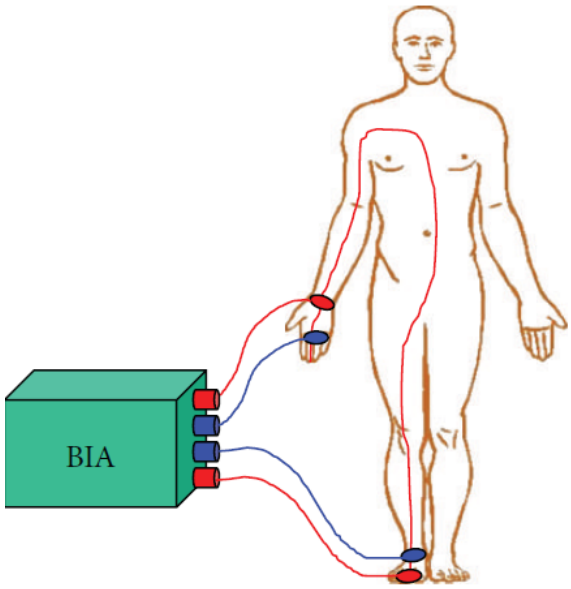


Fig. 1. Bioelectrical Impedance Analysis [5]

E. BIA Based Existing Models for TBW Measurement

Several regression models based on bioelectrical impedance analysis for TBW measurement were also developed. Deurenberg *et al*. [15] showed two mathematical models considering

50 kHz and 100 kHz frequency. For development of following equations, 139 patient's data were used and D_2O was used as pathological criterion.

$$TBW(L) = 6.69 + 0.35h^2/Z_{100} + 0.17w – 0.11a + 2.66s \quad (9)$$

$$TBW(L) = 6.53 + 0.37h^2/Z_{50} + 0.18w – 0.11a + 2.83s \quad (10)$$

Here, h = height in cm, Z_{100} = impedance at 100 kHz frequency, Z_{50} = impedance at 50 kHz frequency, w = weight in kg, a = age, s = sex (1 for men and 0 for women). Cornish $et\ al.$ [16] also developed a mathematical model using 60 participant's data and compared with results found from D_2O based pathological tests. The equation is given below-

$$TBW(L) = 0.6 + 0.50h^2/R_0 + 0.186w \quad (11)$$

Where, h = height in cm, R_0 = resistance at zero frequency in kΩ, w = weight in kg. Later Sun $et\ al.$ [17] developed two models (one for male and another for female) to measure TBW from 1829 individual's data (734 for male and 1095 for female) and the equations are-

$$TBW(male, L) = 1.203 + 0.449h^2/R_{50} + 0.176w \quad (12)$$

$$TBW(female, L) = 3.747 + 0.450h^2/R_{50} + 0.113w \quad (13)$$

Where, h = height in cm, R_{50} = resistance in kΩ at 50 kHz frequency, w = weight in kg. Liu $et\ al.$ [18] proposed a model for TBW measurement considering 50 kHz frequency using 948 individual's data. But this equation was applicable only for children and the equation was-

$$TBW(L) = -4.57 + 0.23h^2/R_{50} + 0.18w + 0.13a + 0.06h + 0.5s \quad (14)$$

Where, h = height in cm, R_{50} = resistance in kΩ at 50 kHz frequency, w = weight in kg, a = age in years, s = sex (1 for male and 0 for female). Then Algago $et\ al.$ [19] developed another model for TBW measurements in adults considering 50 kHz frequency based on 250 participant's data and the equation was-

$$TBW(L) = 5.68 + 0.27h^2/R_{50} + 0.23w – 0.05a + 4.42s \quad (15)$$

Where, h = height in cm, R_{50} = resistance in kΩ at 50 kHz frequency, w = weight in kg, a = age in years, s = sex (1 for male and 0 for female). Recently A. Diouf $et\ al.$ [20] built a mathematical model for TBW measurement using 151 participant's data. But, this equation was also for children (school going children, age 8 to 11 years) which was-

$$TBW(L) = -2.28 + 0.07h + 0.08w + 0.38h^2/Z_{50} – 0.47s \quad (16)$$

In this model, height (h) in cm, weight (w) in kg, sex (s =1 for male 2 for female), impedance (Z_{50}) in kΩ at 50 kHz frequency were considered.

F. Statistical Analysis

For validation and checking the robustness of the proposed model, statistical analysis is major indicator in which the correlation coefficient (Pearson coefficient), 95% confidence interval (95% CI), 95% limit of agreement (95% LOA), bias (difference between actual and measured values), p-value ($p<0.05$ indicate statistical significance) are measured. General characteristics of subject's data are expresses in terms of mean ± standard deviation (SD), 95% CI and p-value. Besides root mean square error (RMSE) is also calculated using following formula.

$$RMSE = \sqrt{\frac{\sum(x - x')^2}{n}} \quad (17)$$

Where, x denotes measured or given value and x' denotes predicted value. All types of statistical analysis have been done for this research using IBM SPSS v25.0 and MATLAB 2015b. The 95% LOA is used in Bland-Altman plot which is used to estimate the agreement between the predicted values and measured values in which limits of agreement have been determined by summing and subtracting the average differences from 1.96 times of standard deviation [21], [22].

III. Results

A. Proposed Linear Regression Models for TBW Measurment

For the development of linear regression models for TBW measurement, data set has been divided into two parts separating male and female and then each data set has been divided into more two parts (one for model development and another for model validation). To develop the TBW measurement model for male people, 720 male participant's data have been used and rest of 677 male participant's data have been used for model validation. On the other hand, 735 female participant's data have been used for TBW model development for female people and rest of 685 female participant's data have been used for model validation. In these proposed models, bioelectrical impedance at 100 kHz frequency along with age, height, body mass index (BMI) have been used and the general characteristics of participant's data have been given in the Table I to Table IV. Using the data sets, two new equations have been proposed for TBW measurement.

$$TBW(male, L) = -11.25 + 0.011a + 0.46b + 0.005Z_{100} + 0.60h^2/Z_{100} \quad (18)$$

$$TBW(female, L) = -10.56 - 0.002a + 0.36b + 0.006Z_{100} + 0.60h^2/Z_{100} \quad (19)$$

Here, a = age in years, b = BMI in $[kg/m^2]$, h = height in cm, Z_{100} = bioelectrical impedance at 100 kHz frequency. After implementing the new developed equations into the validation data sets, it is seen that the predicted values of TBW are close enough to the actual values which have been explained in details in the performance evaluation section.

TABLE I
GENERAL CHARACTERISTICS OF MALE MODEL DATA

Group	n = 720		
	Mean ± SD	***95% CI***	***p-value***
Age ($years$)	21.23 ± 11.22	20.41 to 22.05	<0.001
Height (cm)	167.58 ± 14.87	166.50 to 168.67	<0.001
BMI (kg/m^2)	24.24 ± 5.76	23.82 to 24.66	<0.001
TBW (L)	39.25 ± 12.03	38.37 to 40.13	<0.001
Z_{100} ($k\Omega$)	486.72 ± 89.45	480.18 to 493.85	<0.001

TABLE II
GENERAL CHARACTERISTICS OF FEMALE MODEL DATA

Group	n = 735		
	Mean ± SD	***95% CI***	***p-value***
Age ($years$)	21.01 ± 11.57	20.17 to 21.85	<0.001
Height (cm)	157.21 ± 10.62	156.44 to 157.98	<0.001
BMI (kg/m^2)	25.25 ± 7.16	24.73 to 25.76	<0.001
TBW (L)	29.17 ± 7.71	28.61 to 29.73	<0.001
Z_{100} ($k\Omega$)	566.38 ± 84.44	560.26 to 572.50	<0.001

TABLE III
GENERAL CHARACTERISTICS OF MALE VALIDATION DATA

Group	n = 677		
	Mean ± SD	***95% CI***	***p-value***
Age ($years$)	20.37 ± 10.98	19.54 to 21.20	<0.001
Height (cm)	165.47 ± 15.52	164.29 to 166.64	<0.001
BMI (kg/m^2)	23.44 ± 5.33	23.04 to 23.84	<0.001
TBW (L)	37.14 ± 11.91	36.24 to 38.03	<0.001
Z_{100} ($k\Omega$)	498.84 ± 90.84	491.99 to 505.70	<0.001

TABLE IV
GENERAL CHARACTERISTICS OF FEMALE VALIDATION DATA

Group	n = 685		
	Mean ± SD	***95% CI***	***p-value***
Age ($years$)	21.00 ± 11.35	20.15 to 21.86	<0.001
Height (cm)	157.42 ± 10.29	156.65 to 158.19	<0.001
BMI (kg/m^2)	24.59 ± 6.53	24.10 to 25.08	<0.001
TBW (L)	28.51 ± 7.29	27.96 to 29.06	<0.001
Z_{100} ($k\Omega$)	579.14 ± 85.92	572.69 to 585.58	<0.001

B. Performance Evaluation

The performance evaluation of the proposed models have been done according to the results of the statistical analysis. From the point of view of correlation coefficient (Pearson coefficient), the more closer to unity the more matching will be found among the actual values and the measured values. The proposed mathematical models exhibit the correlation coefficient 0.998 and 0.997 ($p<0.001$) for male and female respectively. Besides, the correlation coefficient, the lower range of 95% CI, 95% LOA of errors are also required along with lower value of bias and absolute error for proper validation of the models. The 95% CI are 0.06 to 0.17 L and -0.07 to 0.18 L for male and female data. The 95% LOA for male data are -1.36 to 1.59 L and for female data are -1.17 to 1.12 L. The values of bias and RMSE are 0.11 L and 0.76 L for male people and -0.03 L and 0.58 L for female data which indicate lower values. The linear correlation curves for both male and female have shown in Fig. 2 and Fig. 3 respectively. TBW (female) Using the averages of actual and predicted

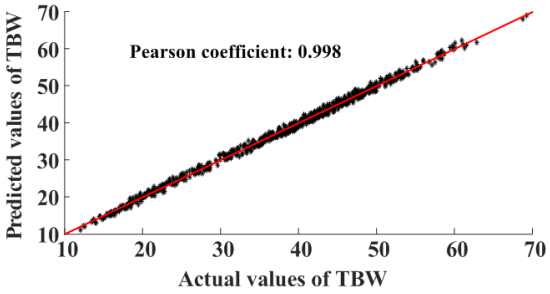


Fig. 2. Correlation curve between actual values and predicted values of TBW (male).

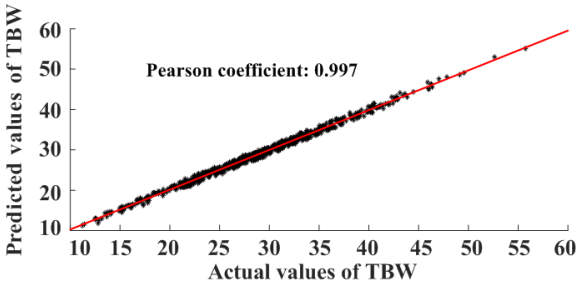


Fig. 3. Correlation curve between actual values and predicted values of TBW (female).

values of TBW and the 95% LOA, Bland-Altman plot has been shown in Fig. 4 and Fig. 5. From the figures, it is seen

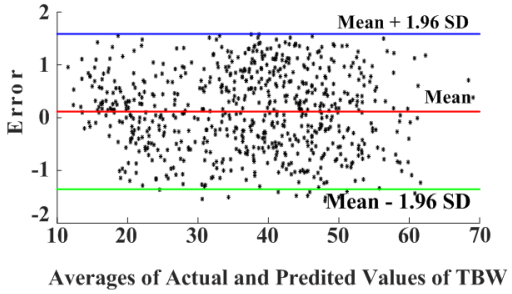


Fig. 4. Bland-Altman plot of TBW (male).

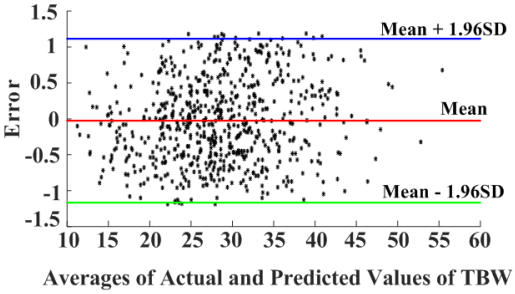


Fig. 5. Bland-Altman plot of TBW (female).

TABLE V
STATISTICAL COMPARISON OF PROPOSED AND CURRENT MODELS (MALE)

Ref. No.	n = 677				
	Correlation Coefficient	***95% LOA (L)***	***Absolute Error (L)***	***Bias (L)***	***RMSE (L)***
Hume *et al.* [11]	0.793	-8.16 to 8.26	3.44 ± 2.39	0.05	4.19
Watson *et al.* [12]	0.794	-12.10 to 7.46	4.56 ± 5.50	-2.32	5.50
Chumlea *et al.* [13]	0.768	-13.91 to 1.26	6.55 ± 3.47	-6.32	7.41
Medoua *et al.* [14]	0.826	-11.12 to -12	4.98 ± 3.20	0.43	5.92
Deurenberg *et al.* [15] (100 kHz frequency)	0.950	-7.80 to 5.42	2.93 ± 2.05	-1.19	3.57
Deurenberg *et al.* [15] (50 kHz frequency)	0.950	-8.16 to 5.13	3.03 ± 2.14	-1.51	3.71
Cornish *et al.* [16]	0.850	-3.4 to 1.00	1.40 ± 0.86	-1.20	1.64
Sun *et al.* [17]	0.840	-3.02 to 2.42	1.17 ± 0.81	-0.3	1.42
Liu *et al.* [18]	0.88	-3.52 to 8.86	3.24 ± 2.58	2.67	4.14
Aglago *et al.* [19]	0.92	-8.32 to 5.64	3.10 ± 2.21	-1.34	3.80
Diouf *et al.* [20]	0.890	-3.42 to 10.77	4.17 ± 3.03	3.67	5.15
Proposed	0.998	-1.36 to 1.59	0.64 ± 0.41	0.11	0.76

TABLE VI
STATISTICAL COMPARISON OF PROPOSED AND CURRENT MODELS (FEMALE)

Ref. No.	n = 685				
	Correlation Coefficient	***95% LOA (L)***	***Absolute Error (L)***	***Bias (L)***	***RMSE (L)***
Hume *et al.* [11]	0.793	-7.06 to 5.32	2.60 ± 1.99	-0.87	3.27
Watson *et al.* [12]	0.794	-7.30 to 2.97	2.89 ± 1.77	-2.16	3.39
Chumlea *et al.* [13]	0.768	-10.71 to 0.49	5.29 ± 2.50	-5.11	5.85
Medoua *et al.* [14]	0.826	-10.16 to -0.13	5.23 ± 2.37	-5.15	5.75
Deurenberg *et al.* [15] (100 kHz frequency)	0.950	-5.48 to 2.07	2.20 ± 1.32	-1.70	2.57
Deurenberg *et al.* [15] (50 kHz frequency)	0.950	-5.60 to 1.84	2.30 ± 2.68	-1.88	2.68
Cornish *et al.* [16]	0.850	-3.4 to 1.00	1.40 ± 0.86	-1.20	1.64
Sun *et al.* [17]	0.840	-3.79 to 1.54	1.51 ± 0.90	-1.12	1.76
Liu *et al.* [18]	0.88	-5.10 to 2.46	1.85 ± 1.42	-1.32	2.33
Aglago *et al.* [19]	0.92	-4.36 to 1.96	1.68 ± 1.11	-1.20	2.00
Diouf *et al.* [20]	0.890	-3.43 to 5.64	1.91 ± 1.71	1.10	2.56
Proposed	0.998	-1.17 to 1.12	0.48 ± 0.32	-0.02	0.58

that the correlation curves show the almost exact matching between actual values and predicted or measured values. On the other hand, the Bland-Altman plots show that the almost all the averages of actual and measured data with respect to errors are present within the 95% LOA. So, it can certainly be said that the proposed regression models can be used for TBW measurement with excellent accuracy.

IV. DISCUSSION

The robustness of the proposed models are also done by comparing the correlation coefficient, absolute errors (mean ± SD), bias with 95% CI, RMSE and 95% LOA found from proposed models and existing models through the data sets used for validation purposes. All the results have been given in Table V and Table VI. The proposed models, developed for male and female people have been compared with existing models (anthropometry and bioelectrical impedance based models) where it is seen that the proposed models show closest matching with actual results as the correlation coefficients are 0.998 and 0.997 for male and female respectively whereas other models show coefficient values below 0.90 except one model. The correlation coefficient found in [15] is 0.950, but the 95% limit of agreement (LOA) is high which is not expected as the lower range of LOA is expected. Besides this model, other models also show larger range of 95% LOA. From the error analysis, it is seen that, the proposed models show lowest absolute error, bias and root mean square (RMSE), numerically the values are 0.64 ± 0.41 L, 0.11 L and 0.76 L respectively for male and 0.48 ± 0.32 L, -0.02 L and 0.58 L respectively for female whereas these values are higher for other models than new ones. The maximum values of different errors have been found in [14] and the minimum have been found in [18] which are greater than the errors of proposed models and hence certainly it is seen that the models proposed in this models can be used for TBW measurement with better accuracy.

V. CONCLUSION

To propose any model, there are some prerequisites such as number of data used, accuracy along with maintaining of the standard range of different types of statistical analysis like correlation coefficient, bias, absolute error, 95% confidence interval of limit of agreement, root mean square error. The models proposed here have fulfilled almost all the required criteria and from the comparison with other existing models it is seen that newly proposed models can easily be implemented for practical implementations with better accuracy. As the measurement of TBW has significant importance for checking nutrition condition, detecting diseases detection in medical sciences, these proposed models will be helpful for accurate

measurement within very short time. Bioelectrical impedance analysis is well known quick and easy non-invasive method. So, proper modeling using bioelectrical impedance will be helpful to render latest medical technology at low cost to the mass people of the world as the large number of population in the world cannot afford the treatment related costs.

ACKNOWLEDGMENT

National Health and Nutrition Examination Survey (NHANES) 2016, United States of America (USA).

REFERENCES

[1] Adrian Mondry, "Total body water content in South East Asian hypertensive patients", *Journal of Hypertension*, vol. 36, no. 3, pp.24, 2018.

[2] W.C. Chumlea *et al*., "Total body water reference values and prediction equations for adults", *Kidney International (Elsevier)*, vol.59, no.6, pp. 2250-2258, 2001.

[3] C. Guyton and J.E. Hall, *Texbook of Medical Physiology (11th Edition)*, Elsevier Saunders Publisher, 2006.

[4] K. Cumming *et al*., "Bioelectrical impedance analysis is more accurate than clinical examination in determining the volaemic status of elderly patients with fragility fracture and hyponatraemia", *Journal of Nutrition, Health and Aging (Springer)*, vol. 18, no. 8, pp. 744-750, 2014

[5] T. K. Bera *et al*., "Bioelectrical Impedance Methods for Noninvasive Health Monitoring: A Review", *Journal of Medical Engineering*, vol. 2014, pp. 1-28, 2014.

[6] J. Simons *et al*., "Bioelectrical impedance analysis to assess changes in total body water in patients with cancer", *Clinical Nutrition (Elsevier)*, vol. 18, pp. 35-39, 1999.

[7] S. R. Abbas *et al*., "Effect of Change in Fluid Status Evaluated by Bioimpedance Techniques on Body Composition in Hemodialysis Patients," *Journal of Renal Nutrition*, vol. 28, no. 3, pp. 183-190, 2018.

[8] United States Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, National Health and Nutrition Examination Survey (NHANES), 2003-2004. Ann Arbor, MI: Inter-university Consortium for Political and Social Research, 2016-07-11.

[9] S. Forest *et al*., "Flow Cytometry of Body Fluid Specimens: Diagnostic Value and Clinical Utility in Patients With or Without Prior Hematologic Malignancy," *American Journal of Clinical Pathology*, vol. 147, no. suppl-2, pp. S178-S179, 2017.

[10] D. J. Wilkinson *et al*., "Internal comparison between deuterium oxide (D_2O) and L-(ring-13C6) phenylalanine for acute measurement of muscle protein synthesis in humans," *Physiological Reports*, vol. 3, no. 7, pp. e12433, 2015.

[11] R. Hume and E Weyers, "Relationship between total body water and surface area innormal and obese subjects", *Journal of Clinical Pathology*, vol.24, pp. 234-242, 1971.

[12] P.E. Watson, I.D. Watson and R. Batt, "Total body water volumes for adult males and females estimated from simple anthropometric measurements", *American Journal of Clinical Nutrition*, vol. 33, pp. 27-39, 1980.

[13] W.C. Chumlea, *et al*., "Total body water volume for white children and adolescents and anthropometric prediction equations: The Fels Longitudinal Study", *Kidney International (ELsevier)*, vol. 68, no. 5, pp. 2317-2322, 2005.

[14] G.N. Medoua *et al*., "Validity of anthropometry- and impedance-based equations for the prediction of total body water as measured by deuterium dilution in Cameroonian haemodialysis patients", *Clinical Nutrtion ESPEN (Elsevier)*, vol. 10, no. 5, pp. 167-173, 2015.

[15] Deurenberg *et al.*, "Multi-frequency impedance for the prediction of extracellular water and total body water", *British Journal of Nutrition*, vol. 73, pp. 349-358, 1995.

[16] Cornish *et al*., "Evaluation of multiple frequency bioelectrical impedance and Cole– Cole analysis for the assessment of body water volumes in healthy humans", *European Journal of Clinical Nutrition (Nature)*, vol. 50, pp. 159-164, 1996.

[17] Sun *et al.*, "Development of bioelectrical impedance analysis prediction equations for body composition with the use of a multicomponent model for use in epidemiologic surveys", *American Journal of Clinical Nutrition*, vol. 77, pp. 331-340, 2003.

[18] A. Liu *et al*., "Validation of bioelectrical impedance analysis for total body water assessment against the deuterium dilution technique in Asian children," *European Journal of Clinical Nutrition (Nature)*, vol. 65, pp. 1321-1327, 2011.

[19] K. E. Aglago *et al*., "Development and validation of bioelectrical impedance analysis equations for predicting total body water and fat-free mass in North-African adults," *European Journal of Clinical Nutrition (Nature)*, vol. 67, pp. 1081-1086, 2013.

[20] A. Diouf *et al*., "Validity of bioelectrical impedance analysis in predicting total body water and adiposity among Senegalese school-aged children", *Journal of Public Library Open Science (PLOS) One*, 2018

[21] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement", *Lancet*, vol.1, pp. 307-310, 1986.

[22] J. M. Bland and D. G. Altman, "Comparing methods of measurement: why plotting difference against standard method is misleading", *Lancet*, vol. 346,pp. 1085-1092, 1995.

Improved Piecewise Constant Approximation Method for Compressing Data Streams

Atik Mahbub
Department of Computer Science & Engineering
East West University
Dhaka, Bangladesh
atikmahbub100@gmail.com

Farhana Haque
Department of Computer Science & Engineering
East West University
Dhaka, Bangladesh
farhanatisha23@gmail.com

Habibul Bashar
Department of Computer Science & Engineering
East West University
Dhaka, Bangladesh
basharewu@gmail.com

Mohammad Rezwanul Huq
Department of Computer Science & Engineering
East West University
Dhaka, Bangladesh
mrhuq@ewubd.edu

***Abstract*— Stream data, obtained from the various sensors and devices with high speed, used daily, are indispensable parts in our day-to-day life. There is an immense necessity to store these data comprising of large volume. Data compression is one of the effective ways of storing data efficiently. In this paper, different data compression techniques such as PCA, APCA have been improved and thus we propose a new hybrid model. In the proposed model, the window size is fixed but data can be compressed efficiently than PCA. Besides, the delay time is minimized compared to the other existing techniques since the model is developed by considering both the delay time and several data points to be compressed. Our experiment shows that the proposed method performs twice as better than PCA in terms of compression and minimizes RMSE by 10%-25% compared to APCA method.**

Keywords—Data Compression, Stream data, Window

I. INTRODUCTION

Stream data has entrusted to be one of the most valuable concepts in Modern information technology. Experts are predicting that the content of stream data will become a huge amount of data that has been collected from social media or any other online resources [5], [12]. As Stream data are ejected from cell phones, vehicles, appliances, buildings, meters, machinery, medical equipment and many other machines continuously [3], it becomes huge data and turnout cumbersome to store this large volume of data. Sometimes new data removes old data due to lack of space in storage. Compression of data activates quick processing of data by accessing queries immediately over compressed data [1], [2], [10] using few I/O as compressed data usually stored using lessen number of disk pages [4], [6]. Our aim is to compress the continuously coming stream data to make the memory space free from over-burdened data [3] and Data should be compressed in such way that compressed value should be close to actual value [9]. The data values are divided into certain segments, which are known as windows and a threshold to support error rate is considered. An improved of Piecewise Constant Approximation (PCA) technique is being followed, where window size of data is fixed-length but it gives optimal result in consideration of delay time and total compressed data points on running process rather than PCA [7], [8].

II. RELATED WORK

Piecewise Constant Approximation (PCA) & Adaptive Piecewise Constant Approximation (APCA) are two exoteric techniques to compress data optimally in model-based compression [9], [10]. Model Based Compression depends on two main properties, which are continuity of the sampling methods or processes and interrelations of different sampling processes [11].

PCA distributes the entire dataset into piecewise fixed length segments, which are known, as windows [7]. For each window, if the difference between the maximum and minimum value of respective windows is less than twice of the maximum error rate threshold than the whole window will represent one single average value. PCA algorithm gives more data points as window size is fixed. However, the delayed time of PCA is low.

APCA is not as straightforward as PCA [8]. Here, the window sizes vary according to the condition. It scans data in a selective order and inserts data into a window until the difference between maximum and minimum value is less than twice of maximum error rate threshold. If we consider the best case where all off the data values maintaining threshold value, the device will not get any data for a long time. Though APCA gives fewer data points but the time, which can be delayed of through APCA is higher.

III. PROPOSED METHOD

Like PCA, this method also divides the entire dataset into piecewise fixed length segments known as windows, W and it is users define value. A maximum error support threshold, € is also in under consideration. The maximum value, 1 ϵ W and the minimum value, V_{min} ϵ W of each window is calculated. If the maximum and minimum value of the window is less than twice of the maximum error rate threshold then the data values will be compressed.

The equation looks like the following.

$$V_{max} - V_{min} < 2 * €$$ (1)

The above equation can be named as "**Compression Condition**". If the above condition is true, all the data points, $V_i \in W$ is represented by a constant. That constant can be named as "**Representative Value**".

$$\text{Representative Value, } V = \frac{V_{max} + V_{min}}{2}$$ (2)

A. Phases of the Model

The whole method is partitioned into two phases. One is Merge phase and another is a Split phase. Initially, if the data points of the window follow the compression condition, then it will go to the Merge phase. Otherwise, it will go to the split phase.

a) Merge Phase: When the values of a window follow the compression condition then the window is saved into the buffer and the next window is formed with next consecutive data points. Now, V_{max} & V_{min} are calculated from this current window and buffer elements. If the difference between V_{max} & V_{min} is also followed the compression condition, then these two windows will be merged together and temporarily store the window in the buffer and form the next window in the same manner and follow the same process. After each window is compressed, the value of the buffer will be updated by adding the current window's data points. If the compression condition is not violated for at most N consecutive windows, then the data points of these N consecutive windows will be merged and a representative value will be calculated that will be sent to the server. Here N indicates at most how many windows will be merged. It is an arbitrary number and user input. If the values of N consecutive windows do not meet the compression condition, then after each window is compressed, the window stored in the buffer will calculate its representative values and will send into the server. Then it clears the buffer.

b) Split Phase: If the compression condition breaks in PCA, then the full window is sent to the server. But in this method, instead of sending the whole window, it splits the data points of the current window and tries to compress them. Thus, data points are reduced by more than PCA. It also scans data points of a window into a selective order like APCA. Initially, V_{max} & V_{min} of the current window and their differences are calculated. Temporarily the current window is saved in the buffer. After that, the difference violates compression condition or not, is checked. There may be two types of cases at this stage. If V_{max} & V_{min} the current window does not violate the condition, then it will go case 1 otherwise the current window will go the split phase in case 2.

In the first case, since data points of the current window follows the condition, it forms the next window and checks the next window is eligible for merging or not in the previously described way. If the next window will not eligible for merging data points, then a representative value of the window that stored into the buffer will be calculated and sent to the server. After that, it clears the buffer.

In the second case, the first two data points of the buffer element are compared and it is seen that if those data points break the compression condition or not. If the condition breaks, the first data is sent is to the server and the second and third data are compared and again check the condition and thus, the process continues, if there are values in the above window. And if the first two data points of the buffer fulfill the compression condition, then those data points will be compared with the third data. In this way, all the values of that window will be checked sequentially until the buffer becomes empty and compare it to see if the compression condition is met or not. If the condition satisfies then data will be compressed. Finally, the buffer is cleared.

B. Flow Chart

As the entire process is partitioned into phases, it has three flow charts.

a) Main Procedure

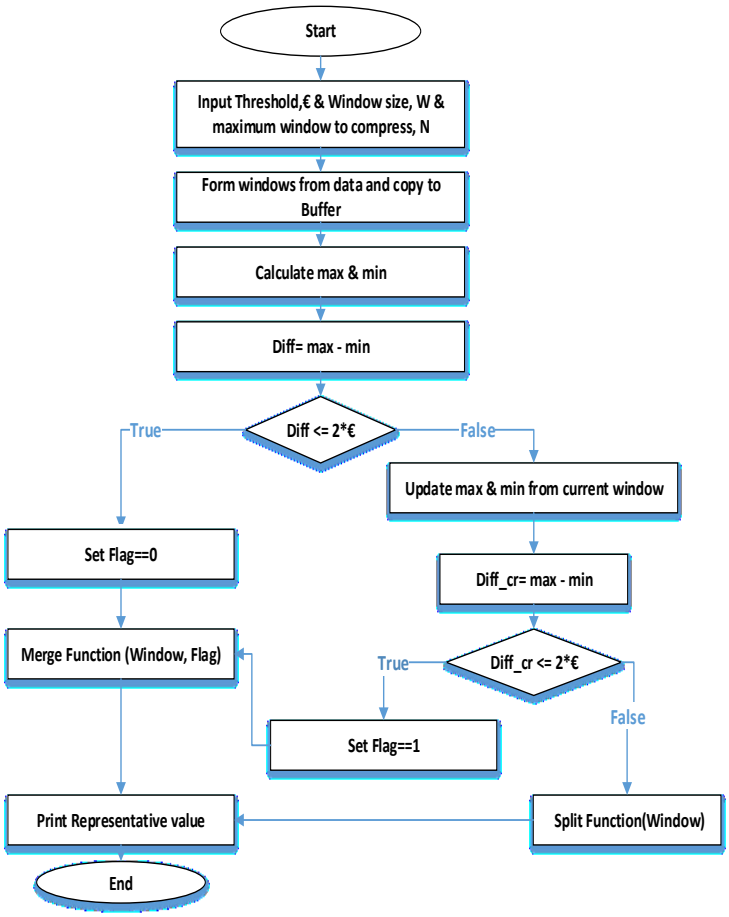


Fig. 1. Flowchart of the Main Procedure

Fig. 1. explains the complete workflow of our proposed method. Initially, the program takes input window size as W; at most how many windows can be merged as N & threshold as € from users. Then build windows from the dataset and copy the window to the buffer. Now, the maximum and minimum value of the window and their differences are

calculated and if the difference is less than twice of threshold and it initiates flag=0 and then calls the merge function with window and flag as an argument. Otherwise maximum and minimum is updated from the buffer the last window and their difference is calculated. If the updated difference is less the twice of threshold then it set flag=1 and call the merge function but if the difference is not less than 2*€ call the split function with a window as an argument. The representative value that calculated finally will send to the server or user.

b) Merge Phase

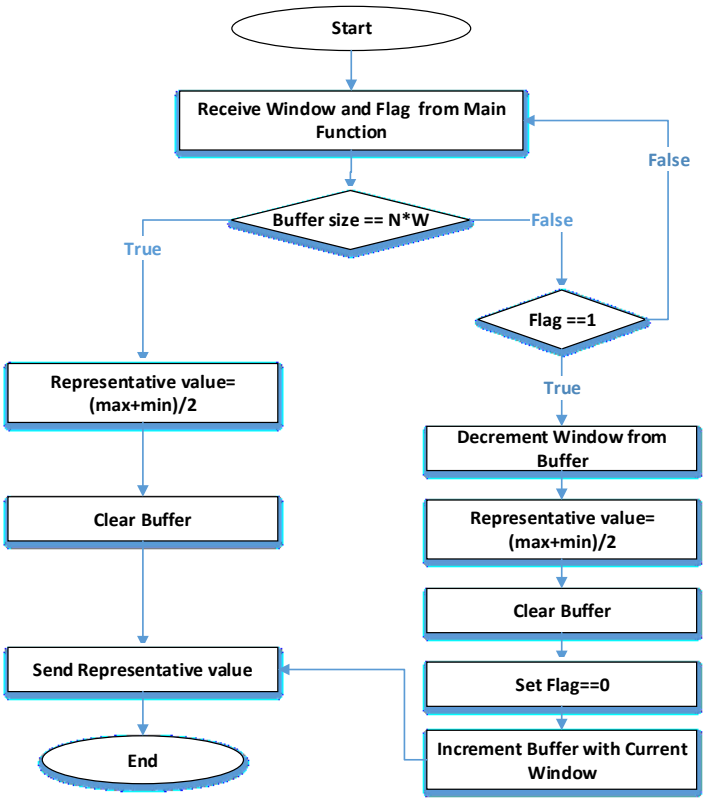


Fig. 2. Flowchart of the Merge Phase

Fig. 2. shows the operations during the merge phase. In the Merge phase, receiving window and flag value it is checked that if buffer size and N*W are equal then the representative value will be calculated directly because buffer size is full. Then it will clear the buffer and sent the representative value to the main function as a return value. If buffer size and N*W are not equal, then it will check flag value. If flag=1 that means current window violates the compression condition, so it decrements the current window from the buffer, calculates the representative value and clears the buffer. After that flag will set as 0 to indicate that buffer can receive window now and increment buffer with current window value and finally representative value will send to the main function as a return value.

c) Split Phase

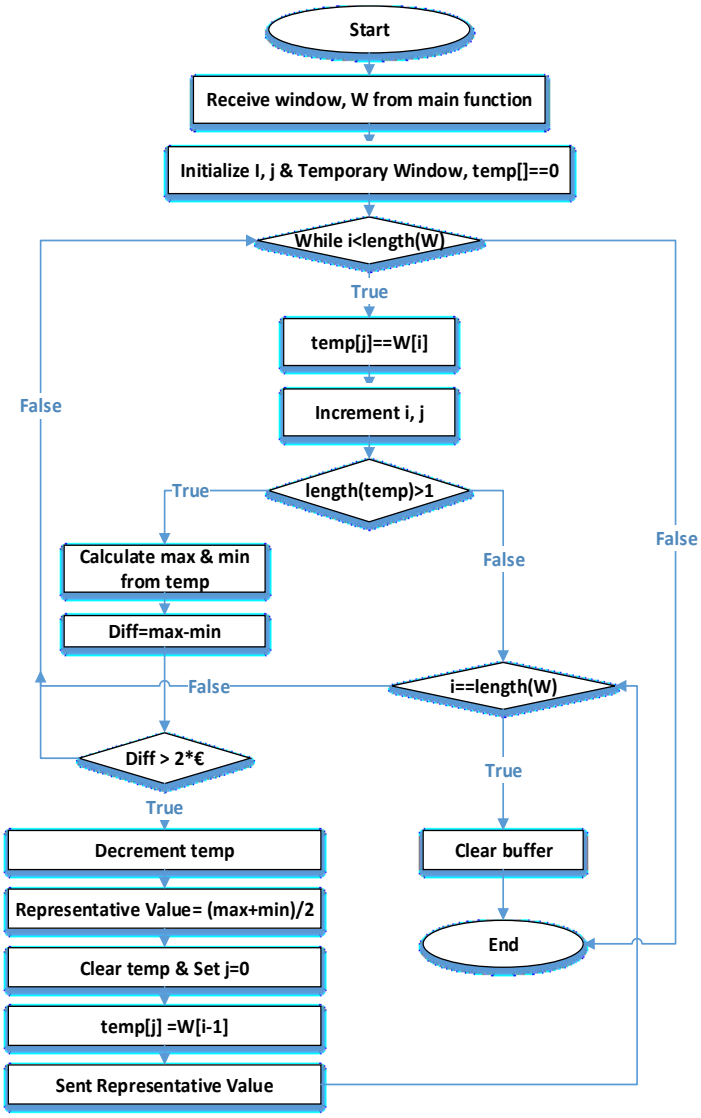


Fig. 3. Flowchart of the Split Phase

Fig. 3. shows the operations during the split phase. In the Split phase, window will receive as an argument from the main function. Then a temporary array is initiated and data of window will be copied to this temporary array. When there is more than one data in the temporary window then it will calculate the difference between the maximum & minimum value of the array. If the difference is less than 2*€ then it will go for next data otherwise pop the last value and calculate the representative value. After that clear the temporary array and current data point will be assigned into its zero indexes. This process will continue for every data point in the window. After scanning all data point's window will be cleared from the buffer.

IV. IMPLEMENTATION

In this method, fewer data points are sent to the server compare to PCA. In APCA, data values are sequentially searched. In the best case all the data values are followed compression condition then the algorithm will execute continuously. If the dataset is too large and a large number of data values are followed compression condition, the server will get any data for a long time so that delay time of APCA is larger. It is necessary to find results in an optimal way. This method is proposed considering both delay time and the number of data points. In PCA, data points are more but less delay time but in APCA data points are low but delay time is high.

A. Datasets

For evaluating Model-based Compression, the data values those have low spike those data will give the optimal result. Basically, we have taken three types of the environmental dataset from Kaggle. A variant number of stream data signals are contained by each and every data set. We have taken water's pH value as our first dataset. Around 10500 a number of data points in total. Their range is in between 0 to 7. We have chosen the second dataset is the one-year temperature of a country. There are roughly 11 thousand data points in total. The third dataset contains 5 years of high temporal resolution (hourly measurements) data of air pressure with more than 21000 number of data points.

B. Running Example

Data source generates data continuously. Here, window size, W is 5, error support threshold, € is 0.5 and N is 3 that indicates at most 3 windows can be merged for compression. Initially, first window forms with consecutive 5 data points from a data source which are 7.3, 7.7, 7.5, 7.2 and 7.1 as the window size are 5 and the window is saved to the buffer. The maximum value, V_{max} & the minimum value, V_{min} of the first window, is 7.7 & 7.1. While the difference between V_{max} & V_{min} is 0.6 which is less than 2*€ so this window can be compressed. Now, it checks the next window is eligible for merging or not.

Again, the second window formed with next consecutive 5 data points for example 7.4, 7.5, 7.6, 7.1, and 7. Now, V_{max} & V_{min} are calculated from this current window and buffer elements. Here, V_{max} is 7.7 & V_{min} is 7. As the difference is 0.7 so it can also be compressed. Temporarily this window also stores in the buffer and form the next window.

In a similar way, the third window is formed with next consecutive 5 data points for instance 7.4, 7.5, 7.6, 7.1, and 6.5 then again, V_{max} & V_{min} is calculated from this current window and buffer elements. Now, V_{max} is 7.7 & V_{min} is 6.5 and their difference is 1.1 so that it violates the compression condition. But it first two windows didn't violate the condition. Hence, at first, we merge the window of the buffer; calculate the representative value from windows that stored in the buffer and send a single representative value to the server instead of sending 10 data points. After that, it clears the buffer.

Now, V_{max} & V_{min} of the current window and their differences are calculated. After that, the difference violates compression condition or not, is checked. There may be two types of cases at this stage. If V_{max} & V_{min} of the current window does not violate compression condition, then it will go case 1 otherwise the current window will go the split phase in case 2.

Case 1: As V_{max} & V_{min} of current window do not violate compression condition, temporarily this window saves in the buffer, form the next window and checks the next window is eligible for merging or not in the previously described way. In this example, the data points of the current window break the compression condition.

Case 2: Since the difference of V_{max} & V_{min} violates the condition, this window is not compatible to compress. In split phase, temporarily this window saves in the buffer and it scans each data points and checks with the compression condition. Initially, the first two data points which are 7.4 & 7.5 is got and is checked whether the difference between these two data points follows the condition or not. The difference is less than twice of € and It maintains the condition. Now, it takes the next data point, 7.6 and checks if the difference between these three data points follows the condition or not. The similar process continues until the data points breaks the compression condition. Here, for the first 4 data points of the window, the compression condition is not violated.

The last data point, 6.5 violates the conditions so it makes a window from the first 4 values and sends a representative value to the server. The last data point also sends to the server. Then it clears the buffer and takes the next window. In this case, only two data points are sent to the server instead of sending 5 data points for this window. If more data points come, then the rest of windows will follow the same process of merging and splitting windows. Thus, compression occurs in the proposed method.

C. Prototype Development

We have made a desktop application with a user interface where the user can select any data source and select a window size, maximum window to compress as Max Window, error threshold value according to their wish. The application calculated total compression data, their compression ratio and RMSE of compression using this proposed technique.

V. EXPERIMENTAL RESULTS

The proposed compression method is evaluated based on various performance factors like compressed data points, Compression ratio and approximation error.

A. Evaluation Measure

a) Compressed Data Points: Compressed data points indicate total how many values are compressed for all the data. Lessen compressed data points indicate the optimal result.

b) Compression ratio: The fraction of data storage values among compressed data and similar original data can be prescribed as compression ratio. The higher compression ratio indicates the higher ability to data compression.

$$\textbf{Compression ratio} = \frac{\textbf{size\_of\_actual\_data}}{\textbf{size\_of\_compressed\_data}} \quad (3)$$

c) Approximation Error: At the point when compression methods demonstrate identical performance, approximation error can turn into a subordinate measurement to consider. To consider this viewpoint, the approximation errors of data compression are illustrated using the root mean square error (RMSE). RMSE generally used to assess the quality of data approximation.

$$\textbf{RSME} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(V_i - \widehat{V_i})^2} \quad (4)$$

Here, v_i means a raw data value in a given sensor data signal, and $\widehat{v_i}$ means the relating approximated value in the compressed signal. Better approximation occurs when the value of RMSE is low. As every stream data signal contains specific range where data values are limited, we standardized the RMSE esteems (from 0 to a given error tolerance €) at each and every experiment set, to have a typical view to exploring the experimental results.

B. Experimental Results

The evaluation is measured with various usages of compression techniques and error support threshold is computed. The absolute error changes over a generally unique range since every data set pass through different domain values. In this manner, it becomes difficult to compare all compression techniques in a similar way. To give the general picture, the relative error is used in the experiments. Besides, in the Merge Phase of the technique, the maximum number of consecutive windows that can be combined together depends on the machine. This varies for a different machine. For our three chosen dataset, we consider Error support threshold is 0.5, the maximum window to compress is 4 and the window size is 5. The results are presented in Table I and Figure 4, 5, 6.

As Table I shows, for dataset 1, dataset 2 and dataset 3 PCA gives the maximum compressed data points compared to the other two techniques. The proposed method gives much fewer data points compared to PCA.

Compare to the other two techniques APCA gives the highest value for each of the datasets. So APCA gives the most optimal result as we know the higher compression ratio indicates a higher ability to compress data.

Fig. 4. shows the experimental results in terms of total data points after the compression. APCA performs better than our proposed method which is not surprise, since it is optimal in terms of total data points after compression. However, our proposed method performs 100% better than PCA which is a major improvement. Fig. 5. shows the compression ratio obtained from these results.

Fig. 6. shows the experimental results in terms of RMSE of the compression algorithms. The RMSEs of data compression is computed to reflect the quality of compression. PCA gives the lowest RMSE value. The RMSE value of PCA and proposed methods is very close. As the lower RMSE value is the better approximation so PCA and proposed method the provides better result than APCA.

The proposed technique gives better result when considering all the performance parameters. It gives much less compressed data points than PCA and it gives higher compression ratio compared to PCA. Besides, RMSE value of this technique is smaller APCA. However, in a few cases, delay time and RMSE of compression is higher than PCA.

TABLE I. EXPERIMENTAL RESULTS

PCA			
	Dataset 1	Dataset 2	Dataset 3
Compressed Data Points	4765	4453	12519
Compression Ratio	2.27	2.41	1.71
RMSE of compression	0.21	0.20	0.30
APCA			
	Dataset 1	Dataset 2	Dataset 3
Compressed Data Points	1606	1387	4911
Compression Ratio	6.67	7.74	4.35
RMSE of compression	0.28	0.31	0.46
Proposed Method			
	Dataset 1	Dataset 2	Dataset 3
Compressed Data Points	2332	2244	6712
Compression Ratio	4.55	4.78	3.18
RMSE of compression	0.21	0.29	0.41

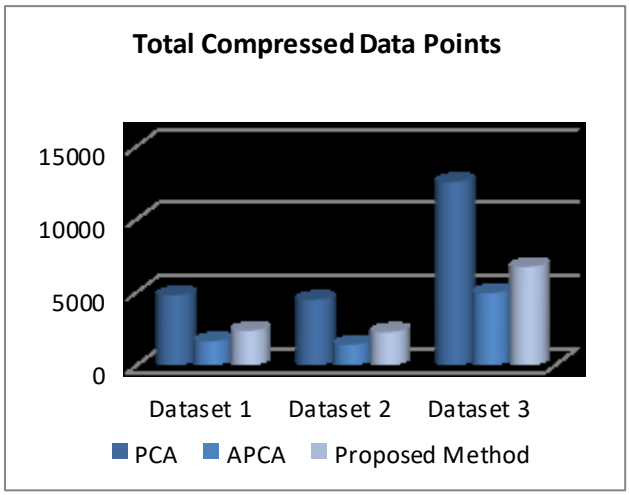


Fig. 4. Comparison of Total Compressed Data Points

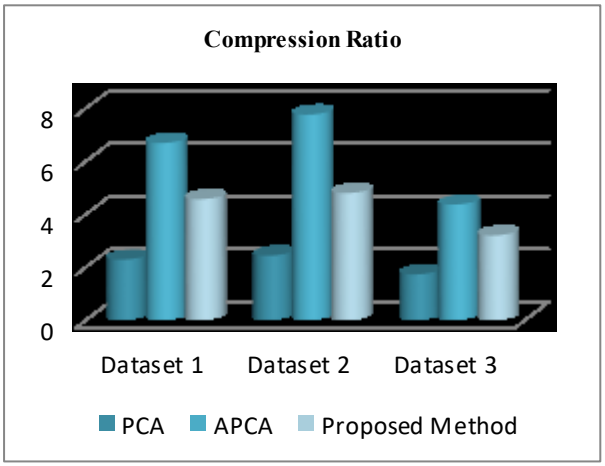


Fig. 5. Comparison of Compression Ratio

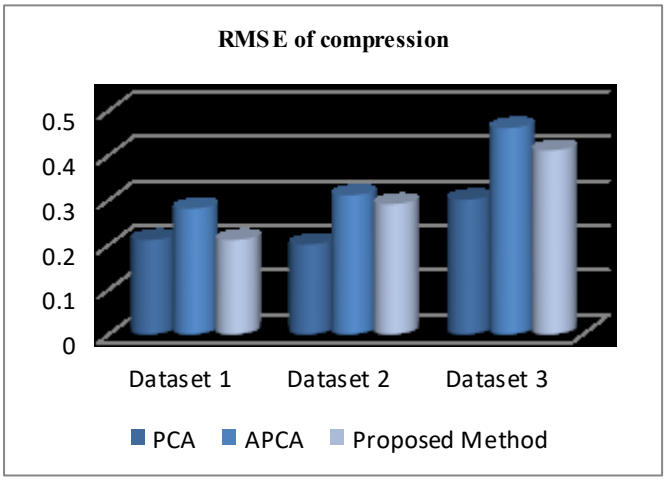


Fig. 6. Comparison of RMSE of Compression

VI. Conclusion and Future Work

This proposed technique is an improvement over PCA. Like PCA, widow size is constant here, but data point is also minimized notably. Considering two criteria such as compression ratio and RMSE value of compression this proposed technique gives most optimal result. Besides the delay time of this method is quite low so that machine or server does not wait long to get the values. Our future concentration is that the machine can automatically take decision about the window size which provides to get optimal result and that window size will be taken accordingly while algorithm will be run. Besides, we will try to apply this algorithm in diverse dataset and minimum more data points as much as can.

References

[1] X. Lian and L. Chen, "Efficient Processing of Probabilistic Reverse Nearest Neighbor Queries over Uncertain Data," in *Proc. VLDB Endowment, vol. 18, no. 3*, 2009, pp. pp. 787-808.

[2] H. Ding et al., "Querying and Mining of Time Series Data: Experimental Comparison of Representations and Distance Measures," in *Proc. VLDB Endowment, vol. 1, no. 2*, 2008, pp. 1542-1552.

[3] Y. Kotidis, and N. Roussopoulo A. Deligiannakis, "Compressing Historical Information in Sensor Networks," *ACM SIGMOD Int'l Conf. Management of Data*, pp. 527-538, 2004.

[4] K. Sayood, *Introduction to Data Compression. Elsevier.*: 2000.

[5] D. Chu et al., "Approximate Data Collection in Sensor Networks Using Probabilistic Models," in *Proc. Int'l Conf. Data Eng. (ICDE)* , 2006, p. 48.

[6] T. Westmann et al, "The Implementation and Performance of Compressed Databases," in *SIGMOD Record*, vol. vol. 29, no. 3, pp-67, 2000, pp. 55-67.

[7] An Evaluation of Model-Based Approaches to Sensor Data Compressio, "Nguyen Quoc Viet Hung, Hoyoung Jeung, and Karl Aberer," *Ieee Transaction on Knowlegde and Data Enginering*, vol. 25, november 2013.

[8] Y. Cai and R. Ng, "Indexing Spatio-Temporal Trajectories with Chebyshev Polynomials," in *Proc. ACM SIGMOD Int'l Conf. Management of Data*, 2004, pp. 599-610.

[9] I. Lazaridis and S. Mehrotra, "Capturing Sensor-Generated Time Series with Quality Guarantees," in *Proc. Int'l Conf. Data Eng. (ICDE)*, 2003, pp. 429-440.

[10] E. Keogh et al, "Locally Adaptive Dimensionality Reduction for Indexing Large Time Series Databases ," in *Proc. ACM SIGMOD Int'l Conf. Management of Data*, 2001, pp. pp. 151-162.

[11] N. Meratnia, P. Havinga Y. Zhang, "Outlier Detection Techniques for Wireless Sensor Networks: A Survey," *J. IEEE Comm. Survey & Tutorials*, vol. 12, no. 2, pp. 159-160, pp. 159-160, second quarter 2010.

[12] Maria ndrawan-Santiago, David Taniar Prajwol Sngat, "Sensor data management in the cloud:Data storage,data ingestion,and data retrieval," 2017.

Mathematical Models for Fat Free Mass Measurement Based on Bioelectrical Impedance Analysis

Tahasin Ahmed Fahim and Md. Shamim Anower
Department of Electrical and Electronic Engineering
Rajshahi University of Engineering & Technology, Bangladesh
Email: tahasin.fahim.bd@ieee.org

***Abstract*—This paper shows new mathematical models for fat free mass (FFM) measurements for both male and female people through bioelectrical impedance analysis. In this research 1300 male and 1350 female (total 2650) data have been considered. Bioelectrical impedance at 100 kHz along with physical parameters like age, height, body mass index (BMI) have been used for the model development. The proposed models have been analyzed statistically. In the statistical analysis, correlation (Pearson) coefficients, 95% limit of agreement (LOA), absolute errors, bias, root mean square error (RMSE) have been used. The results show that the correlation (Pearson) coefficients are 0.998 ($p<0.001$) for male and 0.997 ($p<0.001$) for female people which indicate very good matching with actual data. The intervals of LOA are only -1.90 kg to 1.82 kg and -1.61 kg to 1.56 kg for male and female data respectively and most of errors found by proposed models remain within the limit of agreement. The absolute errors (mean ± Standard Deviation) are (0.81 ± 0.45) kg and (0.69 ± 0.44) kg for male and female data respectively whereas the bias are -0.03 kg for both male and female population only. The RMSE are also very low and which are 0.91 kg for male and 0.79 kg for female people. Comparing the results of this research with existing models it is seen that proposed mathematical models give better results and hence the proposed models can be used for FFM measurement.**

***Index Terms*—Bioelectrical impedance analysis, body compositions, fat free mass, fat free mass index**

I. Introduction

Body composition is the combination of fat mass (FM) and fat free mass (FFM). Comparatively lower portion of body fat and the higher portion of fat free mass are desirable for propoer body composition. Fat free mass consists of total body water, visceral protein, and bone minerals [1]. Generally muscle tissue, under skin and around organs contain fat. The body fat is considered as a proportion of total body weight then the ratio of fat mass and fat free mass can easily measured. American Council on Exercise (ACE) fixed the normal range of body fat percentage for different categories [2]. According to ACE, 2% to 5% body fat for male and 10% to 13% body fat for female are required as essential body fat, 6% to 13% for male athletes and 14% to 20% for female athletes, 14% to 17% for fitness of male and 21% to 24% for fitness of female, 18% to 24% for male and 25% to 31% for female are the maximum acceptable range and the obesity condition is considered beyond the maximum acceptable ranges of body. There are many diseases which occur due to obesity, but no specific disease can be detected by body fat percentage. Besides body mass index (BMI) which is the ratio of body weight to square of height is another parameter for obesity identification, but it is not also specific parameter for diseases. The ratio of fat free mass to square of body height is called fat free mass index (FFMI) and the normal range of FFMI is 18.2 to 20 kgm^{-2} for male and 15 to 16.6 kgm^{-2} for female people [3]. It is one of the prime parameter for different diseases detection.

Fat free mass has significant effects on various types of diseases like cardiac related diseases, anorexia nervosa (eating disorder), chronic obstructive pulmonary disease (COPD), end stage renal disease etc. [4-7]. Fat free mass index is one of the major parameter for chronic heart failure and cardiac complications after cardiac surgery prediction [4]. So, for the correct prediction or detection of such diseases, proper measurements of FFM and FFMI are necessary and if FFM can be estimated correctly then FFMI can easily calculated dividing by the square of the height. Usually, FFM is measured by one of the three methods such as densitometry, dual energy x-ray absorptiometry (DXA) and bioelectrical impedance analysis. Bioelectrical impedance analysis is the more easier and quicker method to measure FFM comparing other methods, but to use bioelectrical impedance analysis accurately, proper model development is mandatory. The accuracy is not still satisfactory for the existing FFM measurement models using bioelectrical impedance analysis. In this paper, new mathematical models for FFM measurement have been developed and validated.

II. Materials and Methods

A. Subjects

From National Health and Nutrition Examination Survey (NHANES), 2003-2004 (version 7, updated in July 2016) performed over US citizen, a data set of 2650 (1300 male and 1350 female) participant's data have been considered for FFM measurement model development [8]. United States Department of Health and Human Services, Centers for Disease Control and Prevention and National Center of Health Statistics conducted this survey in their Mobile Examination Center (MEC) and documented all the results jointly. This survey was operated based on various types of clinical tests results along

with physical and electrical data. For model development and validation, the results of FFM found from dual energy x-ray absorptiometry (DXA) of NHANES data set have been used as standard method for reference. HYDRA Bio-Impedance Spectrum Analyzer (Model 4200) developed by Xitron Technologies, Inc., San Diego, California was used for bioelectrical impedance data collection. The examinations were conducted by expert health technicians maintaining protocol which was mentioned in the Body Composition Procedures Manual on the NHANES website.

B. Dual Energy X-ray Absorptiometry (DXA)

Usually bone mineral density (BMD) is measured by Dual-energy X-ray absorptiometry in which two different energy level x-ray beams are sent to the patient's bones [9]. Measurement of body composition using DXA is possible with satisfactory accuracy, but this procedure is not suitable for all time due to high cost and requirement of complex setup to perform experiment. A low resolution of fat shadow image can be found by DXA from which the total amount of fat mass can be estimated. Finally, after subtraction of fat mass from body mass, fat free mass is found [10].

C. Bioelectrical Impedance Analysis (BIA)

Bioelectrical impedance analysis is a tetra polar electrodes based method. Among four electrodes, two electrodes were connected on right hand. An electrode was connected on wrist and another was connected on the bone of little finger of right hand. Rest of two electrodes were connected on the right foot. One was connected on bone of big toe side of ankle of the right foot and another was connected on base of the second toe of the right foot. A pair of electrodes was for sending constant current to human body and another pair was for measuring voltage drop. Using voltage drop and sending current, bioelectrical impedance analyzed [11, 12].

D. BIA Based Existing Models for FFM Measurement

There are many linear regression models for FFM measurement based on bioelectrical impedance analysis though the accuracy levels are not satisfactory. Among them, four current models of FFM measurement based on bioelectrical impedance analysis have been demonstrated in this section which show better accuracy comparing with other existing models. Kyle $et\ al.$ [13] showed mathematical model for FFM measurement at 50 kHz frequency on the basis of 343 patient's data which were applicable for both male and female individual. The model equation is given below.

$$FFM(kg) = -4.10 + 0.518h^2/R_{50} + 0.23w + 0.13X_{C50} + 4.23s \quad (1)$$

Here, h = height in cm, R_{50} = resistance in kΩ at 50 kHz frequency, X_{C50} = reactance in kΩ at 50 kHz frequency, w = weight in kg, s = sex (1 for men and 0 for women). Sun $et\ al.$ [14] developed two models (one for male and another for female) to measure FFM from 1829 individual's data (734 for male and 1095 for female) and the equations are:

$$FFM\_male(kg) = -10.68 + 0.65h^2/R_{50} + 0.26w + 0.02R_{50} \quad (2)$$

$$FFM\_Female(kg) = -9.53 + 0.70h^2/R_{50} + 0.17w + 0.02R_{50} \quad (3)$$

Where, h = height in cm, R_{50} = resistance in kΩ at 50 kHz frequency, w = weight in kg, a=age in years. A Liu $et\ al.$ [15] built a mathematical model for FFM measurement using 948 participant's data. But, this equation was for school going children (age 8 to 11 years) which was given below

$$FFM(kg) = -6.95 + 0.29h^2/R_{50} + 0.25w + 0.26a + 0.09h \quad (4)$$

In this model, height (h) in cm, age (a) in years, weight (w) in kg, resistance (R_{50}) in kΩ at 50 kHz frequency were considered. Later K.E. Aglago $et\ al.$ [16] developed a mathematical model for FFM measurement using 250 participant's (adult) data which was given below.

$$FFM(kg) = 7.47 + 0.37h^2/R_{50} + 0.31w - 0.06a + 0.09h + 6.04s \quad (5)$$

In this model, height (h) in cm, age (a) in years, weight (w) in kg, sex (s =1 for male 0 for female), resistance (R_{50}) in kΩ at 50 kHz frequency were considered. Kafri $et\ al.$ [17] used multi frequency bioelectrical impedance analysis to measure FFFM in stroke or transient ischaemic patients (250 patients) but no specific model was presented by measuring at 5, 50, 100 and 200 kHz frequencies. Recently Geesje $et\ al.$ [18] showed mathematical model for FFM measurement at 50 kHz frequency on the basis of 103 patient's data which were applicable for both male and female individual. The model equation is given below.

$$FFM(kg) = -1.87 + 0.53h^2/Z_{50} + 0.31w \quad (6)$$

Here, h = height in cm, $Z_{50} = resistance in k\Omega$ at 50 kHz frequency, w = weight in kg.

E. Statistical Analysis

Different types of statistical analysis such as the correlation coefficient (Pearson coefficient), 95% confidence interval (95% CI), 95% limit of agreement (95% LOA), bias (difference between actual and measured values), p-value ($p<0.05$ indicate statistical significance) and root mean square error (RMSE) have been considered for model validation. The general characteristics of people's data have been expressed as mean ± standard deviation (SD), 95% CI and p-value. The formula of RMSE is given below.

$$RMSE = \sqrt{\frac{\sum(x - x')^2}{n}} \quad (7)$$

Where, x indicates the actual data and x' indicated measured data. IBM SPSS and MATLAB software have been used for all types of statistical analysis have been performed for the model development and validations. Bland-Altman plot has been drawn to verify the agreement between the measured data and actual data where agreement limit has been calculated through addition and subtraction of the differences of mean from 1.96 times of standard deviation [19].

III. Results

A. Proposed Mathematical Models for FFM Measurment

To propose the mathematical models for FFM measurement, age, BMI, bioelectrical impedance at 100 kHz frequency and impedance index have been considered. To develop the model, 700 male and 700 female data have been considered. After model development, male FFM model has been validated by 600 data and female FFM model has been validated by 650 data. Table I to IV represent the general characteristics of subject's data which have been used for model development and validation Following equations have been proposed for FFM measurement. have been proposed for FFM measurement.

$$FFM(male, kg) = -13.14 - 0.017a + 0.62b + 0.004Z_{100} + 0.60h^2/Z_{100} \quad (8)$$

$$FFM(female, kg) = -10.52 - 0.004a + 0.47b + 0.004Z_{100} + 0.60h^2/Z_{100} \quad (9)$$

Here, a = age in years, b = BMI in kg/m^2, h = height in cm, Z_{50} = bioelectrical impedance at 50 kHz frequency.

TABLE I
General Characteristics of Male Model Data

Group	n = 700		
	Mean ± SD	***95% CI***	***p-value***
Age ($years$)	21.22 ± 11.18	19.35 to 21.10	<0.001
Height (cm)	167.61 ± 14.81	163.95 to 166.53	<0.001
BMI (kg/m^2)	24.22 ± 5.76	22.96 to 23.83	<0.001
FFM (kg)	50.25 ± 16.75	48.91 to 51.60	<0.001
Z_{100} ($k\Omega$)	486.74 ± 89.75	492.03 to 506.73	<0.001

TABLE II
General Characteristics of Female Model Data

Group	n = 700		
	Mean ± SD	***95% CI***	***p-value***
Age ($years$)	20.84 ± 11.49	19.99 to 21.69	<0.001
Height (cm)	157.14 ± 10.67	156.25 to 157.93	<0.001
BMI (kg/m^2)	25.26 ± 7.17	24.73 to 25.79	<0.001
FFM (kg)	39.02 ± 10.48	38.24 to 39.80	<0.001
Z_{100} ($k\Omega$)	565.89 ± 84.25	559.63 to 572.14	<0.001

B. Performance Evaluation

To propose any model, validation by statistical analysis is required. The major verification tools of statistical analysis are correlation coefficient, 95% limit of agreement (LOA),

TABLE III
General Characteristics of Male Validation Data

Group	n = 600		
	Mean ± SD	***95% CI***	***p-value***
Age ($years$)	20.23 ± 10.91	19.35 to 21.10	<0.001
Height (cm)	165.24 ± 16.10	163.95 to 166.63	<0.001
BMI (kg/m^2)	23.40 ± 5.42	22.96 to 23.83	<0.001
FFM (kg)	50.25 ± 16.75	48.90 to 51.60	<0.001
Z_{100} ($k\Omega$)	499.38 ± 91.64	492.03 to 506.73	<0.001

TABLE IV
General Characteristics of Female Validation Data

Group	n = 650		
	Mean ± SD	***95% CI***	***p-value***
Age ($years$)	20.94 ± 11.40	20.07 to 21.82	<0.001
Height (cm)	157.40 ± 10.26	156.61 to 158.20	<0.001
BMI (kg/m^2)	24.52 ± 6.60	24.01 to 25.03	<0.001
FFM (kg)	37.93 ± 9.89	37.17 to 38.70	<0.001
Z_{100} ($k\Omega$)	580.16 ± 85.62	573.56 to 586.75	<0.001

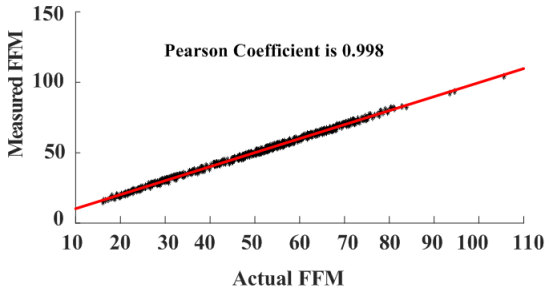


Fig. 1. Correlation Curve of FFM (male).

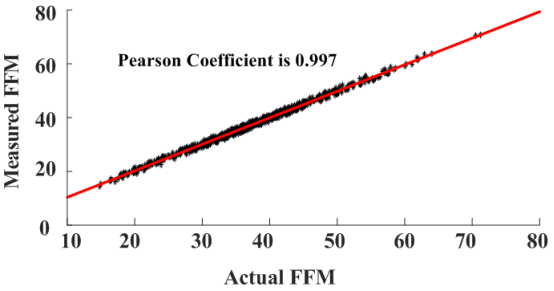


Fig. 2. Correlation Curve of FFM (female).

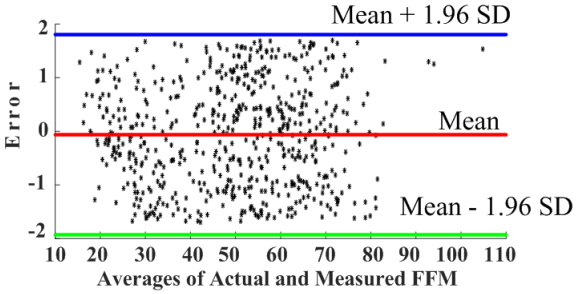


Fig. 3. Bland-Altman plot of FFM (male).

TABLE V
STATISTICAL COMPARISON OF PROPOSED AND CURRENT MODELS (MALE)

Ref. No.	n = 600				
	Correlation Coefficient	***95% LOA*** (kg)	***Absolute Error*** (kg)	***Bias*** (kg)	***RMSE*** (kg)
Kyle *et al.* [13]	0.97	-9.71 to -6.12	3.61 ± 2.56	-1.80	7.08
Sun *et al.* [14]	0.83	-5.87 to 3.73	1.97± 1.45	-0.10	2.97
A-Liu *et al.* [15]	0.88	-5.18 to 13.68	4.96 ± 4.07	4.25	3.70
K.E.Aglago *et al.* [16]	0.92	-12.03 to 7.49	4.44 ± 3.20	-2.27	3.22
Geesje *et al.* [18]	0.92	-1.47 to 7.86	3.27 ± 2.26	3.20	3.99
Deurenberg *et al.* [15] (50 kHz frequency)	0.950	-8.16 to 5.13	3.03 ± 2.14	-1.51	3.71
Cornish *et al.* [16]	0.850	-3.4 to 1.00	1.40 ± 0.86	-1.20	1.64
Sun *et al.* [17]	0.840	-3.02 to 2.42	1.17 ± 0.81	-0.3	1.42
Liu *et al.* [18]	0.88	-3.52 to 8.86	3.24 ± 2.58	2.67	4.14
Aglago *et al.* [19]	0.92	-8.32 to 5.64	3.10 ± 2.21	-1.34	3.80
Diouf *et al.* [**?**]	0.890	-3.42 to 10.77	4.17 ± 3.03	3.67	5.15
Proposed	0.998	-1.90 to 1.82	0.81 ± 0.45	-0.03	0.91

TABLE VI
STATISTICAL COMPARISON OF PROPOSED AND CURRENT MODELS (FEMALE)

Ref. No.	n = 650				
	Correlation Coefficient	***95% LOA*** (kg)	***Absolute Error*** (kg)	***Bias*** (kg)	***RMSE*** (kg)
Kyle *et al.* [13]	0.97	-10.80 to -2.76	6.78 ± 2.02	-6.77	4.43
Sun *et al.* [14]	0.83	-7.41 to 2.73	2.53 ± 1.56	-1.84	2.45
A-Liu *et al.* [15]	0.88	-8.10 to 4.68	2.77 ± 2.45	-1.74	6.41
K.E. Aglago*et al.* [16]	0.92	-6.57 to 1.71	2.80 ± 1.60	-2.43	5.47
Geesje *et al.* [18]	0.92	-4.18 to 2.10	1.45 ± 1.25	-1.04	1.91
Deurenberg *et al.* [15] (50 kHz frequency)	0.950	-5.60 to 1.84	2.30 ± 2.68	-1.88	2.68
Cornish *et al.* [16]	0.850	-3.4 to 1.00	1.40 ± 0.86	-1.20	1.64
Sun *et al.* [17]	0.840	-3.79 to 1.54	1.51 ± 0.90	-1.12	1.76
Liu *et al.* [18]	0.88	-5.10 to 2.46	1.85 ± 1.42	-1.32	2.33
Aglago *et al.* [19]	0.92	-4.36 to 1.96	1.68 ± 1.11	-1.20	2.00
Diouf *et al.* [**?**]	0.890	-3.43 to 5.64	1.91 ± 1.71	1.10	2.56
Proposed	0.997	-1.61 to 1.56	0.69 ± 0.44	-0.03	0.79

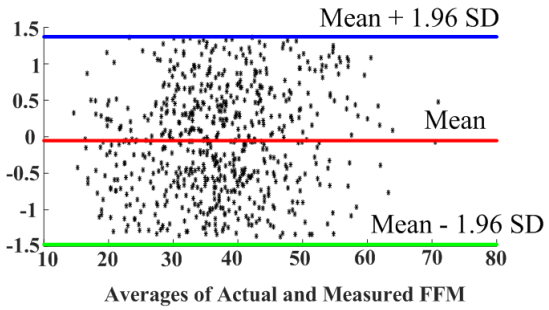


Fig. 4. Bland-Altman plot of FFM (female).

absolute error, bias and root mean square error (RMSE). The proposed FFM models exhibit the correlation coefficient 0.998 and 0.997 ($p<0.001$) for male and female respectively. The 95% LOA for male data are -1.90 to 1.82 kg and for female data are -1.61 to 1.56 kg. The values of absolute error (mean ± SD), bias and RMSE are 0.81 ± 0.45 kg, 0.03 kg and 0.91 kg for male people and 0.69 ± 0.44 kg, -0.03 kg and 0.79 kg for female data which indicate lower values. The linear correlation curves for both male and female have shown in Fig. 1 and Fig. 2 respectively. Bland-Altman plot has been shown in Fig. 3 and Fig. 4 by the averages of actual and predicted values of FFM and the 95% LOA. All the results follow the standard limit or range of statistical analysis parameters. So, it is expected that the proposed regression models can be used for FFM measurement with better accuracy.

IV. DISCUSSION

The proposed models have been compares with five related models. All types of statistical analysis have been performed over all the models which have been selected for comparison with proposed models. Table V and Table VI show all the results found from the comparative analysis. From the analysis it is observed that the values of correlaion coefficients are lower than the proposed models. The smallest correlation coefficient value has been found in the reference [14]. The largest value of correlation coefficient among the current models has been estimated to be 0.97 in reference [13] which is lower than proposed models. For the proposed models, the 95% limits of agreement are also small and all the errors remain within the limit whereas the other models show larger 95% limit of agreement. In reference [13], for male model, the abosute error has been found maximum whereas for the female model, the RMSE has been found maximum. In reference [15], for male model, the RMSE has been found maximum whereas for the female model, the absolute error has been found maximum. The values of bias have been found maximum in reference [13]

for male and in reference [15] for female. Comparing all the results of the models it can be said that the proposed models for FFM measurement for both male and female people will be helpful for the implementation of the bioelectrical impedance analysis based diagnosis scheme practically.

V. Conclusion

To monitor the health condition along with disease detection, fart free mass is one of the major parameter of the body composition. Generally, DXA is used to measure body fat distribution from which fat free mass is estimated, but this process complex and costly procedure and require huge space for arrangement. So, bioelectrical impedance based method will be a good alternative method due its simplicity, portability and quick process. The main concern issue of bioelectrical impedance based method is the accuracy. From the analysis, it can easily be said that the implementation of the proposed models will greatly improve the accuracy of the bioelectrical impedance analysis method. The fat free mass index, another parameter measured from fat free mass is also important for disease detection. The proper measurement of fat free mass index depends on the fat free mass. Using the propose models, the measurement of fat free mass index will also be possible with better accuracy. Since the validation results through the statistical analysis for both male and female people have proper results, these models can be used for every person to measure FFM as well as the different types of algorithms to detect different types of diseases can be developed in future based on this models.

Acknowledgment

National Health and Nutrition Examination Survey (NHANES) 2016, United States of America (USA).

References

[1] S. F. Khalil *et al.*, "The Theory and Fundamentals of Bioimpedance Analysis in Clinical Status Monitoring and Diagnosis of Diseases," Sensors (MDPI), vol. 14, no. 6, pp. 10895-10928, 2014.

[2] N. D. Muth, "Guidelines for percentage of body fat loss," American Council on Exercise, 2009.

[3] H. R. Hull *et al.*, "Fat-free mass index: changes and race/ethnic differences in adulthood", International Journal of Obesity, vol. 35, no. 1, pp. 121-127, 2011.

[4] D. Kinoshita *et al.*, "Fat Free Mass Index is an Important Predictor for Future Cardiac Events in Patients After Cardiac Surgery", American Heart Association Circulation, vol. 132, no. 3, 2015.

[5] M.V. Moreno *et al.*, "Assessment of body composition in adolescent subjects with anorexia nervosa by bioimpedance", Medical Engineering Physics Journal, vol. 30, pp. 783-791, 2008.

[6] L. Silveira *et al.*, "The relationship between fat-free mass index and pulmonary hyperinflation in COPD patients", Respirology (Wiley), vol. 19, no. 8, pp. 1204-1208, 2014.

[7] S. M. Vine *et al.*, "Bioimpedance spectroscopy for the estimation of fat free mass in end-stage renal diseases," European Journal of Clinical Nutrition and Metabolism (Elesevier), vol. 6, pp. 1-6, 2011.

[8] United States Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, National Health and Nutrition Examination Survey (NHANES), 2003-2004. Ann Arbor, MI: Inter-university Consortium for Political and Social Research, 2016-07-11.

[9] B. C. Wingo *et al.*, "Comparison of segmental body composition estimated by bioelectrical impedance analysis and dual-energy X-ray absorptiometry," Clinical Nutrition ESPEN (Elsevier), vol. 28, pp. 141-147, 2018.

[10] Dual energy X-ray absorptiometry - Bone mineral densitometry, International Atomic Energy Agency, 2017.

[11] T. K. Bera, "Bioelectrical Impedance Methods for Noninvasive Health Monitoring: A Review", Journal of Medical Engineering, vol. 2014, pp. 1-28, 2014.

[12] T. A. Fahim *et al.*, "An Embedded System for Observing the Stages of Lung Cancer in Non-Invasive Way", 6th IEEE Global Conference on Consumer Electronics, pp. 285-286, 2017.

[13] U.G. Kyle, L. Genton, L. Karsegard, D. O. Solsman and C. Pichard, "Single prediction equation, for bioelectrical impedance analysis in adults," Nutrition (Elsevier), vol. 17, pp. 248-253, 2001.

[14] S. S. Sun *et al.*, "Development of bioelectrical impedance analysis prediction equations for body composition with the use of a multicomponent model for use in epidemiologic surveys," American Journal of Clinical Nutrition, vol. 77, pp. 331-340, 2003.

[15] A. Liu *et al.*, "Validation of bioelectrical impedance analysis for total body water assessment against the deuterium dilution technique in Asian children," European Journal of Clinical Nutrition (Nature), vol. 65, pp. 1321-1327, 2011.

[16] K. E. Aglago *et al.*, "Development and validation of bioelectrical impedance analysis equations for predicting total body water and fat-free mass in North-African adults," European Journal of Clinical Nutrition (Nature), vol. 67, pp. 1081-1086, 2013.

[17] M. W. Kafri, J. F. Potter and P. K. Myint, "Multi-frequency bioelectrical impedance analysis for assessing fat mass and fat-free mass in stroke or transient ischaemic attack patients," European Journal of Clinical Nutrition (Nature), vol. 68, no. pp. 677, 2014.

[18] Geesje H. Hofsteenge, Mai JM Chinapaw and Peter JM Weijs, "Fat-free mass prediction equations for bioelectric impedance analysis compared to dual energy X-ray absorptiometry in obese adolescents: a validation study", BiomMed Central (Springer Nature), vol. 15, pp. 158-166, 2015.

[19] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement", Lancet, vol.1, pp. 307-310, 1986.

Statistical Approach for Classifying Sentiment Reviews by Reducing Dimension using Truncated Singular Value Decomposition

Asmaul Husna
Department of Information Technology
University of Information Technology & Sciences
Dhaka, Bangladesh
asmaul.husna@uits.edu.bd

Habibur Rahman
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
habib_cse_ruet@yahoo.com

Emrana Kabir Hashi
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
emrana.kabir@ruet.ac.bd

***Abstract*—Human sentiments, the attitudes, thoughts, or the judgments backed by feelings are the most important factors for the field of science and business which can be expressed through the web. The rapid growth of unstructured sentiment data has created a very big challenge to provide a fast and cost-efficient model for sentiment analysis. The aim of this paper is to find positive and negative emotions from the public comments rapidly and effectively. In this paper, to vectorize the text reviews and to reduce the dimension of the vectorized data, Term Frequency-Inverse Document Frequency (TF-IDF) and Truncated Singular Value Decomposition (TSVD) is used respectively. An efficient approach has been provided using the logistic regression algorithm employed in a vector space model of affective knowledge which gives fast learning speed. In result, this proposed method has reduced the learning time by 42.63% with the reduction of dimension from 12209 to 100 and has reported almost the same accuracy which is 85.7% and 89.3% for the movie review and yelp restaurant review dataset respectively. In comparison to existing methods, this method has shown almost the same accuracy after dimension reduction. Our result showed that our proposed system is useful.**

***Index Terms*—Sentiment Analysis, Logistic Regression, Truncated Singular Value Decomposition, Vector Space Model**

I. Introduction

In this modern era, people are very much dependent on social networks, blogs and other online media for expressing their opinions and sentiments. The availability of these web communities gives opportunity to people for sharing their opinions and ideas in a cost and time efficient way. This huge source of information and opinions has provided a powerful communication medium to share knowledge and to get advantage from others experience. At present, it is a key factor for the organizer to distill the knowledge from unstructured web data of public opinions and identify the mind of the customers about their products, policies, and organizations.

Sentiment analysis is a process of extracting positive and negative opinions from the review of people about anything. But the review of people which means the unstructured web data is hardly accessible to computers and for this, extraction the reviews into positive and negative is difficult. For solving this problem, a different type of machine learning algorithm exists. Such algorithms are very good at retrieving text, splitting them into parts, counting their words and finally giving a positive or negative result [1], [2]. Most of the researchers have only focused on which algorithm is superior to others for extracting reviews but they have not considered time and cost [10]–[16]. But as the amount of social data is increasing rapidly, time and cost is also a big issue for sentiment analysis.

In this paper, we have focused on not only extracting positive and negative reviews but also the optimal use of time, storage space, and cost. There are several types of feature extraction method and dimension reduction technique. Among them, Term Frequency-Inverse Document Frequency(TF-IDF) and Truncated Singular Value Decomposition(TSVD) have been used for feature extraction and dimension reduction respectively which perform the main role in reducing time and cost. Logistic regression has been used as a machine learning algorithm for extracting sentiment for its fast learning speed. Present related works and their limitations, our proposed framework, evaluation the performance have been described in the next sections subsequently. Finally, we have concluded our research with future work.

II. Related Works

Sentiment analysis is one of the recent hot topics for social, political, scientific and business world. Sentiment analysis depends on the natural language texts. Natural language texts are using in a wide range from questions and answering to mathematical word problem solving [3]. For classify the

sentiments, Support Vector Machine (SVM) [4]–[6], Nave Bayes (NB) [7], [8] and Maximum Entropy (ME) [9], [10] are the most popular machine learning algorithms.

In paper [11], 82.9% accuracy has been got for movie reviews using SVM with a bag of words feature sets. Nave Bayes, ME have also been used but SVM is superior among them. In paper [12], they have trimmed out objective content from movie reviews and used SVM and NB bag of words classifier to determine review polarity with an accuracy of 87.15%. They have examined the relation between subjectivity detection and polarity classification. For subjectivity detection, a graph has been built where source and sink corresponded to the class of subjective and objective sentences, respectively. Semi-supervised learning for imbalanced sentiment classification has been addressed in [13]. They have only focused on the imbalanced problem in the initial labeled data.

A novel approach using SVM to find out the overall positive and negative scores for a particular feature has presented in [14]. SVM has been showed as a best than other machine learning algorithm. In paper [15], Delta TF-IDF has been introduced for subjectivity detection and sentiment polarity classification, as well as documents of varying sizes. They have shown their improved result using SVM. Different corpora available for scientific research in opinion mining have been compared in [16]. They have applied machine learning algorithm (SVM) for different corpora.

For aspect-based sentiment analysis in [17], SVM and Logistic Regression [21] have been used as supervised machine learning algorithm. But they analyzed their result only for higher dimension. After analyzing above papers, we can see that in most paper researchers used SVM as a machine learning algorithm for sentiment mining. But in this paper, we have used logistic regression algorithm and compared with SVM using dimension reduction technique.

III. Proposed System

The goal of our system is to analysis sentiment with cost-effective way. Fig. 1 gives the overview of our proposed system.

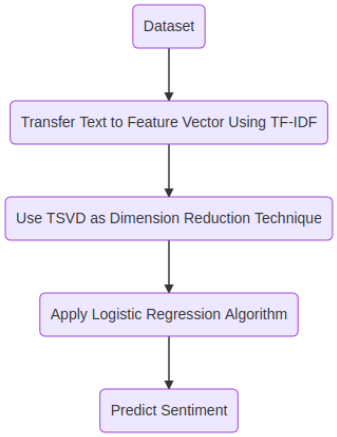


Fig. 1. Flow of the proposed system

In the following sections, we discussed the detailed methodology:

A. *Dataset:*

We have used the movie review dataset and the yelp restaurant review dataset [24]. Movie review dataset is derived from the benchmark corpus developed by Pang and Lee [18]. Those corpus includes 1000 positive and 1000 negative movie reviews. We have divided the datasets into train set and test set using 10 fold cross-validation.

B. *Text to Feature Vector:*

Vector space model represents the text documents as vectors of identifiers. It is used in information filtering, information retrieval etc. Term Frequency-Inverse Document Frequency (TF-IDF) [20] is one of the examples of vector space model. It is used for transforming text to feature vector. Actually, TF-IDF is a numerical statistics that determines how important a word is in a corpus. As 83% of text-based recommended systems in the domain of digital libraries use TF-IDF, we have selected it.

Suppose, we have a set of English text document. The term frequency of a term(word) in this document is proportional to the number of occurrence of the term in this document. So it can be defined as $tf(t,d) = F(t,d)$ where $F(t,d)$ is number of occurrences of term, t in document, d. Inverse document frequency diminishes the weight of terms that occur very frequently in the document set and increases the weight of terms that occur rarely. It is defined as the (1):

$$idf(t,D) = \log \frac{N}{N_{t \in D}} \tag{1}$$

Here, N is the total number of files in the corpus D and $N_{t \in d}$ is the number of files in which term, t is present. Finally, *tf-idf* is calculated as the following (2):

$$tf - idf(t,d,D) = tf(t,d) \times idf(t,D) \tag{2}$$

C. *Dimension Reduction:*

Truncated Singular Value Decomposition (TSVD) reduces the time and storage space by converting higher dimension into the lower dimension [23]. By removing multi-collinearity, it improves the performance of machine learning algorithm. We have selected 100 as the lower dimension. TSVD is preferred to other dimension reduction technique for its simplicity and relatively low computational cost. It is particularly suitable for measuring cross-correlations between affective word or concept.

D. *Logistic Regression Algorithm:*

Logistic Regression has proven highly effective in natural language processing and text classification [20], [21]. It was developed by David Cox in 1958 [22].

We have chosen the approach in [20] for Logistic Regression algorithm cause this approach simultaneously avoids overfitting and maintains high learning speed. It is also appropriate when the dependent variable is binary. As we require a binary

efficient classifier, we have chosen it. Besides this, it works by maximizing posterior class probability. So, for the lower dimension, it works efficiently in terms of time and accuracy.

Logistic regression measures the relationship between the categorical dependent variable and one or more independent variables by estimating probabilities using a logistic function.

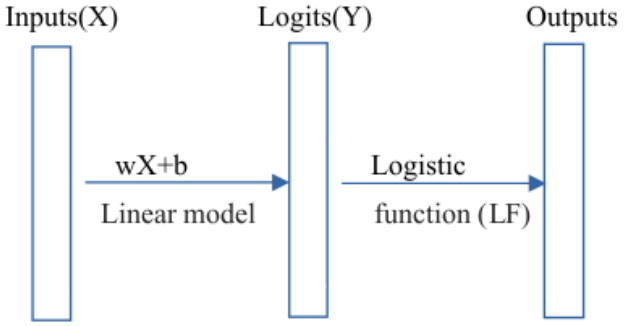


Fig. 2. Model of Logistic Regression Algorithm

where, weight of the corresponding input, w and bias input, b and *LF* represents Logistic Function as follow (3):

$$LF = \frac{e_y}{1 + e_y} \tag{3}$$

IV. Performance Evaluation

For sentiment analysis using machine learning algorithm most of the existing system have used SVM and they have proved that SVM is superior to others. But most of them have not used any dimension reduction technique and have not considered time, dimension and cost. So, we have compared our system with SVM with dimension reduction technique. Time has also been considered in this performance evaluation. Finally, we have compared our result with the results of the existing systems.

A. *Hardware and Software Configuration:*

- **Operating System:** Linux
- **Programming Language:** Python 2.7
- **Processor:** Intel Core i5-3317U CPU @ 1.70GHz
- **RAM:** 4GB

B. *Result Analysis and Comparison:*

Dimension of the training data has been evaluated before using TSVD. Then 10 fold cross-validation has been used. We have got the following result from Logistic Regression and SVM. We have also shown the confusion matrix for both algorithms in I.

TABLE I
Confusion Matrix Without Dimension Reduction

Movie Review Dataset					
		Actual			
Predicted		Neg.		Pos.	
	Classifier	SVM	LR	SVM	LR
	False	850	864	150	136
	True	154	137	846	863
Yelp Restaurant Review Dataset					
		Actual			
Predicted		Neg.		Pos.	
	Classifier	SVM	LR	SVM	LR
	False	907	907	93	93
	True	117	106	883	894

From confusion matrix, we can measure the accuracy like the (4).

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \tag{4}$$

where, T_p =True Positive, T_n = True Negative, F_p = False Positive, and F_n = False Negative.

TABLE II
Result Analysis without Dimension reduction

Movie Review Dataset			
Classifier	**Accuracy(%)**	**#Features**	**Training Time**
SVM	84.8	12209	8.1166726
LR	**86.35**	12209	**0.0763286**
Yelp Restaurant Review Dataset			
Classifier	**Accuracy(%)**	**#Features**	**Training Time**
SVM	89.5	12209	1.4578088
LR	**90.05**	12209	**0.069607299**

After the dimension reduction, we have shown the fixed dimension and the training time for both algorithms in the following table III.

TABLE III
Confusion Matrix after Dimension Reduction

Movie Review Dataset					
		Actual			
Predicted		Neg.		Pos.	
	Classifier	SVM	LR	SVM	LR
	False	818	849	182	151
	True	208	135	792	865
Yelp Restaurant Review Dataset					
		Actual			
Predicted		Neg.		Pos.	
	Classifier	SVM	LR	SVM	LR
	False	846	901	154	99
	True	103	115	897	885

TABLE IV
RESULT ANALYSIS AFTER DIMENSION REDUCTION

Movie Review Dataset			
Classifier	**Accuracy(%)**	**#Features**	**Training Time**
SVM	80.5	100	0.5638779
LR	**85.7**	100	**0.0437935**
Yelp Restaurant Review Dataset			
Classifier	**Accuracy(%)**	**#Features**	**Training Time**
SVM	88.65	100	0.5866717
LR	**89.3**	100	**0.0398641**

Now comparing Table II and Table IV, we can see that accuracy of Table IV slightly decreases. But in terms of dimension and training time, Table IV is superior. When SVM has been used without reducing dimension, it is better than Logistic Regression in terms of accuracy. But when dimension has been reduced, Logistic Regression is better than SVM.

Fig. 3 shows the comparison of training time before and after the dimension reduction for SVM and LR.

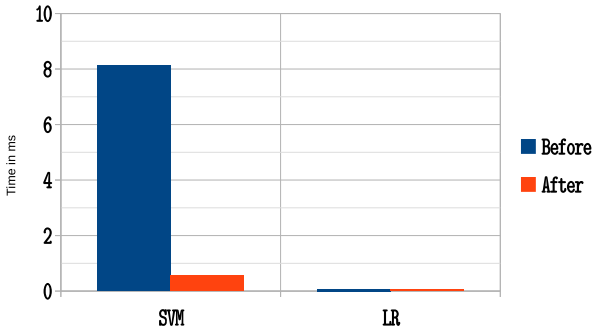


Fig. 3. Comparison of Training Time

Dimension of movie review data has been decreased from 12209 to 100 using TSVD. For this, storage space and cost both have been decreased. Besides this, from table IV we can see that Logistic Regression is 42.63% fast according to training time but the accuracy is almost same.

In terms of the number of features, our proposed system uses 100 features which is reduced by TSVD. On the otherhand, previous two methods [11], [14] have not used any dimension reduction techniques. Fig. 4 shows the comparison of features.

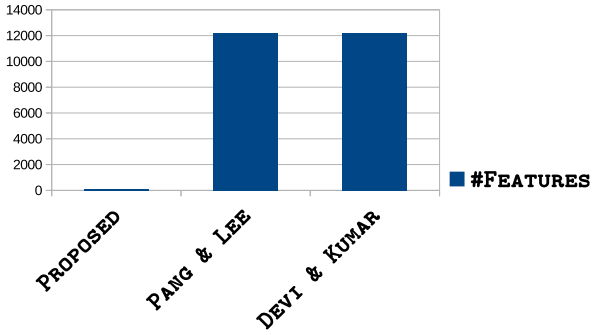


Fig. 4. Comparison of Features

For movie review dataset, our proposed systems shows an accuracy of 85.7% while [11] was 82.9%. Compared to [11], our system is superior. But [14] was 90.85% accurate which is superior to our system and [11]. Fig. 5 depicts the comparison:

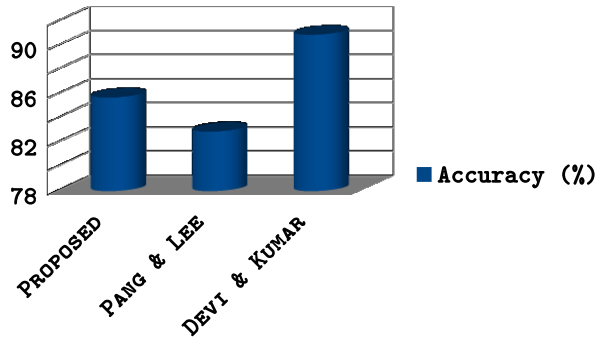


Fig. 5. Comparison of Accuracy

In conclusion, our proposed system is more accurate than [11] and [14] for movie review dataset in terms of features.

V. CONCLUSION AND FUTURE WORK

The main goal of this paper is to introduce a new framework for sentiment analysis which is both time and cost efficient by dimension reduction. The dimension also increases for sentiment analysis from large-scale data. For this learning time of any algorithm increases which is an important factor for measurement performance of any system. The system with faster speed and reasonable accuracy are more appropriate for sentiment analysis than the slower system with high accuracy. By analyzing above result, we can come to a conclusion that the accuracy of our system is almost same to the system using SVM after using dimension reduction technique. But in terms of training time, Logistic Regression is better than SVM. Besides this, the dimension of the proposed system is comparably very low. So our system is cost-effective and we have successfully provided a time and cost efficient model for sentiment analysis.

Unigram TF-IDF has been used for transforming text to feature vector. Bigram TF-IDF and deep learning can be used in future for more accurate result.

ACKNOWLEDGMENTS

The authors of this paper are grateful to all the members of the Syntax and Semantic Research Group, Bangladesh. Authors are also thankful to the Rajshahi University of Engineering and Technology, and to the University of Information Technology and Sciences.

REFERENCES

[1] B. Pang and L. Lee, "Opinion mining and sentiment analysis," Foundations and Trends in Information Retrieval, vol. 2, pp. 1-135, January 2008.

[2] B. Liu, "Sentiment analysis and opinion mining," Morgan and Claypool, April 2012.

[3] Habibur Rahman, Julia Rahman, and Asmaul Husna, "WALGES: Weighted Probability Based Scoring Approach for Solving Algebraic Word Problems using Semantic Parsing", in Proc. of the International Conference on Electrical, Computer and Communication Engineering, February 2019.
[4] T. Joachims, "Text Categorization with Suport Vector Machines: Learning with Many Relevant Features," in Proceedings of the 10th European Conference on Machine Learning, April 1998, pp. 137-142.
[5] T. Joachims, "Making large-scale support vector machine learning practical," MIT Press Cambridge, MA, USA , 1999.
[6] E. Leopold and J. Kindermann,"Text Categorization with Support Vector Machines. How to Represent Texts in Input Space?," Machine Learning, vol. 46, pp. 423-444, March 2002.
[7] D. D. Lewis, "Naive (Bayes) at Forty: The Independence Assumption in Information Retrieval," in Proceedings of the 10th European Conference on Machine Learning, April 1998, pp. 4-15.
[8] S. B. Kim, K. S. Han, H. C. Rim and S. H. Myaeng, "Some Effective Techniques for Naive Bayes Text Classification," IEEE Transactions on Knowledge and Data Engineering, vol. 18, pp. 1457-1466, November 2006.
[9] A. L. Berger, V. J. D. Pietra and S. A. D. Pietra,"A maximum entropy approach to natural language processing," Computational Linguistics, vol. 22, pp. 39-71, March 1996.
[10] K. Nigam, J. Lafferty and A. McCallum, "Using maximum entropy for text classification," in Proc. of the IJCAI Workshop on Machine Learning for Information Filtering, 1999, pp. 6167.
[11] B. Pang, L. Lee, and S. Vaithyanathan,"Thumbs up? sentiment classification using machine learning techniques," in Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing (EMNLP), July 2002, vol. 10, pp. 79-86.
[12] B. Pang and L. Lee, "A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts," in Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics, July 2004, Art. no. 271.
[13] S. Li, Z. Wang, G. Zhou and S. Y. M. Lee, "Semi-supervised learning for imbalanced sentiment classification," in Proceedings of the Twenty-Second international joint conference on Artificial Intelligence, vol. 3, pp. 1826-1831, July 2011.
[14] D. V. N. Devi, C. K. Kumar and S. Prasad, "A feature based approach for sentiment analysis by using support vector machine," IEEE 6th International Conference on Advanced Computing, February 2016.
[15] J. Martineau and T. Finin, "Delta TFIDF: an improved feature space for sentiment analysis," in Proceedings of the Third International ICWSM Conference, 2009.
[16] M. R. Saleh, M. T. M. Valdivia, A. M. Raez and L. U. A. Lopez, "Experiments with SVM to classify opinions in different domains," Expert Systems with Applications: An International Journal, vol. 38, pp. 14799-14804, November 2011.
[17] N. U. Pannala, C. P. Nawarathna, J. T. K. Jayakody, L. Rupasinghe and K. Krisnadeva, "Supervised Learning Based Approach to Aspect Based Sentiment Analysis," IEEE International Conference on Computer and Information Technology, December 2016.
[18] B. Pang and L. Lee, "Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales," in Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics, June 2005, pp. 115-124.
[19] J. Beel, B. Gipp, S. Langer and C. Breitinger, "Research-paper recommender systems: a literature survey," Digital Libraries, vol.17, pp. 305-338, November 2016.
[20] A. Genkin, D. D. Lewis and D. Madigan, "Large-scale Bayesian logistic regression for text categorization," Technometrics, vol. 49, pp. 291-304, 2007.
[21] S. Aseervatham, A. Antoniadis, E. Gaussier, M. Burlet and Y. Denneulin, "A sparse version of the ridge logistic regression for large-scale text categorization," Pattern Recognition Letters, vol. 32, pp. 101-106, January 2011.
[22] D. Cox, "The regression analysis of binary sequences," Royal Statistical Society, vol. 20, pp. 215-242, 1958.
[23] P. C. Hansen, "The truncatedSVD as a method for regularization," Bit, vol. 27, no. 4, pp. 534553, 1987.
[24] Yelp.com. (2019). Yelp Dataset. [online] Available at: https://www.yelp.com/dataset [Accessed 15 Apr. 2019].

Building Predictive Models for Infant Mortality Rate in Bangladesh

Mst Lucky Sultana Lovely
Department of Computer Science and Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
luckysultanacsepust@gmail.com

Mst Irin Sultana
Department of Computer Science and Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
irinsultanacse@gmail.com

Md Mahmudul Hasan
Department of Computer Science and Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
mahmudul.cse@pust.ac.bd

***Abstract*—Infant Mortality Rate (IMR) is one of the key indicators for a developing country, Bangladesh in particular, to measure the health and social status of its population. Many risk factors are associated with IMR that requires close investigation. This paper aims to build predictive models by identifying such potential risk factors responsible for infant mortality in Bangladesh. The World Bank data repository, based on "World Development Indicator (WDI) 2015", has been analyzed. We identify significant factors, build multiple stepwise linear regression models, and evaluate cross-coefficient model performances using MAE, MRE and MSE under the shed of k-fold cross validation technique. The performance of the final prediction model yields significant outcome towards the policy makers that could eventually reduce the infant mortality rate in Bangladesh.**

Keywords—Data Mining, Multiple Linear Regression, Risk Factors, Infant Mortality.

I. INTRODUCTION

Bangladesh has an estimated population of 160.996 million which ranks 8th in the world and population density of 1033 people per square [5]. Children who are one of the most vulnerable parts in society have an important role in the health system for each country. Therefore, Infant Mortality Rate (IMR) is a powerful indicator on the nation's health because it is often associated with other health factors and it is also considered as a serious challenge for health systems in developing countries like Bangladesh. Infant mortality is the death of an infant before his or her first birthday. The IMR, defined as the number of newborns dying under a year divided by the number of live births during a year, has in the past been regarded as a highly sensitive (proxy) measure of population health[1]. It is also part of United Nations development index and therefore becomes important to describe it for purposes of evaluation and outlining of health policies. Infant Mortality Rate varies based on changes in socioeconomic, geographic, biologic, demographic, cultural and environmental factors [2]. It is closely related to the well-being of the community and its development. According to these studies, maternal mortality ratio, newborns protected against tetanus, birth rate, crude, population, female, rural population growth, adolescent fertility rate, improved sanitation facilities, improved water source, labor force participation rate, female, population growth, low-birth weight babies, prevalence of anemia among pregnant women and children, smoking prevalence females, Pregnant women receiving prenatal care , GDP per capita (constant LCU), GDP per capita growth, employers female, parental education factors are commonly measured and demonstrated in World Bank data repository where IMR is relatively more serious in Bangladesh which have suffered with health inequalities, unsuitable socioeconomic condition and unfavorable environmental factors tends.

Data mining is the process of sorting through large data sets to identify patterns and establish relationships to solve problems through data analysis and it holds great potential to improve health systems. In Bangladesh health system there is a huge amount of untapped data that can be turned into useful information. In this proposed system, we are tend to investigate more reasonable factors out of literature and data repository that have been unexplored and to build predictive models using high risk factors to solve infant mortality problem through these untapped data analyses [3].

This paper attempts to identify the factors which influence number of children dying before reaching age one year in Bangladesh and also establishes temporal trends in attribute ability by factor as well as Bangladesh for the period 1990-2013; data is used from the "World Bank Indicator (WDI) 2015. In the following, we present the existing works completed under the same roof. Next, we discuss our proposed system following by the system implementation, discussion on result and analyses and finally, concluding with some insights.

II. LITERATURE REVIEW

Infant mortality is a powerful indicator on the nation's socio-economic and health because it is often associated with the number of different analytical frameworks in which affect the outcome through some proximate determinants factors namely socioeconomic variables (i.e. parental education, socio- economic status, place of residence , parental work status) , demographic factors(i.e. age of the mother at the time of birth, adolescent fertility, birth order, birth interval, sex of the child),environmental factors(i.e. source of drinking water, toilet facilities, housing construction material, type of bathing facility) health-seeking behavior (i.e. prenatal care; place of birth; tetanus injection before birth and contraceptive use; vitamin A dose received; low-birth weight babies) [4][5][6].

The multivariate Proportional Hazards framework is a flexible parametric framework for analyzing the impact

variables on infant [7]. Cross-tabulation, simple and multiple logistic regression techniques have been used to investigate the important factors of infant (below 12 months) mortality of Bangladesh, data from the "Bangladesh Demographic and Health Survey (BDHS) 2011" [5]. All these studies, which census data to measure the effect of socio-economic factors, environmental demographic factors, health and maternal behaviors to be significantly related to infant and child mortality [5]. To assess the effects of maternal and environmental factors on infant mortality in Kenya, data selected and analyzed from the" Kenya Demographic and Healthy Survey 2014" was used Mosley and Chen framework (1984, they distinguished fourteen proximate determinants and categorized them into four groups: maternal [fertility] factors, environmental sanitation factors, availability of nutrients to the infant, injuries, and personal illness control factors) to conceptualize the research and for data analysis methods, descriptive statistics and logistic [2].

One of the socio-economic variables is parental education which is strongly linked or related to infant mortality risk and known to have a reverse relationship with infant mortality [2]. Type of place of residence (urban, rural) where poor socio-economic status lead to under privileged health and living conditions especially for mothers and their children [2]. Africa is confronted with numerous health of its population; especially those living in urban areas have access to modern health facilities. Maternal and child mortality is very high, and mostly caused by long distances to health facilities shortage of trained health professionals, high cost of healthcare, and inadequate funding to purchase modern equipment for hospitals, build new hospitals , and recruit enough qualified doctors or specialists. As concerns the demographic variables (Age of the mother at the time of birth) the patterns for mortality by maternal age and birth generally take a U-shaped .Higher mortality risks occurring to children who are born to both relatively old and young mothers. Receiving tetanus injection and vitamin A has strongly highly significant influence on infant mortality [2][5]. Environmental factors on the households, source of water and sanitation facility has a significantly negative influence on infant mortality. Source of water and sanitation facility will have two categories safe/unsafe water and no sanitation facility/presence of sanitation facility respectively. In the household, both two variables are constructed from the mothers' response question on type of source of drinking water and sanitation facility [8]. HIV prevalence among adults aged 15 to 49 years, GDP per capita(in US$), adolescent fertility rate, female literacy rate, maternal mortality, birth rate, crude (per 1,000 people), health expenditure per capita (current US$), was the highest relative risk factors of infant mortality [9][10].

III. Proposed System

The proposed system works on infant mortality to explore risk factors and builds predictive models on infant mortality in Bangladesh. Figure1 shows the proposed system architecture. The system consists of several steps and includes important components. In the following, we describe each component to provide more clear view of the system.

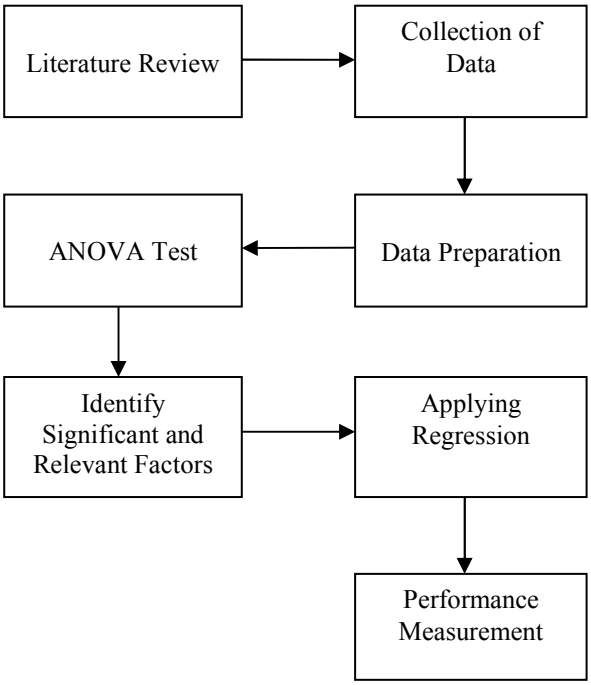


Figure1. Proposed System Model

A. Literature Review

We explore a significant amount research papers on infant mortality accessible via Google Scholar, IEEE, Willy, Elsevier, ACM Digital Library, and Springer Open. Later, the number of papers is narrowed down to relevant papers based on reviewing paper title, abstract and keywords. Then, we dive deep into the rest of the papers to identify various factors that have significant effects on infant mortality and involve both direct and indirect techniques to investigate infant mortality.

B. Source of Data

This study utilizes the data extracted from the "World Development Indicator, 2015 (WDI-2015) "database for Bangladesh [11]. The WDI represents the most accurate development data for each country and for this research the database includes data spanning up to 55 years- from 1960 to 2015.

C. Data Preparation

The database for this study includes data up to 55 years from 1960 to 2015. However, not all indicators have full values across these years. So, we filter out some indicators based on the number of data available in the repository. That eventually brings some indicators for the period 1990-2013.

D. ANOVA Test

After the first attempt on filtering indicators, we analyze significant testing on risk factors of infant mortality. We apply Analysis of Variance (ANOVA) test with an F-statistic measurement.

E. Identifying Significant and Relevant Factors

After ANOVA test, we obtain some significant and relevant factors for infant mortality used in building the model. In the data set we use in this research, some factors that are identified in the literature are not available; so we exclude them from the primary identified list. To investigate the factors that are absent in the literature but could have importance in the real life problems, we dig deeper into the data set and identified some factors that are significant across the IMR. However, not all the significant factors are relevant to the IMR; we exclude them as well. Finally, we include the filtered list of factors identified in the data set to the factors list available in the prior stage.

F. Building Regression Models

Once the factors are identified, we apply stepwise linear regression method using identified significant and relevant factors. Linear Regression is a well known supervised learning algorithm which is used in prediction problems. It finds the target variable by finding a best suitable fit line between the independent and dependent variables. The best fitted line has minimum error from all the points and has high efficiency. The stepwise linear regression is widely used methods for checking the linearity of the correlation among significant factors towards building models and often used as base model to provide insights and to compare with some other prediction algorithms. In this paper, we aim to find the base models with some of its variations and compare among them.

The stepwise regression model consists of considering one or more factors analyzed together and check whether the whole model becomes significant. Following the process, we build enormous amount of models and check which models have higher performance over others based on the predefined performance measurement criteria. A regression equation of the form is shown in equation (1).

$$y = \varepsilon + x_1\beta_1 + x_2\beta_2 + \cdots + x_k\beta_k \qquad (1)$$

where the value of a dependent variable y in terms of a set of k observable variables in $x = [x_1, x_2,\ldots,x_k]$ and an unobservable random variable ε.

E. Performance Measurement Criteria

Measuring performance is a challenging task. We choose some popular error measurement formulas for analyzing the performance of the built models. Those are: Mean Absolute Error (MAE), Mean Relative Error (MRE) and Mean Squared Error (MSE).

$$MAE = 1/n\sum (|\, y_a - y_p|)$$

$$MRE = 1/n\sum((|y_a - y_p|)/y_y)$$

$$MSE = 1/n\sum(|\, y_a - y_p|)^2$$

In the above equations, y_a denotes the actual value from the data set and y_p is the predicted value generated from the model. Moreover, we observe the model R-squared and Adjusted R-squared values for the accuracy of the model.

IV. IMPLEMENTATION

A. Data Preparation

The WDI-2015 dataset contains 1350 indicators and includes data up to 55 years from 1960 to 2015. Indicators are the standards based on which the ups and downs of a country can be evaluated. In data preparation, we extract approximately 650 indicators based on the availability of 25 data for each indicator where the indicators may or may not be completely related to the target problem we deal with in this paper. We ignore indicators that have very small amount of data that would have less significance in building significant models. Indicators are analogous to our future identified factors.

B. Identifying Significant Factors

Among 650 indicators, we are only interested in significant ones, have sufficient data and relevant to the problem. Therefore, we narrow down the indicators further to 550 where data were available for the period of 1990-2013. From the dataset, we apply ANOVA to identify significant factors. Setting the significant testing with p-value < 0.1, we identified 15 factors that do not all exist in literature but have significance on infant mortality.

Further, based on the proposed method, we exclude the significant but irrelevant factors identified from the dataset. The remaining significant factors from the set are added to the list of factors identified in the literature. That becomes totally six factors to be suitable for building the regression model. The histograms of these factors are displayed in Figure 2 to Figure 7.

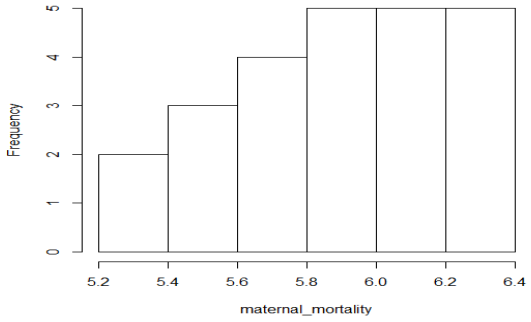


Figure 2: Histogram of Maternal Mortality Rate

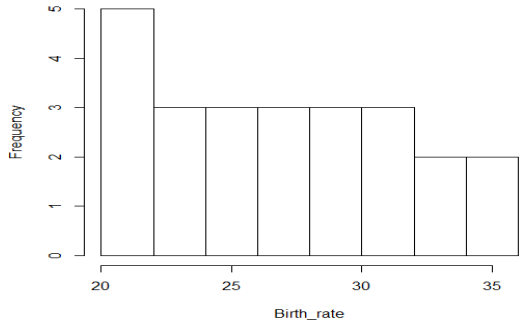


Figure 3: Histogram of Birth Rate

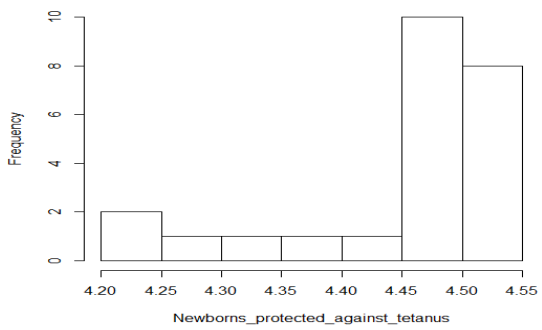


Figure 4: Histogram of Newborns Protected Against Tetanus

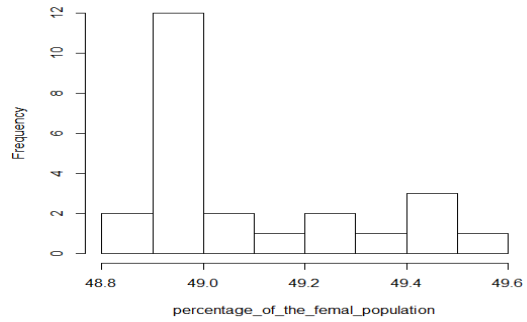


Figure 5: Histogram of Percentage of Female Population

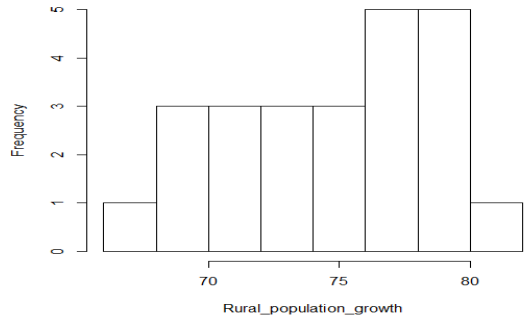


Figure 6: Histogram of Rural Population Growth

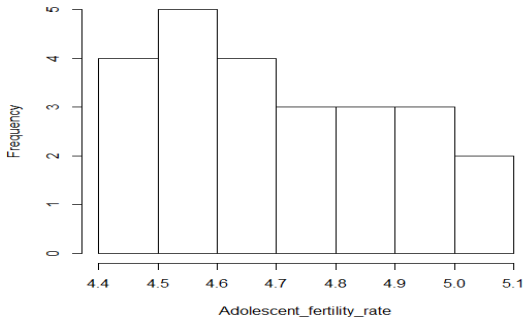


Figure 7: Histogram of Adolescent Fertility Rate

C. Building Models

After stepwise regression analysis we obtain a model consisting six factors: maternal mortality ratio (MMR), newborns protected against Tetanus Injection (TI), Birth Rate crude (BR), Population Female (PF), Rural Population Growth (RPG), Adolescent Fertility Rate (AFR). A set of coefficients are determined from the regression model are shown in Table I and the regression model is presented in equation (2).

$$IMR = 1793575.52 + (135.38 * MMR) - (101.95 * TI) + (12927.15 * BR) - (36006.51 * PF) - (3244.07 * RPG) + (411.94 * AFR) \quad (2)$$

TABLE I: ASSOCIATION BETWEEN INDICATORS AND INFANT MORTALITY

Indicators	Coefficients	P-value
Maternal mortality ratio (modeled estimate, per 100,000 live births).	135.38	0.043
Newborns protected against tetanus (%)	-101.95	0.017
Birth rate, crude (per 1,000 people)	12927.15	0.000
Population, female (% of total)	-36006.51	0.000
Rural population growth (annual %)	-3244.07	0.002
Adolescent fertility rate (births per 1,000 women ages 15-19)	411.94	0.084
Intercept	1793575.52	0.000
R-squared	1	
Adjusted R-squared	1	

Based on the model built using linear regression, we plot residuals vs fitted graph in Figure 8. The normal Q-Q graph has been depicted in Figure 9.

.

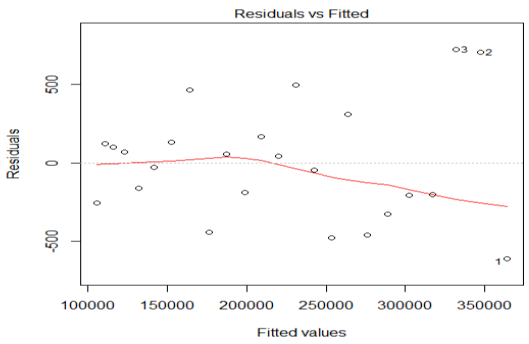


Figure 8: Residual vs Fitted

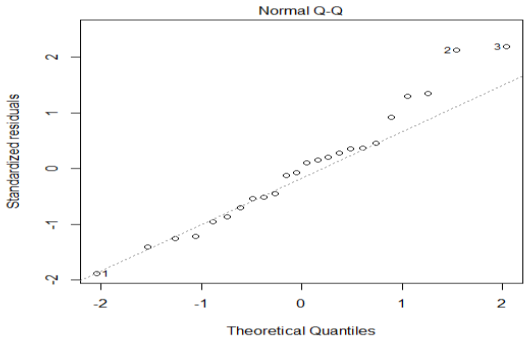


Figure 9: Normal Q-Q Plot

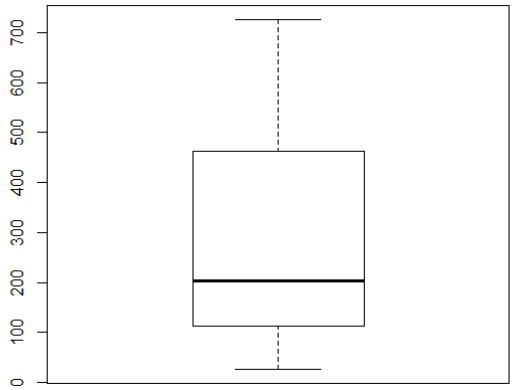


Figure 10: Boxplot of MAE

V. Result and Discussion

The results of main determinant of infant mortality rate in Bangladesh (1990 to 2013) are presented in Table II. The result of R2 statistics showed that the independent variables of model were good predictors for infant mortality rate and also showed that that there were high correlation among dependent variable and explanatory variables which is 1. Regression model parameters displayed in Table I yields from k-fold cross validation where we use k = 4.

TableII : Sample division for K-fold cross validation

Case No.	Testing Set	Testing Set
1	1996-2013	1990-1995
2	1990-1195,2002-2013	1996-2001
3	1990-2001,2008-2013	2002-2007
4	1990-2007	2008-2013

From the above 4 cases, we obtain four different models using the training set and validated using testing set. The resultant values based on the performance criteria have been shown is Table III in row 2. For a deeper analysis, we take two best models from the 4-fold cross validation and the coefficients of these models are applied to the base model to check whether the model has any effect on that particular set of coefficients. The results from those models are presented in row 3 and 4 respectively.

Table III: Performance Comparison

No.	Model	Mean Absolute Error(MAE)	Mean Relative Error(MRE)	Mean Squared Error(MSE)
1	Base	282.921	0.00125	124314.095
2	4-fold Cross Validation	3507.004	0.01051	14824769.5
		461.860	0.00181	328821.992
		262.719	0.00151	94283.755
		1774.334	0.01545	3878313.33
3	Best1	287.985	0.00128	126525.711
4	Best2	323.661	0.00146	164003.317

From Table III, we see that our base model provides more accuracy based on the performance criteria set. The box plots of the errors are presented in Figure 10 to Figure 12. The box plots have shown the error rates are minimal and within range.

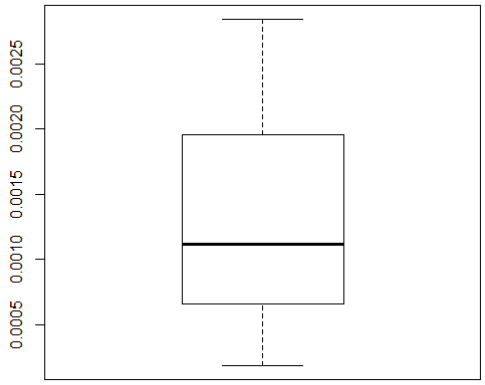


Figure 11: Boxplot of MRE

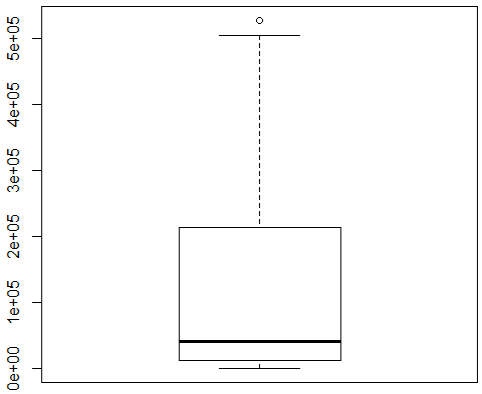


Figure 12: Boxplot of MSE

According to the model, decreasing values in variables such as maternal mortality ratio, birth rate, crude (per 1,000 people), population, female (% of total), rural population growth (annual %), adolescent fertility rate (births per 1,000 women ages 15-19) caused the infant mortality reduction. Tetanus injection is highly significant effect on risk of infant mortality. The results of this study show that tetanus injection is an inverse association with infant mortality. Decreasing rural population growth, population female were associated with a significant lowering of infant mortality rate. The empirical analysis shows that the maternal mortality, adolescent fertility rate and birth rate have a

positive effect on infant mortality rate and infant mortality has been decreased by decreasing values of these variables. In this study, factors such as sanitation facilities, water facilities, Labor force participation rate female, GDP per capita, GDP per capita growth were not significantly related to the infant mortality rate.

VI. Conclusion and Limitation

Infant mortality rate is commonly used as one of the most important indicators of the level of socio-economic development of our country. This indicator is associated with well-being of a population and the one of the main items of the millennium development goals. So, the purpose of this study is to assess the effect of maternal mortality ratio, adolescent fertility rate, birth rate, rural population growth, population female on mortality rate of infant in Bangladesh. This study is designed as a predictive model to identify the factors that significantly influence the risk of infant mortality in Bangladesh. The cross validation and multiple stepwise regression analysis suggests that among the 16 factors only 6 factors are significant. It is recommended that the future study focuses on the predictors of provincial infant mortality rate in Bangladesh along with comparison of the proposed base model with some other popular prediction algorithms.

This study has some threats to the validity. To extract data from WDI-2015 for the period 1990-2013, some data were not available during the period and could not be possible to include them due to the absence of most relevant indicators for infant mortality. For instances, parental education, receiving vitamin A, smoking prevalence female, pregnant women receiving parental care, low birth-weight and similar. The performance can be further investigated by filling the missing values of the factors identified. Overall, the proposed model has a good accuracy on the data set available for identified factors from literature and the WDI-2015 data source, and provides insights on the changes in infant mortality rate with the change in significant factors responsible for the cause.

References

[1] D D Reidpath, P Allotey, *Infant mortality rate as an indicator of population health, J* Epidemiol Community Health 2003;57:344–346, 7 May 2002.

[2] Margaret Nduta Miringu, *Effect of Maternal and Environmental Factors on Infant Mortality Kenya,* Google Scholar, Q50/74579/2014, November 2016.

[3] Senthilkumar D, *Paulraj S, Prediction of Low Birth Weight Infants and Its Risk Factors Using Data Mining Technique*, Proceedings of the 2015 International Conference on Industrial Engineering and Operations Management Dubai, United Arab Emirates (UAE), March 3 – 5, 2015

[4] Clive J. Mutunga, *Environmental Determinants of Child Mortality in Kenya,* World Institute for Development Economices Research, Research Paper No. 2007/83 , December 2007.

[5] Mohammad Sadek, *Infant Mortality Factors of Bangladesh,* LUND University,Sweden,publication/8878528, 6 June 2016

[6] Abdul Hamid Chowdhury, *Determinants of Under-Five Mortality in Bangladesh,* Open Journal of Statistics, 2013, 3, 213-219, Article ID: 33245, June 2013.

[7] Joshua Kembo1 Jeroen K. Van Ginneken2, *Determinants of infant and child mortality in Zimbabwe: Results of multivariate hazard analysis,* Demographic Research, Volume 21, Article 13, Pages 367-384, 01 October 2009.

[8] Puritym, Njagi, *Household Environmental Factors And Child Mortality In Kenya*, Studies and Research Institute Universiy of Nairobi, November 2007.

[9] Benn KD Sartorius and Kurt Sartorius, *Global infant mortality trends and attributable determinants*, Sartorius and Sartorius; licensee BioMed Central Ltd. 2014, 19 October 2014.

[10] Sara Emamgholipour sefidashti1, Majid Nakhaei2, Ali Kazemi Karyani3, Sadegh Ghazanfari4, *Decomposition Socioeconomic Inequality infant Mortality in EMRO Countries*, Original Article (Pages:749-756), 12 Jun 2015.

[11] WorldDevelopmentIndicator,https://data.worldbank.org/country/bangladesh, 29 December 2015.

Design of Twenty-One Level Modular Multilevel Converter Using a Three-level Submodule Block

Ashraf Siddiquee*, Shirajum Munira, Md. Shamim Reza
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka-1205, Bangladesh
*email: ashraf.siddiquee.1996@gmail.com

***Abstract*— In this paper, a modular multilevel converter of twenty one level is designed by using a three level submodule block. The submodule block can be half-bridge or full bridge. In a half bridge submodule block, the switching configuration is more efficient than the full bridge submodule block. But there is problem of DC fault current. Half bridge submodule block is not capable of blocking the DC fault current. In this paper, we have designed such kind of submodule block which is a combination of half bridge and flying capacitor. A capacitor which is floating with the half bridge circuit that is called flying capacitor. We have designed many modular multilevel converters of different levels using this submodule block but the twenty one level modular multilevel converter gives the best output. A simple control mechanism is implemented in this design. So, our proposed design of modular multilevel converter is simple and it shows better performance.**

***Keywords*-Half-Bridge; Flying Capacitor; Modular Multilevel Converter; DC-fault current; Model Predictive Control; Submodule; THD; Spectrum Analysis**

I. INTRODUCTION

In the previous years, for power transmission, however it is long distance power transmission or short distance power transmission, Alternating Current(AC) power transmission was used. However nowadays it is found in research that for long distance power transmission (more than 700km), Direct Current(DC) power transmission is more cost efficient and safe than Alternating Current power transmission. In the DC power transmission, high voltage DC is used. For high voltage DC, low voltage DC has to be converted to high voltage DC. However transformer can only operate in AC. So, for voltage level conversion, DC has to be converted to AC and AC has to be converted to DC. For AC-DC conversion, rectifier is used and for DC-AC conversion, inverter is used. For making pure sinusoidal wave with low harmonic distortion, number of level of the sinusoidal wave has to increased[4]. So, modular multilevel converter is used.

Nowadays, for modular multilevel converter half bridge or full bridge submodule blocks are used. The main drawback of the half bridge submodule block is that the submodule block cannot have any DC fault blocking capability. So, in this paper, the modular multilevel converter is designed using a such kind of submodule block which is a combination of half bridge and flying capacitor. This submodule is a five level submodule[1,2,3], however there is a problem of two level in this five level, which are not stable. So, in this paper the submodule block is used as a three level submodule block. This three level submodule block has the capability of DC fault current blocking.[4,5,6]

Moreover, in this paper, for controlling purpose comparison of reference sine wave is used. Comparing the reference sine wave, by using digital signal processing, the controlling mechanism is implemented in Matlab Simulink. In the previous papers[7,8,9,10], PI control mechanism was implemented. However this control mechanism is not so good for this inverter design due to high high computational time. Again, model predictive control algorithm[11,12] was trying to implement, however that controlling mechanism is too much time consuming for processing the code. For this reason, a simple mechanism is implemented by comparing the reference sinusoidal wave. So, our designed module multilevel converter is simple, cost efficient and better performance.

II. TOPOLOGY OF MODULAR MULTILEVEL CONVERTER

The main topology of three phase modular multilevel converter is shown in figure 1.

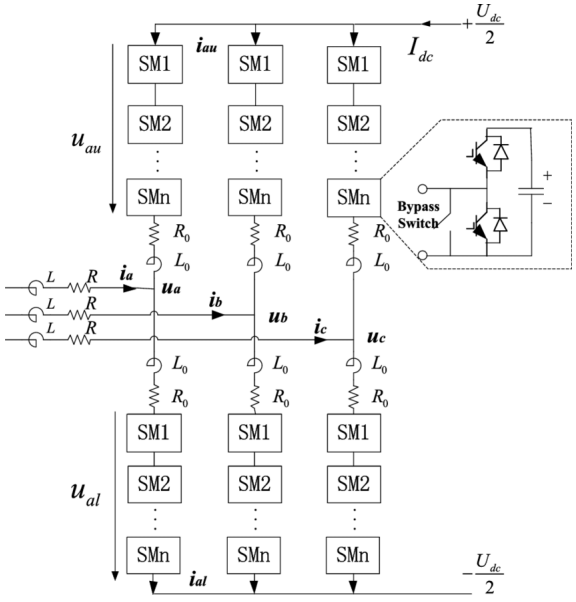


Fig. 1. Topology of Modular Multilevel Converter

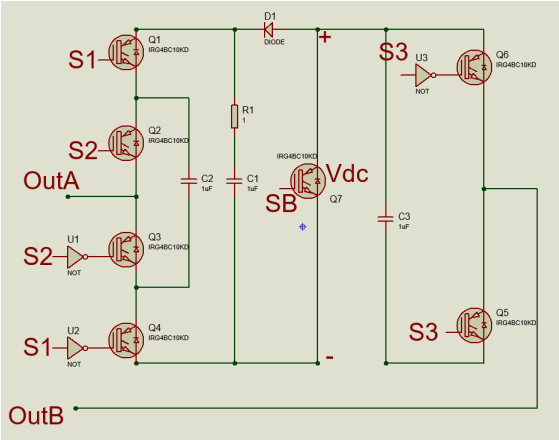


Fig. 2. Three Level Sub-module Block Schematic

Typical MMC topology consists of three-phase legs, each divided in two arms (upper and lower). Further, each arm is a stack of N series connected identical submodules and an arm inductor arm L. A submodule can perform three modes of operations: (i) inserted, (ii) bypassed, and (iii) blocked. When an SM is inserted, it permits current to flow through capacitor and to charge/discharge it. In bypassed mode, switches of SMs operate in such a way that it allows current flow through SM, however bypassing the capacitor, the capacitor remains charged and in block mode of operation, all the switches of SM are off, and current can flow only through anti-parallel diodes, if possible.Any particular submodule can be inserted or bypassed or blocked at any instant, based on switching commands, and thus, any number of SMs can be inserted at a time corresponding to switching scheme, and therefore, with appropriate switching schemes, an alternating voltage can be generated.

III. TOPOLOGY OF THREE LEVEL SUBMODULE BLOCK

In this paper, the modular multilevel converter is designed using a three-level submodule block. The circuit diagram of the three-level submodule block is given in figure 2. This submodule block has two parts. One is a flying capacitor unit and another is half bridge submodule unit. In a previous paper[13], it is proposed that the submodule block will work as a five-level submodule block. However in Matlab Simulink simulation it is seen that the submodule block can work for three-level properly. Other two levels are not stable, that why in this paper, it is used as a three-level submodule block.

There are total 8 switches in this submodule block. They are s1,s2,s3,sB and their compliment switches. So, we have to control 4 switches, then the other complimentary switches will be automatically controlled. Moreover, between these 4 switches, sB will be always on. For this reason, the main concern is to control the other three switches. One of the most useful features of this submodule block is that it has DC fault blocking capability and the full mechanism of the blocking the fault current is described in the paper[14].

TABLE I
SUBMODULE PARAMETERS

Component	Value
C1	$1\mu F$
C2	μF
C3	μF
Battery	40Volt
R1	1ohm

TABLE II
SWITCHING SEQUENCE AND OUTPUT OF THE SUB-MODULE BLOCK

s0	s1	s2	$V_o ut$	Level
0	0	0	0	0
1	1	0	$V_{dc}/2$	1
1	1	1	V_{dc}	2

TABLE III
REFERENCE SINUSOIDAL WAVE CONFIGURATION

Amplitude	200 Volt
Frequency	50 Hz

The parameters of the submodule block is given in table I. The switched controlling output of the submodule block is given in Table II.

IV. CONTROL MECHANISM

In the previous papers to control the modular multilevel converter, PI control mechanism is used. There are many control mechanisms. The popular control mechanisms are model predictive control mechanism, normal switch control mechanism,PI control, PID control etc[15,16,17]. Model predictive control mechanism one of the most popular mechanism, and it is highly used in the control mechanism of the modular multilevel converter. However there is a drawback in the model predictive control mechanism. If the number of levels of the modular multilevel converter is increased, then model predictive control cannot work properly due to high computation. So, model predictive controlling mechanism needs high computation. For this reason, though this control mechanism is used in the previous papers in this paper, this control mechanism is omitted. In this paper, a new control mechanism is proposed, which is very simple to compute.

Model predictive control is used in those cases where the output does not follow any kind of pattern.[18] However in a inverter, the output is known. For this reason, a reference sinusoidal wave is used to predict the output. The full control mechanism is described in this section.

A. Reference Sinusoidal Wave

For controlling purpose, a reference sinusoidal wave is used. The reference sinusoidal wave configuration is given in table III.

B. Controlling Unit

In the control unit, a processor is needed to compute the mathematical calculation. The sampling frequency of the computational time is $2.5 * 10^{-5}$ seconds. There is total 21 levels in this modular multilevel converter. From 0 volt to 200 volt there is 11 levels, from 0 volt to -200 volt there is 11 levels and 0 is a level. That's why there is total 21 level. The more the level the less the harmonics, however if the number of levels is increased, then there is a lag in the output sinusoidal wave. For this reason, the best output is found in the 21 levels. Further it will be discussed more elaborately.

Every sampling time, the reference sinusoidal wave is compared with the levels of output voltage. Then it is decided that in which level the output wave should come. In this process, this sinusoidal wave is generated.

V. BLOCK DIAGRAM AND OTHER ANALYSIS

The circuit is divided into mainly two sections. One is control unit another is sub-module unit. In the control unit, there are two units, one is comparison and level decision unit and another is switching sequence generator unit[19].

A. Control Unit

1) Comparison and Level Decision Unit: In this unit, every sampling time, reference sinusoidal wave is located in which level of the output sinusoidal wave, it is decided. So, in this section comparison of the level with reference sinusoidal wave is done and the decided level is the output of this section.

2) Switching Sequence Generator Unit: In this unit, using the number of level, a switching sequence is generated for the submodule blocks. So, for 21 levels, there is total 10 submodule blocks. Switching sequence for each block is generated in this unit and the switching sequence is sent to the submodule blocks.

B. Submodule Blocks Unit

In this unit, for 21 levels there are total 10 submodules[20]. If there is n submodules, then the number of levels is 2n+1. In this paper, we used total 10 submodule blocks. These 10 blocks are divided into two sections, the upper 5 submodule blocks are for the positive half cycle of the output sinusoidal graph and the lower 5 blocks are for the negative half cycle of the output sinusoidal graph[13].

The DC input of the each sinusoidal graph is 40 volt. If we want 140 volt, then the situation of the submodule blocks is shown in Table IV. Again for -80 volt output, then the configuration of the submodule blocks is shown in Table V.

VI. OUTPUT CIRCUIT

In the output side a series RLC circuit is used. In the RLC circuit the output parameters is set using the written values in table VI. The output circuit is given in figure 3.

TABLE IV
SUBMODULE BLOCKS CONFIGURATION FOR 140 VOLT OUTPUT

Submodule	Switching Sequence	Level	Output(volt)
Submodule1	1,1,1	2	40
Submodule2	1,1,1	2	40
Submodule3	1,1,1	2	40
Submodule4	1,1,0	1	20
Submodule5	0,0,0	0	0
Submodule6	0,0,0	0	0
Submodule7	0,0,0	0	0
Submodule8	0,0,0	0	0
Submodule9	0,0,0	0	0
Submodule10	0,0,0	0	0
Total Voltage Output			140

TABLE V
SUBMODULE BLOCKS CONFIGURATION FOR -80 VOLT OUTPUT

Submodule	Switching Sequence	Level	Output(volt)
Submodule1	0,0,0	0	0
Submodule2	0,0,0	0	0
Submodule3	0,0,0	0	0
Submodule4	0,0,0	0	0
Submodule5	0,0,0	0	0
Submodule6	1,1,1	2	-40
Submodule7	1,1,1	2	-40
Submodule8	0,0,0	0	0
Submodule9	0,0,0	0	0
Submodule10	0,0,0	0	0
Total Voltage Output			-80

TABLE VI
OUTPUT PARAMETERS

Parameters	Value
R	100ohm
L	100mH
C	100uF

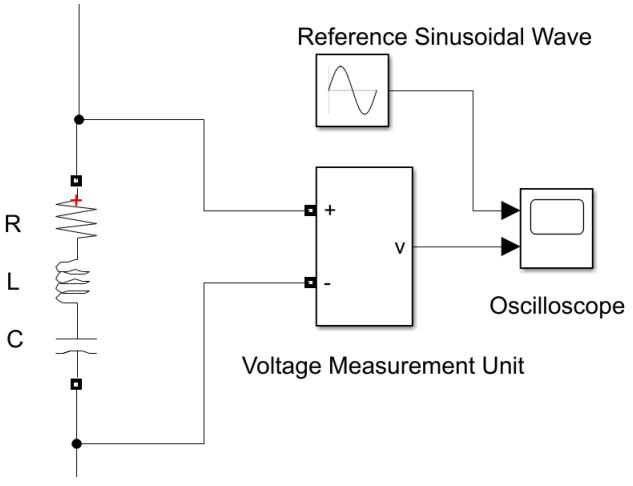


Fig. 3. Output Circuit

In the oscilloscope, there is two input, one is from the output of the Modular Multilevel Converter and another is the reference sinusoidal wave.

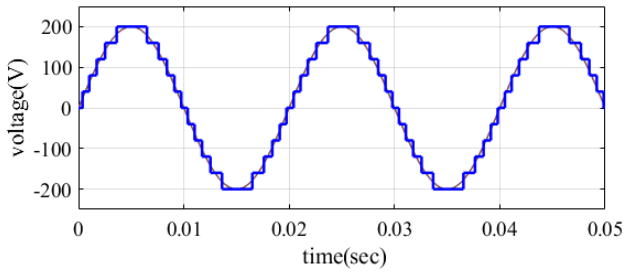


Fig. 4. Output of 11 Level MMC

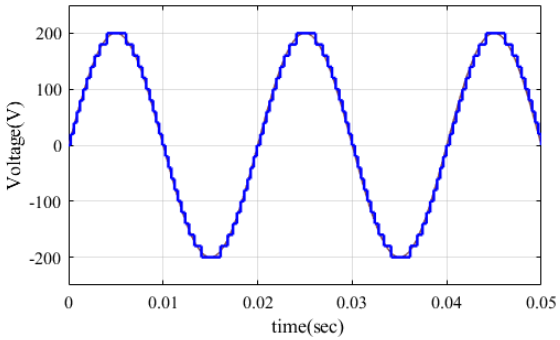


Fig. 5. Output of 21 Level MMC

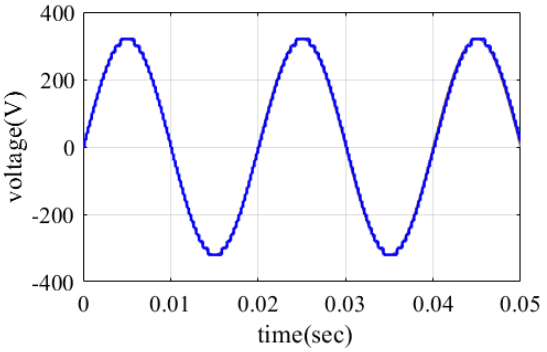


Fig. 6. Output of 33 Level MMC

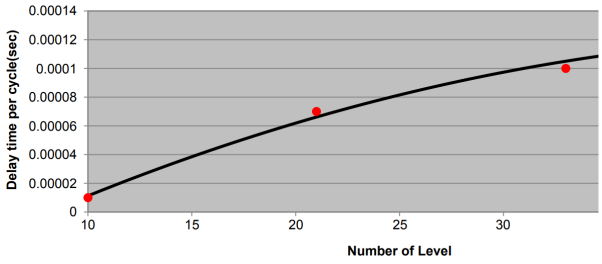


Fig. 7. Delay Time Per Cycle versus Number of Level Graph

VII. OUTPUT ANALYSIS

In this paper, the modular multilevel converter is designed for 21 levels. The Output waveform is given in figure 4.

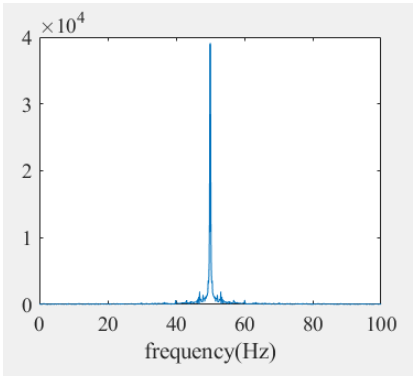


Fig. 8. Spectrum Analysis Graph of 11 Level Modular Multilevel Converter

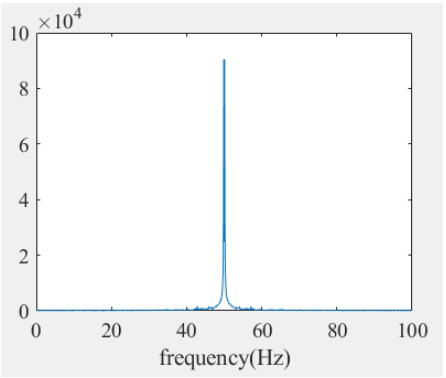


Fig. 9. Spectrum Analysis Graph of 21 Level Modular Multilevel Converter

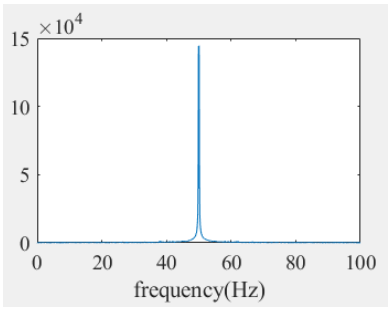


Fig. 10. Spectrum Analysis Graph of 33 Level Modular Multilevel Converter

TABLE VII
TOTAL HARMONIC DISTORTION FOR EVERY OUTPUT OF MODULAR MULTILEVEL CONVERTER

Number of Levels	THD(%)
11	9.42
21	3.97
33	3.42

In this figure, the reference sinusoidal wave is in blue color and the output voltage is in orange color. In this figure, we can see that the output is slightly lagged from the reference sinusoidal wave. This happened because high computation in the processor. However if we decrease the number of

levels, then the output graph will be better. Output graph of 11 levels MMC is given in figure 5. However there is a problem, if we decrease the number of levels, then the Total Harmonic Distortion will be higher. Again, for lower THD if we increase the number of levels, then there will be huge lag in the computation time. For this reason, the output voltage will be too much lagged from the reference sinusoidal wave. The output graph of 33 level MMC is given in figure 6. In this figure, the total harmonic distortion will be lower, however due to high computation time, it is clearly seen that the output wave has huge time lag from the reference voltage. So, lag time of the output sinusoidal wave from the reference sinusoidal wave versus no of levels graph is given in the figure 7. Here we can see that, the more the levels, the more the lag time. After comparing all the matters, total harmonic distortion, time lag etc, we can see that the 21 level modular multilevel converter is best among the others. In this paper, different levels of modular multilevel converter is simulated. From different number levels output, we can see that, the more the level, the better the output. Different spectrum analysis output of different number of levels are shown in figure 8, figure 9 and figure 10. Here it is shown that the best output is for the 33 level modular multilevel converter. Again the total harmonic distortion is determined for every output. The output of THD for every analysis is given in table VII. So, we can see that, the best output is for 33 level. However, in previous it is shown that time lag of the output graph is maximum for the 33 level MMC. So, after all types of analysis, spectrum analysis and time domain analysis, it is decided that the 21 level MMC gives the best output.

VIII. CONCLUSION

This paper proposes a design of twenty-one level Modular Multilevel Converter using a three-level Submodule block. This block is very much helpful for DC fault current blocking. Moreover, in this paper, it also analysed that the best output of a Modular Multilevel Converter is seen for 21 level MMC. Nowadays High Voltage Transmission is very much popular and design of efficient Inverters is very important. So, this paper will help to implement a low cost and more efficient Modular Multilevel Converter. That will help the whole mankind. Further, it is hoped that, practical implementation of the proposed design will be done and then the design will be more practical and more efficient and more beneficial for all of us.

REFERENCES

[1] Fuentes, Ricardo & Dekka, Apparao & Rivera, Sebastian & Wu, Bin. (2016). Modular Multilevel Converter based on 5-level submodule with DC fault blocking capability. 3482-3487. 10.1109/IECON.2016.7793119.

[2] He, Zhixing & Guo, Peng & Shuai, Zhikang & Xu, Qianming & Luo, An & Guerrero, Josep. (2019). Modulated Model Predictive Control for Modular Multilevel AC/AC Converter. IEEE Transactions on Power Electronics. PP. 1-1. 10.1109/TPEL.2019.2895224.

[3] Lee, Joonhee & Jung, Jae-Jung & Sul AE, Seung-Ki. (2019). Balancing of Submodule Capacitor Voltage of Hybrid Modular Multilevel Converter (MMC) under DC Bus Voltage Variation of HVDC System. IEEE Transactions on Power Electronics. PP. 1-1. 10.1109/TPEL.2019.2896336.

[4] Hagiwara, Makoto & Akagi, Hirofumi. (2008). PWM Control and Experiment of Modular Multilevel Converters. Ieej Transactions on Industry Applications. 128. 154 - 161. 10.1109/PESC.2008.4591917.

[5] Perez, Marcelo & Fuentes, Ricardo & Rodriguez, Jose. (2012). Decoupled current control of modular multilevel converter for HVDC applications. 1979-1984. 10.1109/ISIE.2012.6237396.

[6] E. Solas, G. Abad, J. Barrena, A. Ca andrcar, and S. Aurtenetxea, Modelling, simulation and control of modular multilevel converter, in Power Electronics and Motion Control Conference (EPE/PEMC), 2010 14th International, sept. 2010, pp. T290 T296.

[7] Z. Yan, H. Xue-hao, T. Guang-fu, and H. Zhi-yuan, A study on mmc model and its current control strategies, in Power Electronics for Distributed Generation Systems (PEDG), 2010 2nd IEEE International Symposium on, june 2010, pp. 259 264.

[8] R. Lizana, C. Castillo, M. Perez, and J. Rodriguez, Capacitor voltage balance of mmc converters in bidirectional power flow operation, in 38th Annual Conference on IEEE Industrial Electronics Society (IECON), 2012, pp. 49354940.

[9] Crispino, Las & G. B. Rolim, Luis. (2017). Model predictive control of a modular multilevel converter with a reduced number of states. 1-6. 10.1109/COBEP.2017.8257227.

[10] Wang, Yue & Cong, Wulong & Li, Ming & Li, Ning & Cao, Mu & Lei, Wanjun. (2014). Model predictive control of modular multilevel converter with reduced computational load. 1776-1779. 10.1109/APEC.2014.6803546.

[11] R. Li, J. E. Fletcher, L. Xu, D. Holliday, and B. W. Williams, A hybrid modular multilevel converter with novel three-level cells for dc fault blocking capability, IEEE Transactions on Power Delivery, vol. 30, no. 4, pp. 20172026, Aug 2015.

[12] Sharifabadi, Kamran & Harnefors, Lennart & Nee, H.-P & Norrga, Staffan & Teodorescu, Remus. (2016). Design, Control and Application of Modular Multilevel Converters for HVDC Transmission Systems. i-xxv. 10.1002/9781118851555.fmatter.

[13] Bayat, Hasan & Yazdani, Amirnaser. (2019). A Hybrid MMC-Based Photovoltaic and Battery Energy Storage System. IEEE Power and Energy Technology Systems Journal. PP. 1-1. 10.1109/JPETS.2019.2892418.

[14] Verdugo, Cristian & Ignacio Candela, Jose & Blaabjerg, F & Rodriguez, Pedro. (2019). Three-Phase Isolated Multi-Modular Converter in Renewable Energy Distribution Systems. IEEE Journal of Emerging and Selected Topics in Power Electronics. PP. 1-1. 10.1109/JESTPE.2019.2895465.

[15] Riar, Baljit & Geyer, Tobias & Madawala, Udaya. (2015). Model Predictive Direct Current Control of Modular Multilevel Converters: Modeling, Analysis, and Experimental Evaluation. Power Electronics, IEEE Transactions on. 30. 431-439. 10.1109/TPEL.2014.2301438.

[16] Riar, Baljit & Geyer, Tobias & Madawala, Udaya. (2015). Model Predictive Direct Current Control of Modular Multilevel Converters: Modeling, Analysis, and Experimental Evaluation. Power Electronics, IEEE Transactions on. 30. 431-439. 10.1109/TPEL.2014.2301438.

[17] Vahedi, Hani & metri, julie & Kanaan, Hadi & Al-Haddad, Kamal. (2016). Real-Time Implementation of Model Predictive Control on 7-Level Packed U-Cell Inverter. IEEE Transactions on Industrial Electronics. 63. 10.1109/TIE.2016.2542133.

[18] Wu, Zhi & Chu, Jiawei & Gu, Wei & Huang, Qiang & Chen, Liang & Yuan, Xiaodong. (2018). Hybrid Modulated Model Predictive Control in a Modular Multilevel Converter for Multi-Terminal Direct Current Systems. Energies. 11. 1861. 10.3390/en11071861.

[19] Perez, Marcelo & Fuentes, Ricardo & Arancibia, David & R. Espinoza, Jose & Rodriguez, Jose. (2015). Decoupled Current Model and Control of Modular Multilevel Converters. IEEE Transactions on Industrial Electronics. 62. 1-1. 10.1109/TIE.2015.2405900.

[20] Vazquez, Sergio & Leon, Jose & Franquelo, Leopoldo & Carrasco, J.M. & Martinez, O & Rodriguez, Jose & Cortes, Patricio & Kouro, Samir. (2009). Model Predictive Control with constant switching frequency using a Discrete Space Vector Modulation with virtual state vectors. Industrial Technology, 2009. ICIT 2009. IEEE International Conference on. 1 - 6. 10.1109/ICIT.2009.4939728.

Comparative Study of Enhancing Stability of a Wind Farm connected to the Grid using PID Controller based STATCOM and Capacitor Bank

*Ajmayen Fayek, Shah Md. Salimullah, Md. Shakhawath Hossain, Rakib Hossain, Md. Shadman Hasan Shakib, Ariful Islam Anik, and Moynul Hossen Khan
Department of Electrical & Electronic Engineering
Bangladesh Army International University of Science & Technology (BAIUST)
Comilla - 3501, Bangladesh
*E-mail: ajmayeneee011@gmail.com

***Abstract*—Comparative study of enhancing stability of a wind farm that is grid connected has been done by moderating voltage fluctuation and achieving reactive power compensation using STATCOM (Static Synchronous Compensator). To control STATCOM a voltage control loop with a PID controller is used. To maintain the control system of STATCOM a PWM technique has been introduced. By using different approach, comparative study regarding stabilization of a wind farm has been performed (i.e. by using capacitor bank only, using STATCOM or using STATCOM along with PID controller during wind speed change. STATCOM with PID controller deals better performance with improved stability.**

Keywords—STATCOM, PID Controller, Capacitor Bank.

I. INTRODUCTION

Recent statistics say that sustainable power source has turned into a period of colossal examination due to constrained characteristic assets. Natural gas (24%) is the fundamental wellspring of vitality in Bangladesh, that is probably going to be exhausted continuously 2020 [1]. Renewable energy sources like wind energy, ocean wave energy[2] etc. are needed to use because these sources of energy are environment friendly.

For power system utilities voltage stability has become a vital concern with respect to the ups and downs in power systems like increase of load, genset reaches reactive power borders [4] etc. In wind turbine generator, induction generator (IG) utilized in light of the fact that it can recoup energy with moderately straightforward controls, it is likewise practical and strong. Since the active power produced by steady speed wind turbines is varied because of the impacts of turbulence, the breeze inclination, the pinnacle shadow and the reactive power request of these squirrel-cage induction generators have likewise fluctuated. Thus, the reactive power and active power varieties can cause change of voltage in the purpose of basic coupling (PCC). Therefore, settled speed wind turbines present testing impacts on the grid performance. In spite of the fact that a bank of capacitors is utilized to keep up reactive power at enduring position, it can't legitimately be controlled to keep the voltage steady. For balancing out the wind generator SVC (static var compensator), unified power flow controller (UPFC), STATCOM with the presence of VSC (voltage source converter) etc. are utilized. In wind farms, STATCOM gives the voltage which is steadier and it is likewise cost-effective. With the ever-increasing use of renewable energy [5] resources, the use of STATCOM has become absolutely essential [6-8].

In this paper, we propose a utilization of six-pulse STATCOM for generating wind energy which is connected to the grid, gives speedier reaction time maintaining prevalent voltage bolster ability. Because of the smooth task of voltage and speedier reaction PID controller have been utilized. The PSCAD simulation comes about demonstrate the relative execution of wind farm utilizing just capacitor bank and utilizing STATCOM or utilizing STATCOM alongside PID controller. Examination of result demonstrates that STATCOM with PID controller offers better execution.

II. MODELLING THE SYSTEM

Wind turbine setup contains the necessary equipment needed to grab the wind's energy. Fig. 1 shows a typical grid-connected wind system where the wind's energy turns 2 or 3 propeller-like vanes prevailing a rotor. The rotor that is associated with the main shaft, simply rotates a generator to generate electricity. For stability enhancement it is very important to design a wind turbine properly.

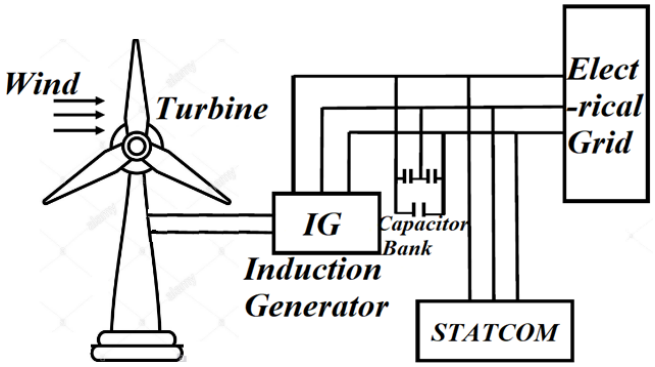


Fig. 1. Induction Generator and Grid connected Wind turbine

III. STATCOM

STATCOM is the device in the family of Flexible AC Transmission Systems (FACTS) used for power flow control, improving transient stability, enhancement of power factor, enrich voltage stability etc. There is a var control mode operation in where STATCOM's output power (reactive) is constant and parameters (different) are independent. Time response of STATCOM is better than static Var Compensator (SVC) [9]. Fig. 2 shows a very simple construction of statcom with six IGBTs and a shunt capacitor. In this study, STATCOM is used to enhance the voltage stability.

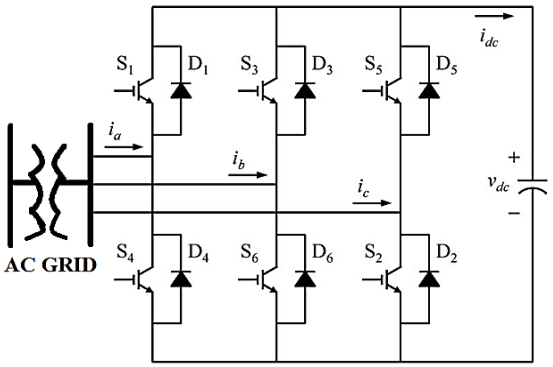


Fig. 2. Circuit Diagram of STATCOM

IV. STATCOM CONTROL STRATEGY

The control strategy basically follows vector control principles [10]. It is applied to the three-phase grid-connected converters. The VSC in STATCOM is produced from a DC link capacitor, in order to reduce converter losses the DC link voltage is controlled by modulating direct axis current of converter. To modulate the transmission of reactive power from STATCOM towards the grid, the voltage of the grid can be controlled by the quadrature current of converter [11]. Lowest fault condition operation rating is the subject to the ratio of connection point's short-circuit power and wind farm rated power [12]. In this paper mainly two types of control strategies are discussed. One is Voltage Control System using PID controller and another is pulse width modulation.

A. Voltage Control Using PID (Proportional-Integral-Derivative) Controller

The PID Controller is a type of control loop feedback system which is mainly used for control systems in light of industrial purpose and continuously modulated control systems. From a research, PID controllers are used in industrial automation at near about 95%. In closed-loop operations, PID maintains the zero error among variable parameters and output [13]. In PID controller there are three terms.

Proportional term is the term where the output changes proportionally with the recent error value. Controller output becomes zero if the error value is zero. When the proportional gain stands high enough the system can become unstable and control action might be too less when the gain is too low.

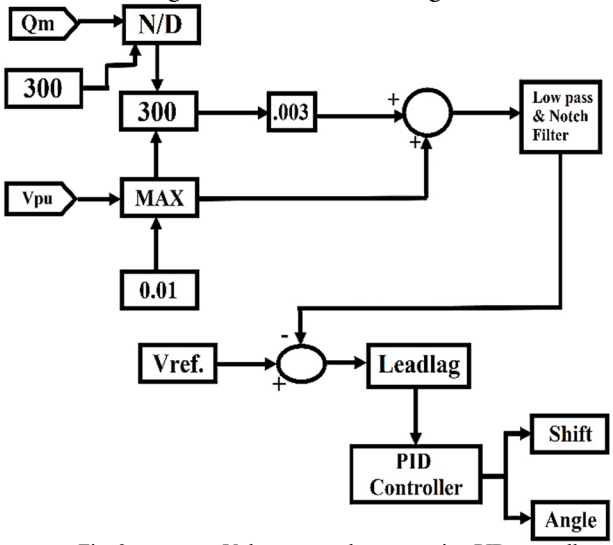


Fig. 3. Voltage control system using PID controller

The integral terminology depends upon both the magnitude and duration of the error. In proportional term, there is a steady state error which is eliminated in this term. It continuously does integration of the error till error value is zero. It has no ability of detecting forthcoming behavior of error.

By anticipating the forthcoming value of error the derivative term resolves the problem. The error is determined from the slope of the error and multiplication of the derivative gain with rate of change. Improving settling time derivative term enhances the stability [13].

For generation of angle order voltage control loop is used. Through the disputation between the reference and filtered voltage, the value of error voltage is found. The filtered voltage passes the reactive power and measured voltage by the notch filter. After that, the error voltage is filtered by PID controller (Fig. 3). In this paper, Voltage Control System has been used both with PID controller and without PID controller. Without PID controller there are voltage fluctuations and the system is unstable. But with PID controller the voltage fluctuations are less and the desired output is gained. This gives a better performance and makes the system stable.

B. Pulse Width Modulation (PWM)

PWM is a system for controlling DC power to an electrical device. It is additionally used to control the power that is provided to the different sorts of electrical devices, for example, AC/DC motors. For the power loss in the exchanging device of the PWM is low. There is for all intents and purposes no currents present when the switch is turn off. When it is on there will be no voltage drop over the switch and power will be exchanged to the load. So the power loss in the two cases will be near zero.

PWM has two principle parts. The initial segment of the PWM is known as the age of the triangular waveforms and the second piece of the PWM is the age of the reference waveforms.

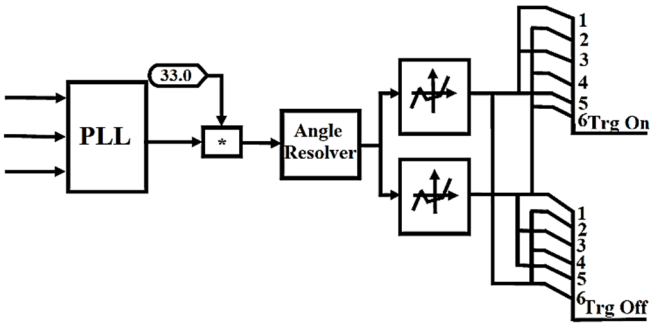


Fig. 4. Part 1: Generation of Triangular waveform.

Part 1: Fig. 4 shows the PLL (Phase Locked Loop) that makes an edge which is called theta and it is multiplied by the carrier frequency. The multiplication result is associated with a client characterized part which is called edge resolver. There is likewise trigger on and trigger off pulse which is utilized for setting off the pulse of the STATCOM.

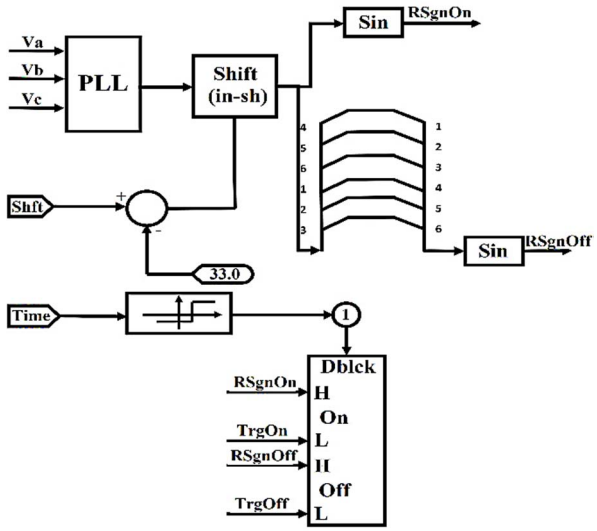


Fig. 5. Part 2: Reference waveform Generation

Part 2: Here again in Fig. 5 theta is made from another Phase Locked Loop (PLL). The yield of the voltage control loop and yield of PLL are associated with a client characterized segment which is meant by six-dimensional arrays.

V. Results and Simulation

In this part, utilizing PSCAD simulation of a wind farm that is connected to the grid alongside induction generator is demonstrated. To repay the reactive power, the wind farm is associated with capacitor bank and STATCOM which is controlled by a PID controller. A similar investigation of stability enhancement is examined. Simulation comes about have been contrasted along and important diagrams.

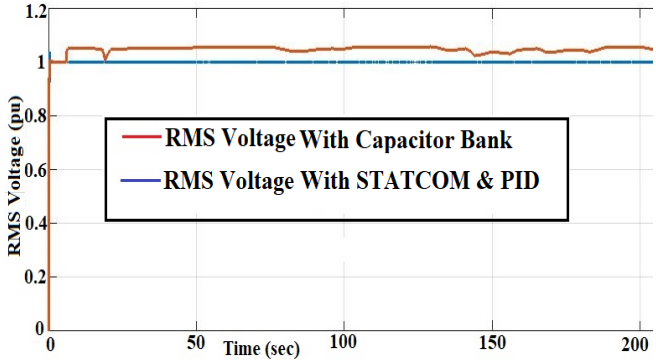


Fig. 6. RMS voltage with capacitor bank & STATCOM

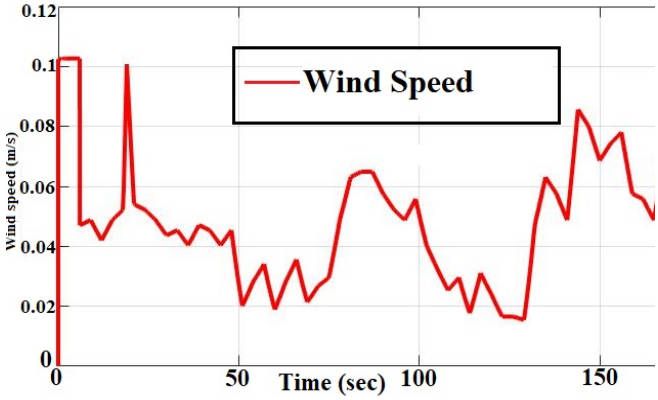


Fig. 7. Wind Speed (m/s)

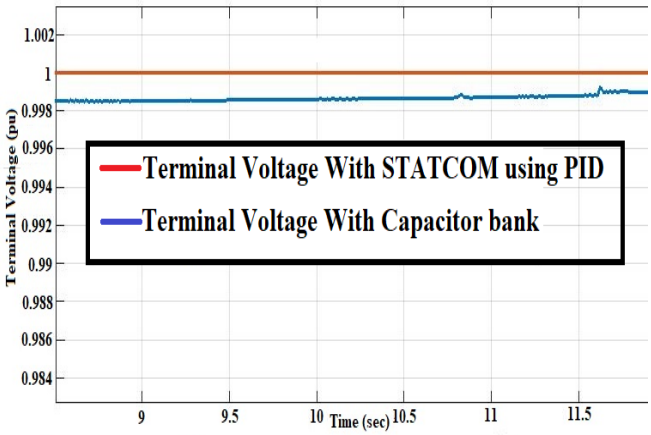


Fig. 8. Induction Generator term end voltage

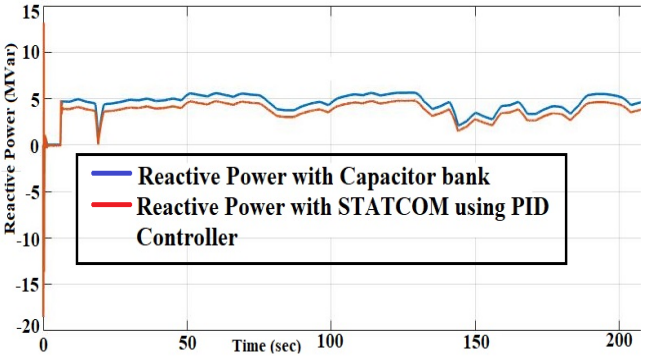


Fig. 9. Reactive power (MVar)

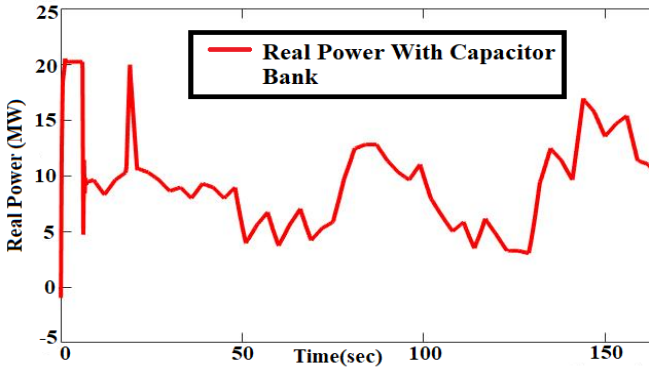


Fig. 10. Real Power with capacitor bank (MW)

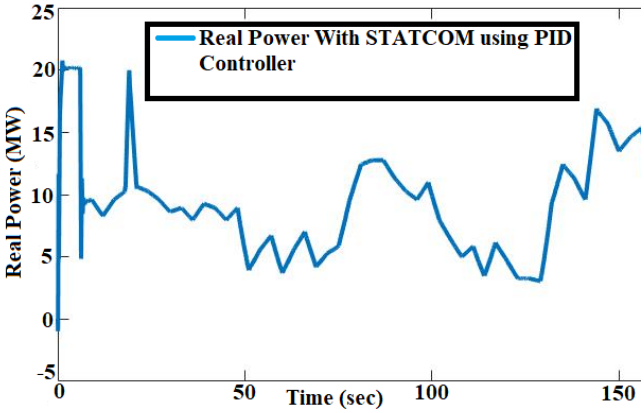


Fig. 11. Real Power with STATCOM using PID Controller (MW)

Fig. 6 demonstrates the RMS voltage information utilizing STATCOM with PID controller and just Capacitor bank. The steadier waveform has been found if there should be an occurrence of STATCOM with PID controller than

Capacitor bank. Fig. 7 shows the nature wind speed variance. Fig. 8 demonstrates the terminal voltage of induction generator utilizing STATCOM with PID controller and Capacitor bank. There is a vacillation estimation of the terminal voltage picked up when capacitor bank is utilized. In any case, when STATCOM alongside PID controller is utilized the terminal voltage vacillation is expelled and it indicates more steady esteem however the wind speed is fluctuating in nature. A superior consequence of a terminal voltage is picked up by utilizing STATCOM instead of Capacitor bank. Fig. 9 demonstrates reactive power with STATCOM and Capacitor bank and Fig. 10 and Fig. 11 demonstrates the real power also.

VI. Conclusion

This study deals with the enhancement of the stability of a wind farm that is grid connected by mitigating voltage fluctuation and attaining reactive power compensation. Different approach has been adopted like adding capacitor bank, STATCOM or using STATCOM along with PID controller etc. to enhance the stability of a wind farm during wind speed change. STATCOM control strategy involves a voltage control loop with a PID controller and a PWM technique. Comparative study showed us that STATCOM with PID controller offers better performance with enhanced stability then that of using only capacitor bank.

References

[1] M. N. M. Chowdhury, S. Uddin, and S. Saleh, "Present scenario of renewable and non-renewable resources in Bangladesh: A compact analysis," International Journal of Sustainable and Green Energy, vol. 3, pp. 164-178, 2014.

[2] S. M. Salimullah, M. M. E. Rafi, and M. Sheikh, "Prospects of Wave power in Bangladesh," American Journal of Engineering Research, vol. 3, pp. 29-35, 2014.

[3] A. Saifullah, M. A. Karim, and M. R. Karim, "Wind Energy Potential in Bangladesh," American Journal of Engineering Research (AJER), e-ISSN: 2320-0847, p-ISSN: 2320-0936, 5 (7), pp. 85-94, 2016.

[4] A. Touhami, Z. S. Ahmed, and H. Samir, "Control of a wind energy conversion system equipped by a STACOM for power quality improvement," in Modelling, Identification and Control (ICMIC), 2016 8th International Conference on, 2016, pp. 115-121.

[5] Md.Monjur Ellahi Rafi, S. M. Salimullah, M.R.I. Sheikh , S.M. Showybul Islam Shakib , F.A. Bhuyan, "Prospects of Wave power in Bangladesh," in International Conference on Electrical, Computer and Telecommunication Engineering (ICECTE2012), RUET, Rajshahi-6204, Bangladesh, 2012, pp. PI-0098.

[6] M. Molinas, J. A. Suul, and T. Undeland, "Improved grid interface of induction generators for renewable energy by use of STATCOM," in Clean Electrical Power, 2007. ICCEP'07. International Conference on, 2007, pp. 215-222.

[7] R. K. Varma, V. Khadkikar, and R. Seethapathy, "Nighttime application of PV solar farm as STATCOM to regulate grid voltage," IEEE transactions on energy conversion, vol. 24, pp. 983-985, 2009.

[8] S. W. Mohod and M. V. Aware, "A STATCOM-control scheme for grid connected wind energy system for power quality improvement," IEEE systems journal, vol. 4, pp. 346-352, 2010.

[9] N. G. Hingorani and L. Gyugyi, "Static shunt compensators: SVC and STATCOM," Understanding FACTS: Concepts and Technology of Flexible AC Transmission Systems, pp. 135-207, 2013.

[10] C. Schauder and H. n. Mehta, "Vector analysis and control of advanced static VAR compensators," in IEE Proceedings C (Generation, Transmission and Distribution), 1993, pp. 299-306.

[11] M. Molinas, J. A. Suul, and T. Undeland, "Low voltage ride through of wind farms with cage generators: STATCOM versus SVC," IEEE Transactions on power electronics, vol. 23, pp. 1104-1117, 2008.

[12] H. Gaztanaga, I. Etxeberria-Otadui, D. Ocnasu, and S. Bacha, "Real-time analysis of the transient response improvement of fixed-speed wind farms by using a reduced-scale STATCOM prototype," IEEE Transactions on power systems, vol. 22, pp. 658-666, 2007.

[13] B. Messner and D. Tilbury, "Control Tutorials for MATLAB and Simulink-Introduction: PID Controller Design.[online] Ctms. engin. umich. edu," ed, 2015.

Load Scheduling of a Refrigerated Warehouse with Homogeneous Compressors Under Dynamic Pricing

Shah Md. Nehal Hasnaeen, Shohanur Rahman, and Md. Forkan Uddin*
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka - 1205, Bangladesh
nehalhasnaeen.eee14@gmail.com, shohanur.naim@gmail.com, mforkanuddin@eee.buet.ac.bd*

***Abstract*—We consider the electrical load scheduling problem for a multi-chambered refrigerated warehouse with each chamber consisting of a screw compressor operating under load-unload control. Considering loading and unloading time duration of the compressors to be fixed, we propose a simple load scheduling scheme where the loading of the compressors starts with a uniform time delay. We provide analytical models to calculate the cumulative load of the warehouse compressors at any given time and the electricity cost under dynamic pricing considering the proposed load scheduling scheme. We then validate the analytical models via simulations. As a benchmark, we consider a warehouse system where the loading time of the compressors starts randomly. We investigate that the total load under the proposed scheduling scheme do not vary significantly with time. However, the traditional system has significant variation in total load with time. We compare the electricity cost of warehouses for the proposed and the traditional systems under a dynamic pricing model. We find that the proposed load scheduling scheme significantly reduces the electricity cost compared to the traditional system. However, the cost reduction decreases with increasing the loading-unloading duty cycle, i.e., with increasing the ratio of loading time duration to the total loading-unloading time duration.**

***Index Terms*—Smart grid, demand management, real time pricing, load scheduling, refrigerated warehouse, screw compressor, load-unload control**

I. Introduction

The conventional electrical grid infrastructure is a centralized system where current flows unidirectionally through the transmission and distribution lines from the plant to the customer. Thus, the intelligence only exists in central locations and partially in substations, not at all in the terminal (i.e. loads) areas. The smart grid, as it is aptly dubbed, is a loose integration of complementary components, subsystems, functions, and services under the pervasive control of highly intelligent management-and-control systems that uses information, two-way, cyber-secure communication technologies, and computational intelligence in an integrated fashion across electricity generation, transmission, substations, distribution and consumption [1]. They enable, among others, new approaches in grid management by applying demand side management and providing grid energy storage, needed for load balancing and for overcoming energy fluctuations caused by the intrinsic nature of renewable energy sources, preventing widespread power grid cascading failures [2], and enabling the integration of plug-in hybrid and electric vehicles (PHEV and PEV) [3].

In the smart grid, two-way flow of electricity and information are supported, which lay the foundation for realizing various functions and management objectives, such as energy efficiency improvement, operation cost reduction, demand and supply balance, emission control, and utility maximization. This paves the way for the demand to be able to be matched with the supply instead of the conventional opposite scenario. One of the ways to execute this is to utilize load side management (LSM).

LSM refers to the shifting of load to accommodate the supply of the grid generation. LSM can shift peak load to off-peak periods which can reduce the cost of a consumer significantly and shifts the pressure of the load to a more convenient period [4]. In this case, the participants offer the flexibility to the operator of the demand-side management system to decrease cost of energy or to maximize their benefits of energy consumption [5]. In order to shift loads during off-peak periods appliances that act automatically, called 'smart appliances' are required [6]. With the development of smart grid, which enables reliable and real-time communications between the energy providers and the consumers, the price can be provided to the consumers daily, hourly, or in even shorter intervals [7]. One of the approaches used for load side management is load scheduling which can reduce overall load of a system as opposed to an unscheduled system, especially if a system is running in an on-off scheme [8].

Refrigerated warehouses are significant consumers of electrical energy in the food industrial sector. It has been estimated that this sector is responsible for 20% of the total electric energy consumption of the food industry [9] and 15% of all energy consumption worldwide [10]. Within cold storage facilities, 60-70% of the electrical energy may be used for refrigeration In general, compressor controls seek to maintain the compressed air discharge pressure within a specified range. With load-unload control, a compressor runs fully loaded, producing compressed air at maximum efficiency until the discharge pressure reaches the upper activation pressure setting, which causes the compressor to unload. When unloaded, the compressor no longer adds compressed air to the system, but the motor continues to run [11].

In [12], the authors propose a control scheme for all refrigeration systems in a supermarket that performs short term

adjustments to achieve overall peak reduction by using an applied control approach that performs short-term adjustments on the refrigeration power demand, while taking into account the temperature limitations of air and stored commodities. In [13], the authors propose a cooperative demand response (CDR) scheme for load management in smart grid for a refrigerated warehouse and via numerical results it is demonstrated that the CDR scheme can reduce the electricity costs and the total energy consumption in comparison with the non-cooperative demand response scheme. The approaches taken for peak reduction are either predictive and real time pricing feedback schemes, or algorithms presented for non-industrial applications. These approaches are very complex and difficult to implement. Developing a very simple load scheduling scheme for avoiding operation complexity is the main motivation of this research work.

In this paper, a multi-chambered refrigeration warehouse is considered for applying a non-predictive load scheduling scheme to reduce overall peak demand. Each chamber of the warehouse consists of one compressor per chamber operating under the same rated power. All compressors operate under the load-unload control scheme with a fixed load-unload time duration. The contributions in this paper are as follows.

- We propose a simple scheduling scheme of the compressor's electric loads of the warehouse where the initial loading of the compressors begins with a uniform delay.
- We provide an analytical model to calculate the total electric load of the compressors under the proposed load scheduling scheme at any instant of time and validate the model via simulation.
- We provide closed form expression of electricity cost under dynamic pricing considering the proposed load scheduling scheme. We also validate the model via simulation.
- The performance of proposed scheduling scheme is compared with the performance of a traditional electric load operation scheme under dynamic pricing and it is demonstrated that the proposed scheme can reduce the variation of electrical load with time and the electricity cost significantly.

The rest of the paper is organized as follows. In Section II, we review the different LSM methods that have been applied in cold storages, refrigerated warehouses and in similar cases. In Section III, we describe the refrigerated warehouse system and propose load scheduling scheme. The analytical models are developed in Section IV. In Section V, we analyse the the performance of the proposed scheme. Section V concludes the paper.

II. Related Work

In the previous studies, some algorithms and methods are proposed for LSM to specifically accommodate the refrigerated warehouse system and other relevant cases and scenarios. The Ernest Orlando Lawrence Berkeley Laboratory [14] explores various reduction techniques and permanent changes in the system to achieve demand response. Detailed energy audits are carried out by Evans and Gigiel [15] on a small number of cold stores and it is shown that energy consumption can dramatically exceed often by at least double. These audits also demonstrate that energy savings of 30-40% are achievable by optimizing the usage of the stores, repairing current equipment and by retrofitting of energy efficient equipment. In [16], the authors consider the case of a typical household user operating a set of appliances and provide a mixed integer linear programming (MILP) model and a heuristic algorithm for optimal scheduling. In [10], the authors develop a 24-hour load-forecast model to achieve dynamic system response to the variations in the ambient weather conditions. In [13], the authors use a CDR scheme for load management in case of refrigerated warehouses which also has an incentive-compatible trigger-and-punishment mechanism to avoid the non-cooperative behaviours of the selfish consumers. In [17], the authors study the application of demand response for thermal storages in various cases. There are also many studies on load scheduling. However, most of the load scheduling schemes have been implied in cases of domestic equipment which have varied rated load and often are not multiple in quantity or not even identical. Such approaches are not suitable for implementation at the industrial level, where in many cases, there are often multiple identical equipment with similar rated loads and hence, a complicated load scheduling scheme might hinder the overall production to a certain degree.

III. System Model and the Proposed Load Scheduling Scheme

A. Refrigerated Warehouse System

We consider a refrigerated warehouse of N compressors. The compressors operate in parallel to meet the aggregate cooling load requirements of the system. The rate and size of all the compressors are assumed to be the same. The set of the compressors is denoted by $\mathcal{N}$.

We assume that each compressor is controlled by a load-unload scheme. The load-unload mode of operation usually generates a distinct step type power signature, ranging between 105-115% of the rated power in the load state and 15-30% of the rated power in the unload state as shown in Fig. 1 labelled as "Actual" [18]. For simplicity, the load-unload process is considered as an on-off process with a fixed duration T. We also consider that the load time and unload time are fixed. The load and unload times are denoted by T_{on} and T_{off}, respectively and hence, $T = T_{on} + T_{off}$. The effective representation of the load-unload control model of a compressor with on-off process is exhibited in Fig. 1 labelled with "Approximation". Thus, on-off process can closely approximate the actual process. Let the power consumption levels during the load and unload periods be L_{on} and L_{off}, respectively. The duty cycle of the on-off process is defined as $d = \frac{T_{on}}{T}$.

B. Real Time Pricing Scheme

For the pricing scheme, a parabolic real time pricing approximation is used where the price of electricity consumption per

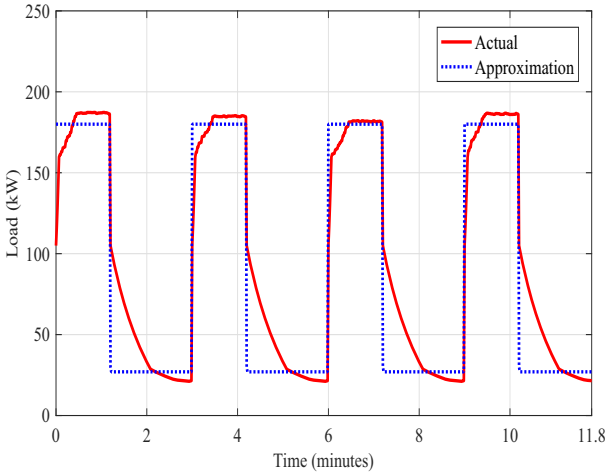


Fig. 1. Load unload operation of a 180 kW screw compressor and comparison with on-off model.

unit varied with the load by a simple parabolic expression as follows.

$$P(t) = aL^2(t) \tag{1}$$

where, $P(t)$ is the price per unit of electricity at time t, $L(t)$ is the total load of the warehouse system at time t and a is a pricing parameter. The electricity cost during a time duration T_d is given as

$$C = \int_0^{T_d} \frac{P(t)L(t)}{3600} dt. \tag{2}$$

Using (1) in (2), we obtain

$$C = \frac{a}{3600} \int_0^{T_d} L^3(t) dt. \tag{3}$$

C. Proposed Load Scheduling

We assume that all the compressors are numbered arbitrary from 1 to N and they are ordered as ascending order in set $\mathcal{N}$, i.e., $\mathcal{N} = \{1, 2, \ldots, N\}$. The proposed scheduling scheme is very simple. The initial loading time, i.e., starting time of a compressor $n \in \mathcal{N}$ is $\frac{(n-1)T}{N}$. The compressor n remains at loading state from time $\frac{(n-1)T}{N}$ to $\frac{(n-1)T}{N} + T_{on}$ and then unloads. The first unloading duration of the compressor n is from $\frac{(n-1)T}{N} + T_{on}$ to $\frac{(n-1)T}{N} + T_{on} + T_{off}$, i.e., $\frac{(n-1)T}{N} + T$. Then, the loading and unloading of the compressor n repeats at time interval T.

IV. Analytical Models

A. Analytical Model of Cumulative Load

Let Δ be the time delay between the loading times of two consecutive compressors, i.e., $\Delta = \frac{T}{N}$. Define, $m = \frac{T_{on}}{\Delta}$. The analytical model is separated into two forms based on the relation of T_{on} and Δ. One form is developed for the integer values of m and the other form is developed for the non-integer values of m.

1) The value of m is Integer: During time 0 to T_{on}, m compressors start one after another and all the compressors remain in loading state after their starting. Thus, power consumption during the time 0 to T_{on} can be written as

$$L(t) = rL_{on}; \quad (r-1)\Delta \le t < r\Delta \tag{4}$$

where, $r = 1, 2, ..., m$. At time T_{on}, i.e., $m\Delta$, two events occur simultaneously: the $(m+1)^{th}$ compressor starts operating and the first compressor drops to unload mode of operation. Since the power consumption of a compressor at unload mode is L_{off}, the increment on the total load at this point is L_{off}, i.e., $L(t)$ at this time is $mL_{on} + L_{off}$. This is the total load until the same event occurs again, i.e., the second compressor goes into unload state and the $(m+2)^{th}$ compressor starts operating at loaded state, resulting in increment of load L_{off}. The time difference between these two events is Δ. Thus, the load function in these cases can be described as,

$$L(t) = \begin{cases} mL_{on} + L_{off}; & m\Delta \le t < (m+1)\Delta \\ mL_{on} + 2L_{off}; & (m+1)\Delta \le t < (m+2)\Delta. \end{cases} \tag{5}$$

This pattern, i.e., starting of one compressor and going into unload mode of operation by another compressor lasts until the N-th compressor starts. The N-th compressor starts at time $(N-1)\Delta$. Thus, power consumption during the time $m\Delta$ to T can be written as

$$L(t) = mL_{on} + qL_{off}; \quad (m+q-1)\Delta \le t < (m+q)\Delta \tag{6}$$

where, $q = 1, 2, ..., (N-m)$.

Once all the compressors start their operation, at Δ time interval one of the compressors shifts to unload mode from load mode and one of the compressors shifts to load mode from unload mode. As a result, for $t \ge T$ the total power consumption is given as

$$L(t) = mL_{on} + (N-m)L_{off}; \quad t \ge T. \tag{7}$$

Thus, the total power consumption of a warehouse is given by

$$L(t) = \begin{cases} rL_{on}; & (r-1)\Delta \le t < r\Delta \\ mL_{on} + qL_{off}; & (m+q-1)\Delta \le t < (m+q)\Delta \\ mL_{on} + (N-m)L_{off}; & t \ge T \end{cases} \tag{8}$$

where, $r = 1, 2, \ldots, m$ and $q = 1, 2, \ldots, (N-m)$ if m is integer.

2) The value of m is Non-integer: Let $M = \lfloor m \rfloor$. If m is not integer, then the process of the load scheduling is the same as described in sub-section IV-A1 from time 0 to $M\Delta$. At time $M\Delta$, the $(M+1)$-th compressor starts, however, the first compressor does not go to unload operation mode until the time reaches to $m\Delta$. Thus, during the time $M\Delta$ to $m\Delta$, the power consumption is $(M+1)L_{on}$ and during the time $m\Delta$ to $(M+1)\Delta$, the power consumption is $ML_{on} + L_{off}$. Similar process continues until the time reaches to T. Thus, the expression of power consumption between time 0 to T can be written in (9), where $r = 1, 2, \ldots, M$ and $q = 1, 2, \ldots, (N-M)$.

$$L(t) = \begin{cases} rL_{on}; & (r-1)\Delta \le t < r\Delta \\ (M+1)L_{on} + (q-1)L_{off}; & (M+q-1)\Delta \le t < (m+q-1)\Delta \\ ML_{on} + qL_{off}; & (m+q-1)\Delta \le t < (M+q)\Delta \end{cases} \tag{9}$$

$$L(t) = \begin{cases} rL_{on}; & (r-1)\Delta \le t < r\Delta \\ (M+1)L_{on} + (q-1)L_{off}; & (M+q-1)\Delta \le t < (m+q-1)\Delta \\ ML_{on} + qL_{off}; & (m+q-1)\Delta \le t < (M+q)\Delta \\ (M+1)L_{on} + (N-M-1)L_{off}; & (N+p)\Delta \le t < (m+p+N-M)\Delta \\ ML_{on} + (N-M)L_{off}; & (m+p+N-M)\Delta \le t < (N+p+1)\Delta \end{cases} \tag{10}$$

When all compressors start to operate at time $(N-1)\Delta$, the power consumptions are $(M+1)L_{on}+(N-M-1)L_{off}$ from time $(N-1)\Delta$ to $(N-1+m-M)\Delta$ and $ML_{on}+(N-M)L_{off}$ from time $(N-1+m-M)\Delta$ to $N\Delta$ (i.e., T). This periodic power consumptions repeats at Δ time interval. Thus, the expression of power consumption for non-integer n can be written as in (10), where $r = 1, 2, \ldots, M$, $q = 1, 2, \ldots, (N-M)$ and $p = 0, 1, 2, 3....$ and so on.

B. Analytical Model of Total Cost

According to the model of the cumulative load, the total power consumption is $mL_{on} + (N-m)L_{off}$ for $t \ge T$ if m is integer. If we assume that T is negligible compared to the total system operation time, i.e., $T_d >> T$, then we can assume that the total load is always $mL_{on}+(N-m)L_{off}$ for an integer value of m. Thus, from (3), the cost can be found as

$$C = \frac{aT_d}{3600}\left(mL_{on} + (N-m)L_{off}\right)^3 \tag{11}$$

when the value of m is integer.

For a non-integer value of m, the instantaneous power consumption is periodic with interval Δ for $t \ge T$. During the interval Δ, the power consumption is $(M+1)L_{on}+(N-M-1)L_{off}$ for a duration $(m-M)\Delta$ and the power consumption is $ML_{on}+(N-M)L_{off}$ for a duration $(M+1-m)\Delta$. Thus, from (3), for a non-integer value of m, the cost can be found as

$$C = \frac{aT_d}{3600}\left[\left((M+1)L_{on} + (N-M-1)L_{off}\right)^3 (m-M) + \left(ML_{on} + (N-M)L_{off}\right)^3 (M+1-m)\right]. \tag{12}$$

V. Results

A. System Parameters

To evaluate the performance of the proposed compressor scheduling scheme, the number of compressors N is taken to be 9. The load-unload period T is considered to be 3 minutes with different duty cycle. The power consumptions of load and unload operation states are taken to be 180 kW and 27 kW, receptively. We consider two different values of pricing factor a, 0.000021 and 0.00021.

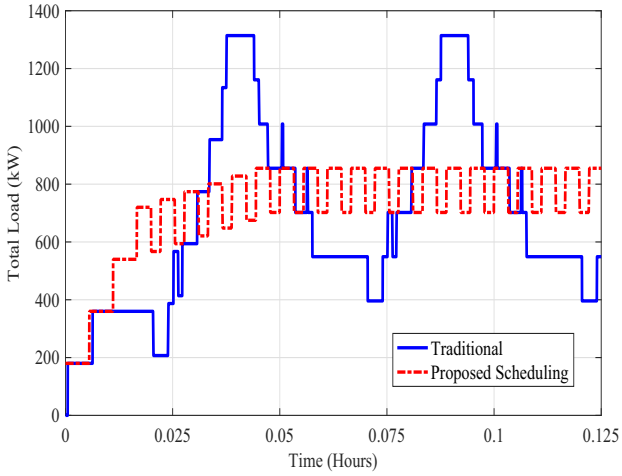


Fig. 2. Instantaneous load of warehouse with 40% duty cycle.

B. Instantaneous Load under the Proposed Load Scheduling Scheme and the Traditional Scheme

To investigate the effect of the proposed load scheduling scheme on the total load, refrigerated warehouse systems are simulated by using the tool Matlab under the proposed scheduling scheme and the traditional operation scheme. Note that in traditional operation the loading time of the compressors start randomly. The 40% duty cycle is considered as an ideal situation for demonstration of total load, price and cost. The instantaneous load of the warehouse is presented in Fig. 2. The results show that the instantaneous load of the traditional system vary significantly. However, the variation of the instantaneous load under the proposed scheduling scheme is not significant. Thus, the stability of the warehouse connected grid can be enhanced by using the proposed load scheduling scheme.

C. Validation of the Analytical Models

In order to validate the analytical model of cumulative load function, the computed values of the instantaneous loads are compared with the simulated values for the both integer and non-inter values of m comprising the values of m to be

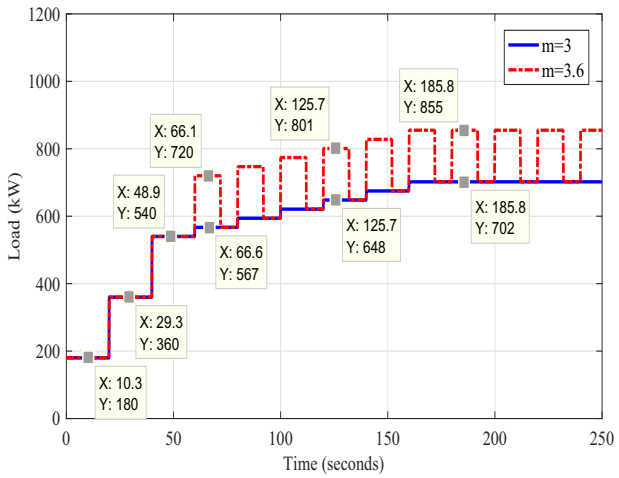


Fig. 3. Instantaneous load obtained by simulation for integer and non-integer values of m.

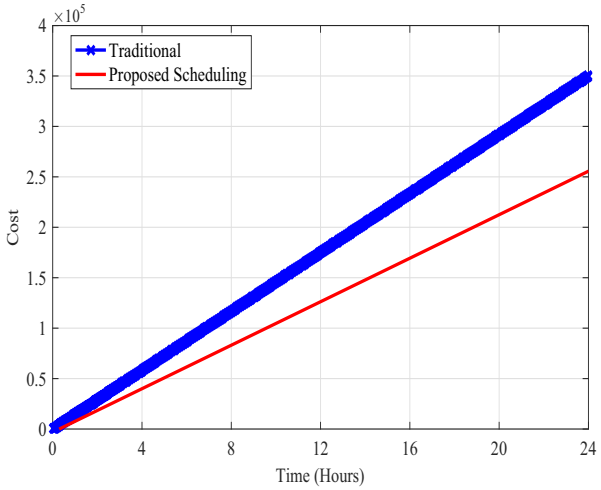


Fig. 4. Cumulative cost progression of the system operating for a day.

TABLE I
ELECTRICITY COSTS USING THE ANALYTICAL MODEL AND SIMULATION

m	Simulation	Analytical	% Difference
3	173160	174200.89	0.60
3.6	256580	258516.48	0.75

3 and 3.6. The duty cycle is adjusted accordingly. In Fig. 3, we present the simulation results and the values of the instantaneous loads are marked at the different time instants. The power consumptions values of those time instants are determined by using the analytical model and found to be the exactly same.

Considering $T_d = 24$ hours, we determine the electricity cost during time T_d using the analytical model and simulation for the two values of m. The comparison of the results is presented in Table I. The results show that the analytically calculated costs are little bit higher than the costs obtained by simulation since higher value of power consumption is considered in analytical model between time 0 to T. However, the difference is limited by 1%. Thus, the developed analytical model of cost is very accurate.

D. Electricity Costs under the Proposed Load Scheduling Scheme and the Traditional Scheme

We compute the instantaneous values of total costs under the proposed load scheduling scheme and the traditional operation scheme considering $a = 0.000021$ and d=40%. The results are shown in Fig. 4, where the thicker line represent the traditional scheme. The results show that electricity cost under the proposed load scheduling scheme is significantly lower compared to the traditional operation and the electricity cost reduction increases with time. We find that the reduction of electricity cost is 27.042% at 24 hours of operation.

E. Cost Reduction by the Proposed Scheduling Scheme

Since the tradition operation is based on random starting of the compressors, the cost reduction by the proposed scheme

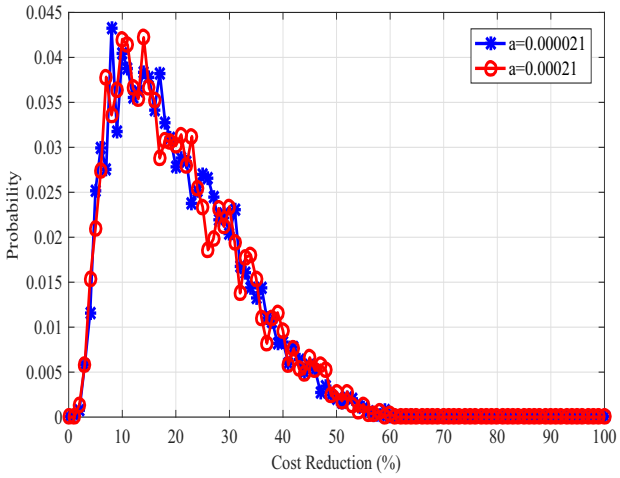


Fig. 5. Probability distribution function of cost reduction.

depends on the random starting times of the compressors. To evaluate the cost reduction statistically, for the traditional scheme, simulations are run for 15,000 times for two values of pricing factor ($a = 0.000021$ and $a = 0.00021$) with 40% duty cycle and electricity costs at 24 hours are determined. The probability distribution function (PDF) of cost reduction by the proposed load scheduling scheme with respect to the traditional operation is then determined from the 15,000 samples of cost reductions which is shown in Fig. 5. Here, only the integer values of the percentage of cost reduction are considered, i.e., PDF is similar to probability mass function (PMF). PDF of x% cost reduction is determined as $N_x/15000$, where N_x is the number of samples provides cost reduction between $(x-1)$% to x%. The results show that the percentage of cost reduction does not significantly depend on the pricing factor a since the costs are simply scaled by the pricing factor in the both systems. Probabilities of cost reduction between 10% to 30% are found to be higher.

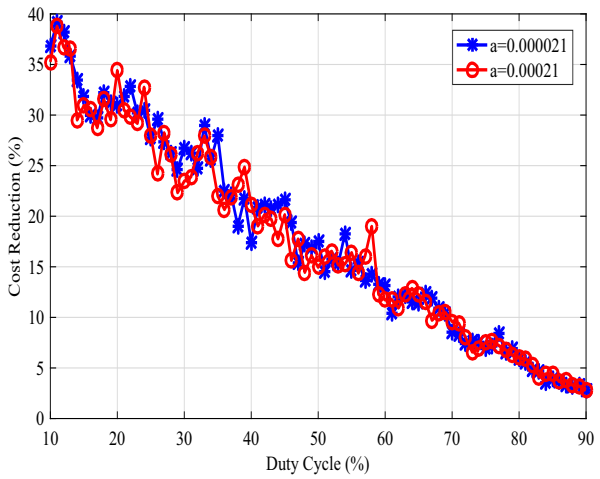


Fig. 6. Cost reduction by the proposed load scheduling scheme by varying the duty cycle.

F. Effect of Duty Cycle on Cost Reduction

To investigate the effect of duty cycle on the percentage of electricity cost reduction, the duty cycle is varied from 10% to 90%, the costs under the proposed load scheduling scheme and the traditional operation are determined. For averaging the randomness of the starting time of the compressors under the traditional operation, 40 simulations are run and averaged for each of the duty cycles. The percentage of electricity cost reductions for different duty cycles are presented in Fig. 6 for the two values of pricing factor. The results show that the effect of pricing factor on the percentage of cost reduction is negligible as shown before. We find that for an increasing duty cycle the percentage of cost reduction decreases. It is due to the fact that load time overlapping of the compressors increases with increasing the duty cycle even when the proposed load scheduling scheme is considered. As a result, the peak load difference between the traditional and the proposed schemes decreases and the percentage of cost reduction decreases if duty cycle is increased.

VI. Conclusion

We have proposed a simple load scheduling scheme for a warehouse of homogeneous refrigerators. We have provided the models for instantaneous load and electricity cost under dynamic pricing considering the proposed load scheduling scheme. We have demonstrated that the proposed load scheduling scheme reduces the peak load as well as the electricity cost significantly. However, the percentage cost reduction decreases if duty cycle of the compressors is increased. The proposed load scheduling scheme can be implemented easily due to its simplicity. This study focuses on only the evaluation of the reduction of peak load and electricity cost under the proposed load scheduling scheme. The future work related to this work is to evaluate the stability of power system under the proposed load scheduling scheme and its comparison with the stability of traditional system.

References

[1] M. Filippini and L. C. Hunt, "Energy demand and energy efficiency in the OECD countries: a stochastic demand frontier approach," *Energy Journal*, vol. 32, no. 2, pp. 59-80, 2011.

[2] V. Calderaro, C. Hadjicostis, A. Piccolo, and P. Siano, "Failure identification in smart grids based on petrinet modeling," *IEEE Trans. Ind. Electron.,* vol. 58, no. 10, pp. 4613-4623, Oct. 2011.

[3] J. A. P. Lopes, Hatziargyriou, J. Mutale, P. Djapic, and N. Jenkinset, "Integrating distributed generation into electric power systems: A review of drivers, challenges and opportunities," *Electric Power Syst. Res.,* vol. 77, no. 9, pp. 1189-1203, Jul. 2007.

[4] M. H. Albadi and E. F. El-Saadany, "A summary of demand response in electricity markets," *Electron. Power Syst. Res.*, vol. 78, no. 11, pp. 1989-1996, 2008.

[5] M. A. Khan, N. Javaid, M. Arif, S. Saud, U. Qasim and Z. A. Khan, "Peak load scheduling in smart grid communication environment," *Proc. IEEE 28th International Conference on Advanced Information Networking and Applications*, pp. 1025-1032, Victoria, BC, 2014.

[6] A. Z. Zhou, F. Zhao and J. Wang, "Agent-based electricity market simulation with demand response from commercial buildings" *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 580-588, Dec. 2011.

[7] K. Ma, G. Hu and C. J. Spanos, "Distributed energy consumption control via real-time pricing feedback in smart grid," *IEEE Transactions on Control Systems Technology,* vol. 22, no. 5, pp. 1907-1914, Sept. 2014.

[8] I. Atzeni, L. G. Ordonez, G. Scutari, D. P. Palomar, and J. R. Fonollosa, "Demand-side management via distributed energy generation and storage optimization," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 866-876, Jun. 2013.

[9] R. Paul Singh, "Benchmarking study of the refrigerated warehouse industry sector in california", *California Energy Commission Public Interest Energy Research Program*, Jul 2008.

[10] P. Rupanagunta, M. L. Baughman and J. W. Jones, "Scheduling of cool storage using non-linear programming techniques," *IEEE Transactions on Power Systems*, vol. 10, no. 3, pp. 1279-1285, Aug. 1995.

[11] C. Schmidt and K. Kissock, "Power Characteristics of Industrial Air Compressors," *Mechanical and Aerospace Engineering Faculty Publications, University of Dayton eCommons*, pp. 1-9, 2003.

[12] M. Glavan, D. Gradišar, I. Humar and D. Vrančić, "Refrigeration control algorithm for managing supermarket's overall peak power demand," *IEEE Transactions on Control Systems Technology*, July 2018.

[13] K. Ma, G. Hu and C. J. Spanos, "A cooperative demand response scheme using punishment mechanism and application to industrial refrigerated warehouses," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1520-1531, Dec. 2015.

[14] D. Scott, R. Castillo, K. Larson, B. Dobbs and D. Olsen, "Refrigerated warehouse demand response strategy guide", *Ernest Orlando Lawrence Berkeley National Laboratory*, Nov. 2015.

[15] J. A. Evans and A. J. Gigiel, "Reducing energy consumption in cold storage rooms," *IIR ICCC*, Cambridge, 2010.

[16] A. Agnetis, G. de Pascale, P. Detti and A. Vicino, "Load scheduling for household energy consumption optimization," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2364-2373, Dec. 2013.

[17] T. Samad and S. Kiliccote, "Smart grid technologies and applications for the industrial sector," *Computers & Chemical Engineering,*" vol. 47, pp. 76-84, 2012.

[18] C. Schmidt and K. Kissock, "Modelling and simulation of air compressor energy use," *ACEEE Summer Study on Energy Efficiency in Industry*, pp. 131-142, 2014.

Compensation for Voltage Disruption in a Smart Grid during Cyber–Attack

M. A. Rahman
Department of Electrical & Electronic Engineering
Pundra University of Science & Technology
Bogura, Bangladesh
Email: mdabdur.rahman.1995@ieee.org

M. S. Rana
Department of Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
Email: sohel.unsw@gmail.com

***Abstract*—Power system is turning into smart grid for ensuring efficient, reliable and economic operation. But, cyber–attack is being introduced because of involvement of cyber infrastructure in modern grid. In this paper, impacts of cyber–attack on proportional integral derivative controlled automatic voltage regulator have been analyzed. A control scheme has been proposed for compensating voltage disruption during cyber–attack. A controlled switching unit has been used in the proposed control scheme which ensures normal operation during an attack.**

***Index Terms*—Smart grid, Cyber–attack, Automatic voltage regulator (AVR), Voltage disruption, Controlled switching unit (CSU).**

I. Introduction

Power system, a complex and critical infrastructure, is responsible for generation, transmission and distribution of electrical energy. At present, the concept of smart grid (SG) is being introduced for achieving optimized use of facilities and improving quality of services of the system by using cyber infrastructure with physical infrastructure [1]. SG enhances efficiency and reliability by using information, communication technologies and computational intelligence. In SG, two–way communication for measurement and control purposes is achieved using smart devices, which leads to automated operation of the system. Utilization of information and communication technologies (ICT) in physical infrastructure has opened up new possibilities including vulnerabilities, which emphasizes on physical application security as well as cyber infrastructure security [2].

ICT introduces possibilities of cyber–attack during data transfer via communication network. An attack may cause cyber switching or data manipulation to hide physical attack using communication network [3], [4]. If the attacker has proper knowledge about the system, reduction in efficiency and increase in operational cost may be caused by false data injection (FDI) [5], [6]. Again, there is possibility of delaying data communication between sensors and controller during cyber–physical attack [7]. There is also risk of cyber–attack on bus and line protection system settings and parameters [8].

Besides these types of attacks, an attack may be launched to disrupt normal operation of the system by disturbing voltage and frequency. The impact of injecting false data for voltage violation in distributed grid system has been discussed in [9], [10]. Impacts of Denial of service cyber–attacks on voltage control schemes in SG has been discussed with countermeasures in [11]. Again, an attacker may attack distributed energy resources control for voltage violation [12], [13]. There is also possibility of attacker with on–line learning ability to find optimal strategy for attacking automatic voltage control of a power system [14].

A. Contribution of This Work

Though there are several studies regarding cyber–attack in power system, there are less concerns about attack for manipulating parameters of control unit. In our previous work [15], impacts of cyber–attack on vulnerable parameters of proportional–integral–derivative (PID) controlled AGC have been analyzed with countermeasure. In this study, cyber–attack on automatic voltage regulator (AVR) has been discussed. A controlled switching unit (CSU) has been proposed as a countermeasure which detects attacks and takes necessary steps to prevent the impacts. The major contributions of this paper are: analyzing the impacts of cyber–attack on PID controlled AVR and proposing a protective model of AVR using a CSU for compensating voltage disruption.

The reminder of this paper is organized as follows: Different models of AVR are discussed and compared briefly in Section II. Vulnerabilities and impacts of cyber–attack on PID controlled AVR are discussed in Section III. The model of CSU based AVR for a single unit of a power system is presented in Section IV. Finally, the paper is concluded in Section V.

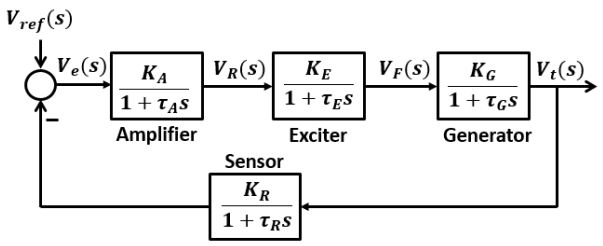


Fig. 1: Block diagram of simplified AVR [16].

II. Model of Automatic Voltage Regulator

During generation, transmission and distribution of electric energy, it is important to maintain stable terminal voltage.

Fluctuation of terminal voltage is related to fluctuation of reactive power. There are several means to maintain stable voltage operation or control reactive power such as field excitation system, transformer load–tap changers, switched capacitors, step–voltage regulators and static var control equipment. The primary way of controlling generator reactive power is using AVR.

In simple AVR (see Fig. 1) voltage magnitude is sensed and compared with reference set point. Then, the error signal is amplified, which controls the exciter field to change exciter terminal voltage. After that, the exciter terminal voltage changes current of the generator, which results in change in generated emf. In this way, the reactive power is adjusted to maintain stable voltage.

A system with following parameters is considered (based on [16]):

TABLE I: PARAMETERS OF AVR SYSTEM

Name	Gain	Time constant
Amplifier	K_A = 10	τ_A = 0.1 s
Exciter	K_E = 1	τ_E = 0.4 s
Generator	K_G =1	τ_G = 1.0 s
Sensor	K_R = 1	τ_R = 0.05 s

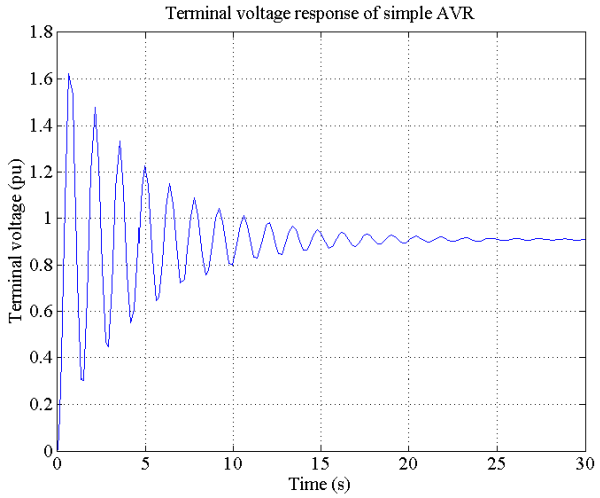


Fig. 2: Step response of simple AVR.

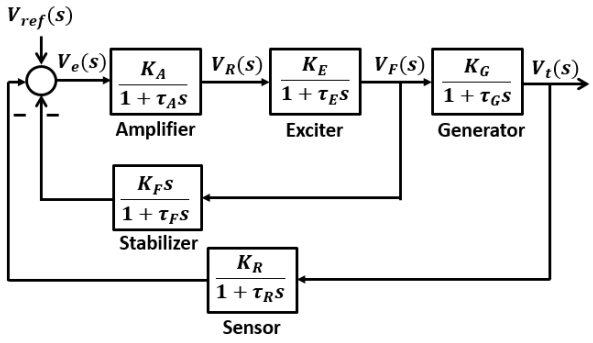


Fig. 3: Block diagram of rate feedback controlled AVR [16].

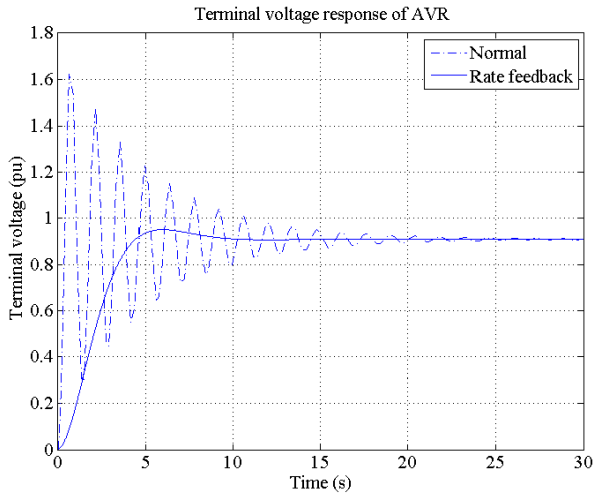


Fig. 4: Comparison of step responses of normal and rate feedback AVR systems.

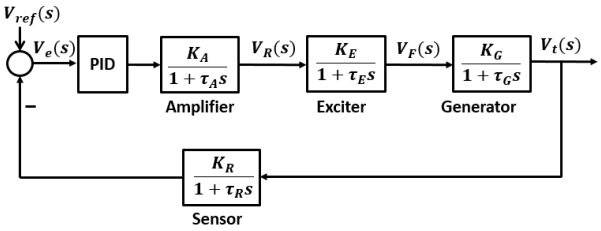


Fig. 5: Block diagram of PID controlled AVR [16].

Fig. 2 shows response of normal AVR system. The major problem in this system is that it has oscillatory response for long period. Another problem in this system is its steady–state error.

To improve the response of simple AVR system rate feed-

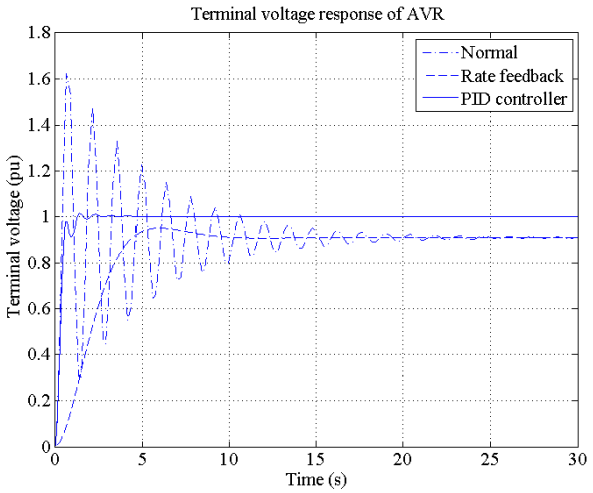
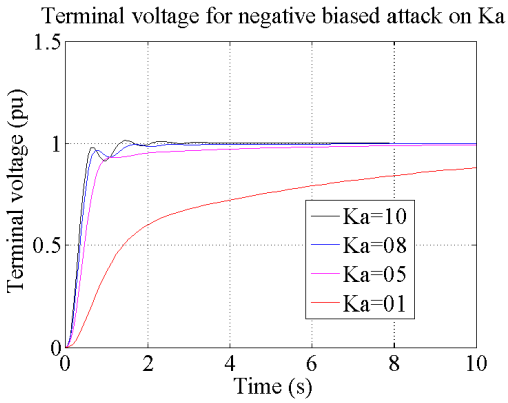
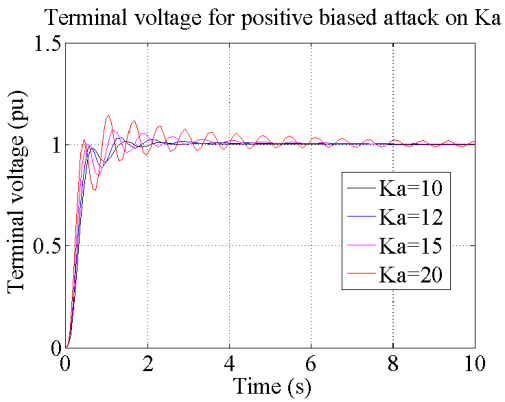


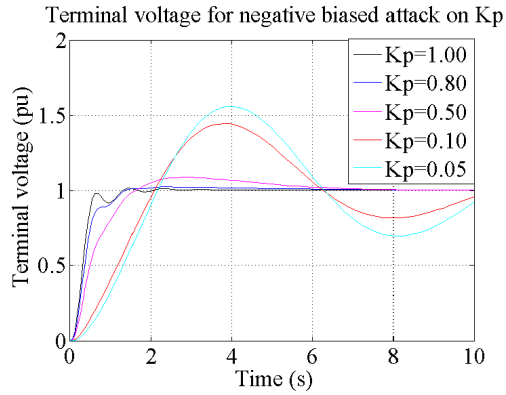
Fig. 6: Comparison of step responses of different AVR systems.



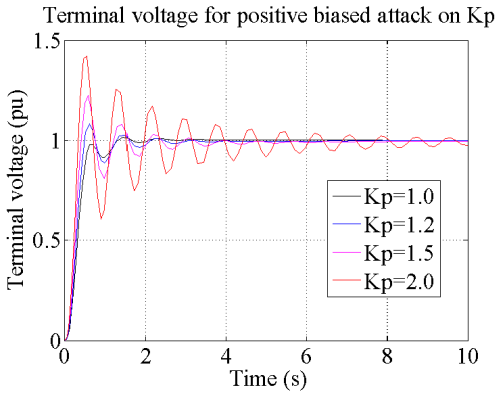
(a) Negative biased attack on K_A.



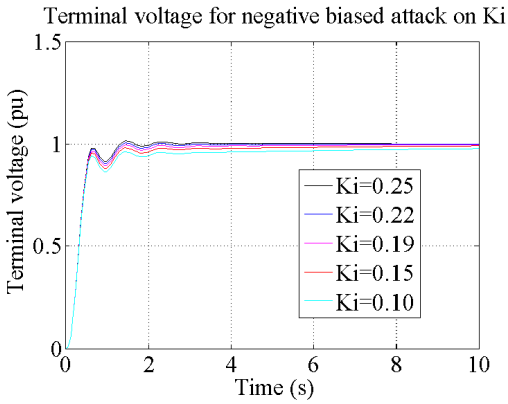
(b) Positive biased attack on K_A.



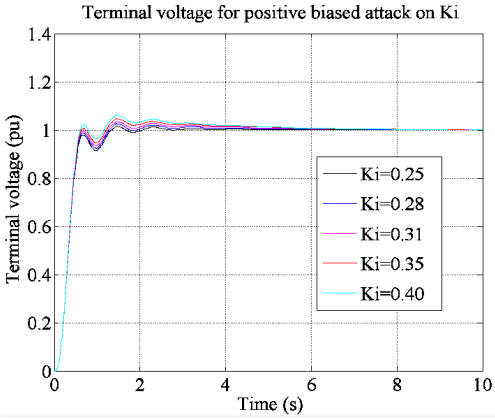
(c) Negative biased attack on K_p.



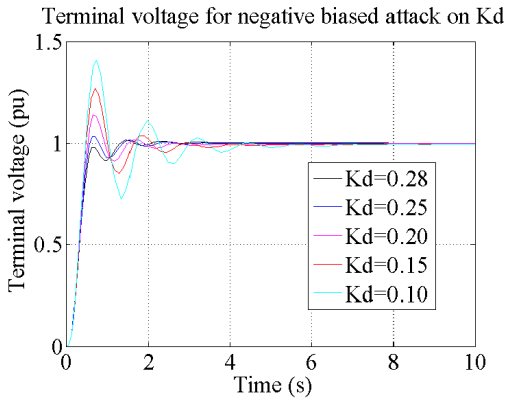
(d) Positive biased attack on K_p.



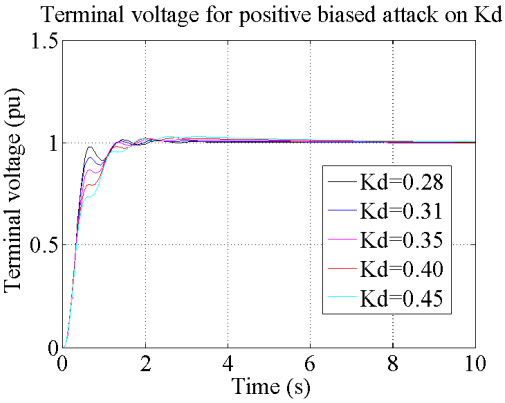
(e) Negative biased attack on K_i.



(f) Positive biased attack on K_i.



(g) Negative biased attack on K_d.



(h) Positive biased attack on K_d.

Fig. 7: Impacts of cyber–attack on PID controlled AVR.

back control can be implemented (see Fig. 3). Let's consider, gain of rate feedback controller is $K_F = 2$ and time constant $\tau_F = 0.04$. Though rate feedback control improves oscillatory response at the beginning of the response of simple AVR system, it does not eliminate steady–state error (see Fig. 4).

To solve these problems of normal AVR systems, PID controller is implemented. Three different control actions of PID controller eliminate steady–state error, reduce settling time and decrease oscillation as well as overshoot. Fig. 5 represents the block diagram of PID controlled AVR. The parameters of PID controller are $K_p = 1$, $K_i = 0.25$ and $K_d = 0.28$. Fig. 6 shows the comparison among normal, rate feedback and PID controlled AVR systems. It is observed that PID controlled AVR is proper choice for maintaining terminal voltage within permissible limit.

III. Vulnerabilities of PID controlled AVR

During cyber–attack, there is possibility that an attacker may attack an AVR system to disrupt voltage stability. Instability in terminal voltage may damage insulation of conductors as well as home appliances. So, it's important to identify the vulnerable parameters for ensuring proper protection for such an attack. In an AVR system, there are various parameters including parameters of PID controller. In a system, controller parameters are vulnerable during cyber–attack [15]. Again, the response of the AVR system is sensitive to proper choice of amplifier's gain (K_A). Hence, vulnerable parameters of a PID controlled AVR system are K_A, K_p, K_i and K_d. An attacker may change them separately or simultaneously. Again, the value of each parameter can be increased (positive biased attack) or decreased (negative biased attack) during a cyber–attack.

The impacts of cyber–attack on the vulnerable parameters are observed separately in Fig. 7. It is observed that negative biased attack has more impact than positive biased attack except for K_p, in which both types of attacks are severe.

IV. Indemnity for voltage disruption

A CSU is proposed in this paper as countermeasure against cyber–attack on AVR. It acts in two steps to ensure protection of vulnerable parameters in AVR. At first step, it monitors them to check occurrence of any disturbance. It compares current value of parameters with default data during monitoring. Default data for the system is kept in secure off–line data storage, so attacker can not cause any change in those parameters. In second step, it performs control action. If there is any attack to change those parameters, it modifies the signals using default data.

Fig. 8 represents protective model for AVR in single unit of power system. It takes two signals and two data as input. It has a feedback loop with amplifier's gain for protecting amplifier gain during an attack. The output is fed to amplifier after implementing control action, where application of control action depends on cyber–attack. Fig. 9 shows the result of implementing CSU in AVR.

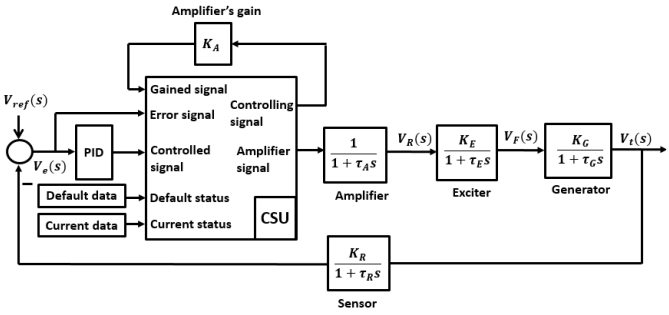


Fig. 8: CSU for AVR in single unit of power system.

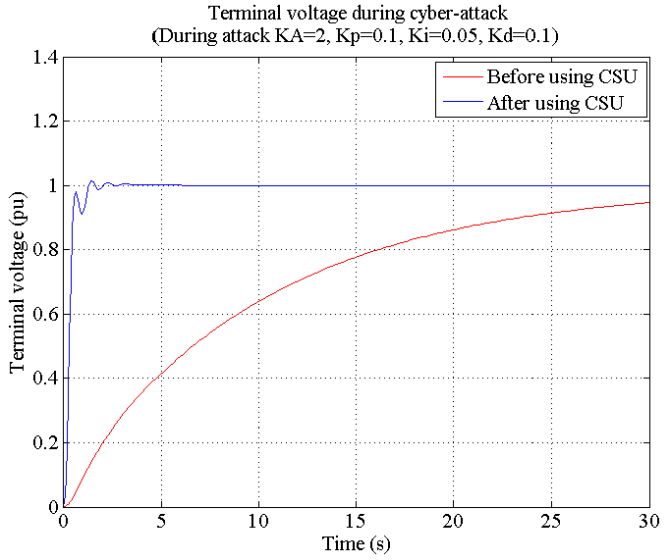


Fig. 9: Terminal voltage during cyber–attack in isolated power system (before and after using CSU).

V. Conclusion

In this work, impacts of cyber–attack on PID controlled AVR are analyzed. The values of vulnerable parameters have been changed manually during simulation for observing negative and positive biased attacks. A CSU based protective model of AVR is proposed as countermeasure for compensating voltage disturbance. The attacks, impacts of attacks and effectiveness of CSU as a countermeasure have been simulated using MATLAB Simulink. However, the proposed model is designed for a single unit of power system. In future work, protective model will be designed for interconnected system. Again, there is scope for designing a CSU for providing protection for AGC and AVR simultaneously.

References

[1] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid – The new and improved power grid: A survey", *IEEE Commun. Surveys & Tutorials*, vol. 14, no. 4, pp. 944–980, Oct. 2012.

[2] Y. Yan, Y. Qian, H. Sharif and D. Tipper, "A survey on cyber security for smart grid communications", *IEEE Commun. Surveys Tutorials*, vol. 14, no. 4, pp. 998–1010, Jan. 2012.

[3] A. K. Farraj, E. M. Hammad, A. Al Daoud and D. Kundur, "A game–theoretic control approach to mitigate cyber switching attacks in smart grid systems", in *Proc. of IEEE Int. Conf. on Smart Grid Commun.*, pp. 958–963, Nov. 2014.

[4] Z. Li, M. Shahidehpour, A. Alabdulwaheb and A. Abusorrah, "Bilevel model for analyzing coordinated cyber–physical attacks on power systems", *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2260–2272, Sep. 2016.

[5] R. Deng, G. Xiao, R. Lu, H. Liang and A. V. Vasilakos, "False data injection on state estimation in power systems – attacks, impacts and defense: a survey", *IEEE Trans. on Ind. Informat.*, vol. 13, no. 2, pp. 411–423, Apr. 2017.

[6] X. Liu, Z. Li and Y. Wen, "Cyber attacks against the economic operation of power systems: a fast solution", *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 1023–1025, Mar. 2017.

[7] A. Farraj, E. Hammad and D. Kundur, "A cyber–physical control framework for transient stability in smart grids", *IEEE Trans. Smart Grid*, Jun. 2016, DOI: 10.1109/TSG.2016.2581588.

[8] X. Liu, M. Shahidehpour, Z. Li, X. Liu, Y. Cao and Z. Li, "Power system risk assessment in cyber attacks considering the role of protection systems", *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 572–580, Mar. 2017.

[9] A. Teixeira, G. Dan, H. Sandberg, R. Berthier, R. B. Bobba and A. Valdes, "Security of smart distribution grids: Data integrity attacks on integrated volt/VAR control and countermeasures", in *Proc. of IEEE American Control Conf.*, pp. 4372–4378, Jun. 2014.

[10] Y. Isozaki, S. Yoshizawa, Y. Fujimoto, H. Ishii, I. Ono, T. Onoda and Y. Hayashi, "Detection of cyberattacks against voltage control in distribution power grids with PVs", *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1824–1835, Jul. 2016.

[11] C. Cameron, C. Patsios, P. Taylor and Z. Pourmirza, " Using Self–Organizing Architectures to Mitigate the impacts of Denial–of–Service Attacks on Voltage Control Schemes", *IEEE Trans. Smart Grid*, Apr. 2018, DOI: 10.1109/TSG.2018.2817046.

[12] P. Ju and X. Lin, "Adversarial Attacks to Distributed Voltage Control in Power Distribution Networks with DERs", in *Proc. ACM Ninth Int. Conf. Future Energy Systems*, pp. 291–302, Jun. 2018.

[13] A. Majumdar, Y. P. Agalgaonkar, B. C. Pal and R. Gottschalg, "Centralized Volt-Var Optimization Strategy Considering Malicious Attack on Distributed Energy Resources Control", *IEEE Trans. Sustainable Energy*, vol. 9, no. 1, pp. 148–156, Jan. 2018.

[14] Y. Chen, S. Huang, F. Liu, Z. Wang and X. Sun, "Evaluation of Reinforcement Learning Based False Data Injection Attack to Automatic Voltage Control", *IEEE Trans. Smart Grid*, Jan. 2018, DOI: 10.1109/TSG.2018.2790704.

[15] M. A. Rahman, M. S. Rana and M. S. Anower, "Indemnity for Frequency Disruption in a Smart Grid during Cyber–Attack", in *Proc. IEEE 2nd Int. C. on Elect. & Electron. Eng.*", Dec. 2016, DOI: 10.1109/CEEE.2017.8412868.

[16] H. Saadat, "Power System Analysis", *McGraw–Hill*, 1999.

Fuzzy Logic based Optimization for Electric Vehicle Charging Station in Bangladesh

Ashish Kumar Karmaker
Department of EEE
Dhaka University of Engineering & Technology, Gazipur, Bangladesh
ashish@duet.ac.bd

Md. Raju Ahmed
Department of EEE
Dhaka University of Engineering & Technology, Gazipur, Bangladesh
mrahmed@duet.ac.bd

***Abstract*— Electric Vehicles are opening a new field of engineering research which is very innovative and promising due to increasing consumer swiftly. As like other developed countries, Bangladesh is also experiencing with the use of battery electric vehicles like easy bike, auto rickshaw and electric bike. However, the charging stations are not sufficient and thus the electric vehicle owner recharges the batteries from the residential connection. Due to the huge electricity consumption throughout the country, the power sector has been fallen into higher system loss and failed to earn profit from this sector. In this paper, fuzzy logic based energy management algorithm is developed for electric vehicle charging station in Bangladesh. The fuzzy optimization techniques provide different charging rate for different conditions. It will save energy and makes electric vehicle business profitable to the owner.**

Keywords—Electric Vehicle, Electric Vehicle Charging Station, Fuzzy logic, Charging rate

I. INTRODUCTION

The worldwide awareness about environmental issues and reduction of fossil fuel moves the transport industry into clean and eco-friendly transportation like electric vehicles. Electric vehicles (EV) are running in three typical modes such as- Hybrid electric vehicle (HEV), Plug-in-Hybrid Electric Vehicle (PHEV), and Battery Electric Vehicle (BEV). Among of these vehicles BEV is widely used in Bangladesh which is driven by an electric motor that takes the power from battery. BEV has required no fossil fuel and provides less emission than HEVs and PHEVs [1].

In Bangladesh, mass EV penetration require large amount of power and creates a huge pressure on the power grid, distribution networks [2]. A study showed that, more than 0.5 million electric vehicles are running in Bangladesh and these ingest 450 MW electric power daily from the national grid [3]. This huge power demand creates a serious problem with the existing demand. Charging of EV requires high power and large time but their range is not satisfactory. In addition, there is no sufficient charging station in Bangladesh. Thus, the EV owner has to charge their batteries from residential connection illegally which becomes a cause of system loss in the power sector. Batteries are charged through a charger from the ac power. The charger is a non-linear equipment which hampers the power quality through production of harmonics, voltage fluctuation etc. When large number of chargers is connected with the distribution networks, the power quality problem arises [4-5]. To solve the above problems regarding EV charging, it is necessary in Bangladesh to develop sufficient charging infrastructure, their charging coordination scheme and charging policy [6].

Several researches were performed on fuzzy logic based electric vehicle charging station for controlling power demand and setting priority level for EV charging [7-9]. However, in this paper, the charging station connected with grid is considered and fuzzy logic optimization principle is applied for determining charging rate at different conditions. The power demanded by EV, time of charging (off peak/ peak hour) and duration of charging are assumed as input of mamdani fuzzy logic tool in MATLAB where charging rate is taken as output variable. Finally a comparison will be shown among fuzzy based charging station with conventional charging station.

II. ELECTRIC VEHICLE CHARGING STATION IN BANGLADESH

In Bangladesh, Electric vehicle charging stations are not sufficient. There are two types of charging stations exists- i.e. public and private charging stations. Government has established few charging stations in different cities of Bangladesh but maximum charging stations are private. These private charging stations have taken a higher charging rate.

In city areas where the electric vehicle parking is not possible for every owner, they can park it during night time & lunch time in private stations. Charging stations categorizes according to the capacity, number of charging points and charging time & cost. With the appropriate use of charging stations in Bangladesh, government can earn additional revenue and also it will increase the opportunity for decreasing unemployment. According to the technology used, charging station may be classified as follows. Such as- grid based charging station and renewable solar based charging station. The charging stations provide full charging of battery by BDT.120-150. Daily load profile for an EVCS located in Gazipur is displayed in fig. 1.

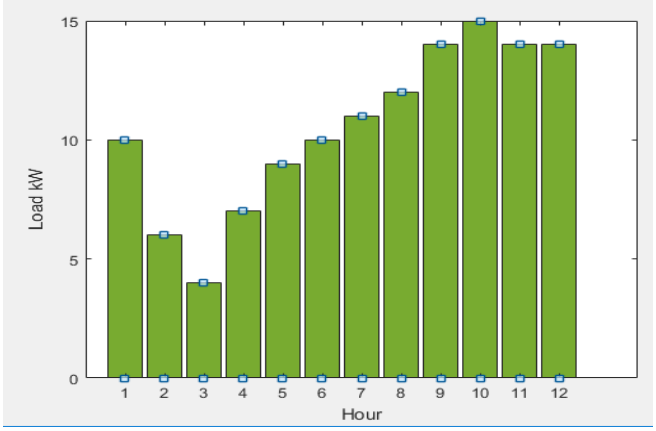


Fig.1: Daily load profile of an EVCS located in Gazipur, Bangladesh.

III. Modeling of Electric Vehicle Charging Station

Rising popularity of electric vehicles due to several positive impacts is admirable but its detrimental impacts on the grid power quality cannot be neglected. EVCS consists of the following parts basically- transformer, rectifier and converter etc. A model of EVCS is shown in fig. 2 below.

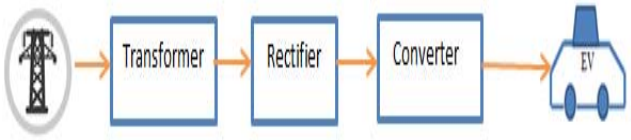


Fig. 2: Electric Vehicle connected with utility grid.

The voltage level is stepping down by the transformer for distribution purposes. Rectifier transforms AC electricity into DC and the DC-DC converter controls the DC power supplied to the EV battery. Power demanded by an EVCS can be determined as the equation (1) below.

$$P_{EVCS} = \frac{P_{EV} * N_i * F_{O.L}}{\cos\theta}; \qquad (1)$$

Where P_{EV} is the individual maximum power absorbed by an EV, N_i is the number of slot for EV connection and $F_{O.L}$ is the overload factor and $\cos\theta$ is the power factor [10]. In case of EV battery, minimum SOC has taken as 30% and maximum SOC is taken as 70%.

A. Power Demand

Electric Vehicle battery takes charge from the power distribution system. The increased power demand affects the stability of the system due to non-linearity. The power demand by an EV can be expressed as in equation (2).

$$P_{EV} = \frac{B.C*(SOC_{\max} - SOC_{\min})}{D_t - A_t}; \qquad (2)$$

Where B.C is the battery capacity, D_t and A_t are the departure and arrival time of EV respectively [11]. Battery SOC is a factor whether the EV takes high or small power. The gross power demand of the EVs is the summation of individual power demand of all EVs which likely signifies as in equation (3).

$$P_{Gross} = \sum_{N=1}^{N} P_{EV}; \qquad (3)$$

B. Duration of Charging

Higher charging time for electric vehicle battery is a great problem for EV penetration. The difference between EV arrival and departure time at EVCS is called duration of charging. In Bangladesh, level 2 AC charging is mostly used which takes large time to recharge batteries approximately 6-8 hours. The charging duration can be useful for calculating charging rate. Actually, in Bangladesh charging stations are insufficient and can provide simultaneous charging for only 30-50 EVs. Thus, maintaining charging duration is necessary. Duration of charging can be expressed as equation (4).

$$T_D = T_{departure} - T_{arrival} - T_{waiting}; \qquad (4)$$

C. Time/period of charging

Electric vehicle charging affects the power network by consuming huge energy. If the electric vehicle charging is performed on peak hour, it will surely provide a negative impact on the power system. Otherwise, electric vehicle charging at off-peak hour can be beneficial for the power system and also for the consumers. In Bangladesh, off-peak period is between 11 pm to 5 pm and peak hours are between 5 pm to 11 pm. Due to having less power demand in the off-peak hour, the charging rate will be lesser than the peak hour charging rate. In the proposed fuzzy based EVCS, the charging rate for peak hour is taken greater than the off-peak hour.

$$T_C = \begin{matrix} T_{off-peak}; & 11pm \rightarrow 5pm \\ T_{peak}; & 5pm \rightarrow 11pm \end{matrix} \qquad (5)$$

IV. Optimization Algorithm for EVCS

The main objective of the optimization algorithm is to minimize charging rate with proper utilization of electricty supply. The objective function of the proposed model is as follows:

$$Minimize\{C_{charging}\} \qquad (2)$$

This objective function works under few constraints as folows:

$$SOC_{\min} \le SOC_i \le SOC_{\max};$$

$$P_S \le P_i(t) \le P_D;$$

$$t_1 \le t \le t_2;$$

The supplied power, P_s depends only on the availability of power supply. This is also a cause of varying electrcity price in the proposed EVCS. The power demand of the EV depends on the SOC and battery capacity. In this model, the minimum SOC is taken as 20% where maximum SOC is 80%. The supplied available power should be less than or eqaul to the power demand of the EVCS. The duration of charging is the time to recharge the batteries which is shown in equation (4). The period of charging should be in between t_1 and t_2. Minimum duration of charging, t_1 is assumed as 3 hours whereas maximum charging time, t_2 is taken as 10 hours.

Power demanded by the EVs can be determined using equation (2). The overall charging cost is the function of the four input parameters such as- power demand, period and duration of charging. Thus, it can be represented by equation given below.

$$F_C = \int_{t=t_1}^{t_2} P_D(i,t) r(t) dt \qquad (5)$$

Where r(t) is the actual charging rate for an EV. Charging time varies from t_1 to t_2. If t is equal to the T_C, then r(t) will be maximum. Also, if the power generation is smaller within the proposed EVCS, the extra demand will be fulfilled by the utility grid. In that case the price will be highest. The definite integral of the r(t) is constant over the time between t_1 and t_2.

V. FUZZY LOGIC BASED EVCS

Mamdani type fuzzy inference model is used in this proposed EVCS for obtaining optimized charging rate at various conditions. Centroid based defuzzification technique is employed in this modeling. Power demanded by EVs, time of charging and duration of charging are the input parameters of fuzzy model and charging rate is the output parameter. Fig. 3 demonstrates a Mamdani fuzzy system having three input variable and an output variable.

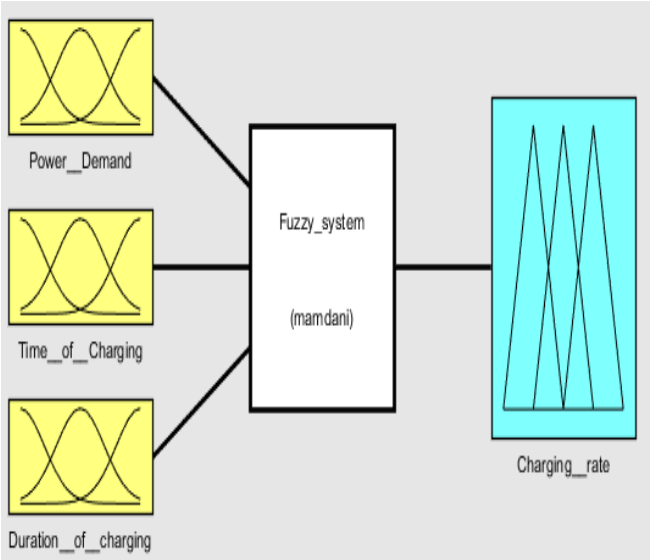


Fig.3: Fuzzy optimization model

A. Input and output parameters

Power demanded by an EV depends on the battery capacity and SOC. In this model, the membership functions of "Power_Demand" are as Very_Low, Low, Medium, High and very_High. In bangladesh, peak hour is defined as 5:00 PM to 11: 00 PM and off-peak hour is the time between 11:00 PM to 5:00 PM. For this reason, in this proposed model membership functions of the "Time_ of_ Charging"are peak_hour and off-peak_hour. Electric vehicles in Bangladesh mainly recharge their batteries in night time. But few of them are charged at day time also. Normally the easy bike and auto-rickshaw battery takes 8-10 hours for full charging. But the EVs come at charging station with above the minimum SOC, requires less time to recharge. Thus, the "Duration_ of_ charging" has membership functions as Very_short, Short, Normal, Large and Very_Large.

Charging rate is the output variable which is used in designing fuzzy system. It has membership functions declared as Very_less, less, Average, Expensive and Very _ Expensive. The input and output parameters with the membership functions are shown in fig. 4 below.

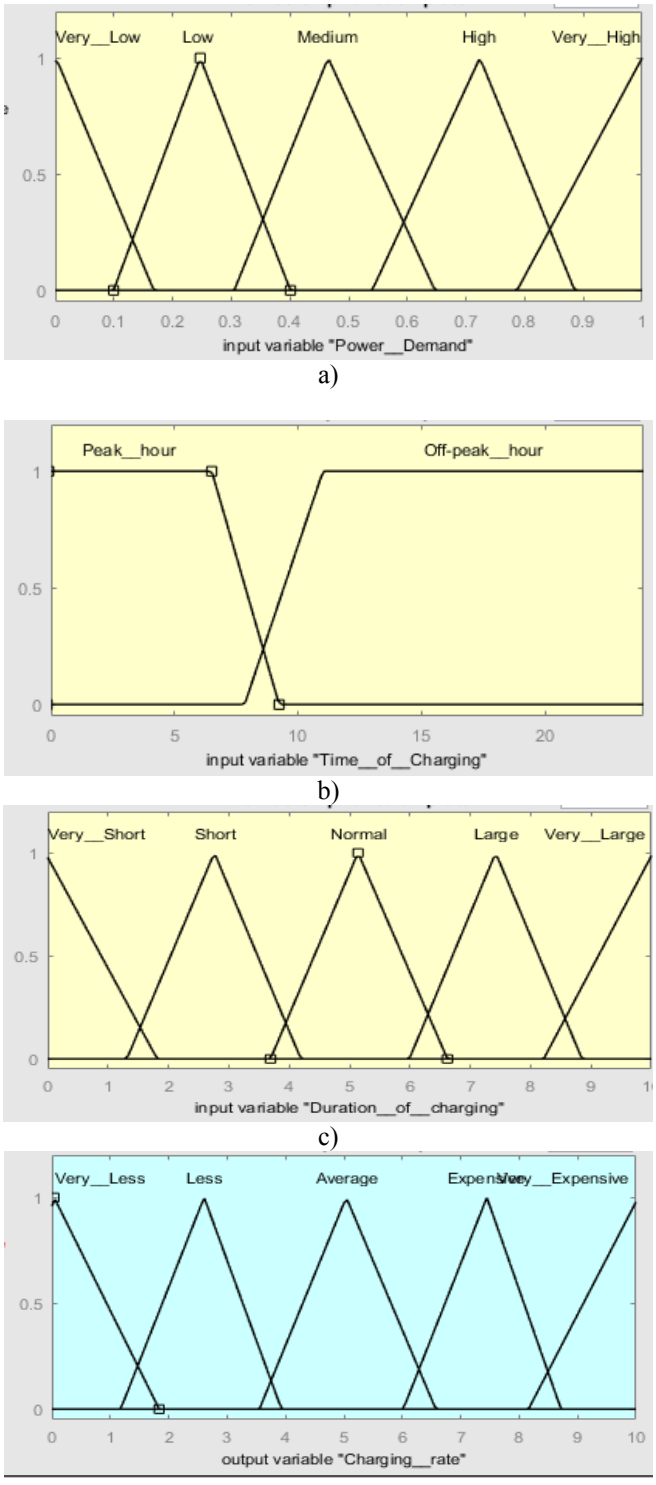


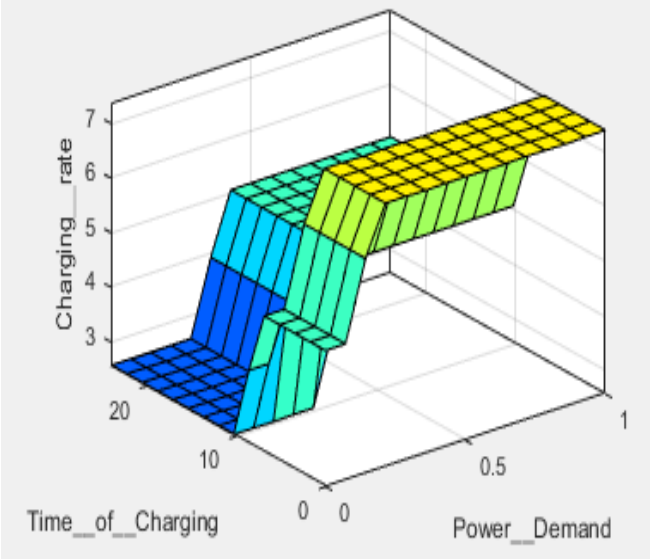
Fig. 4 a) b) c) d): Input & output parameters of Fuzzy with membership functions.

VI. RESULT AND DISCUSSION

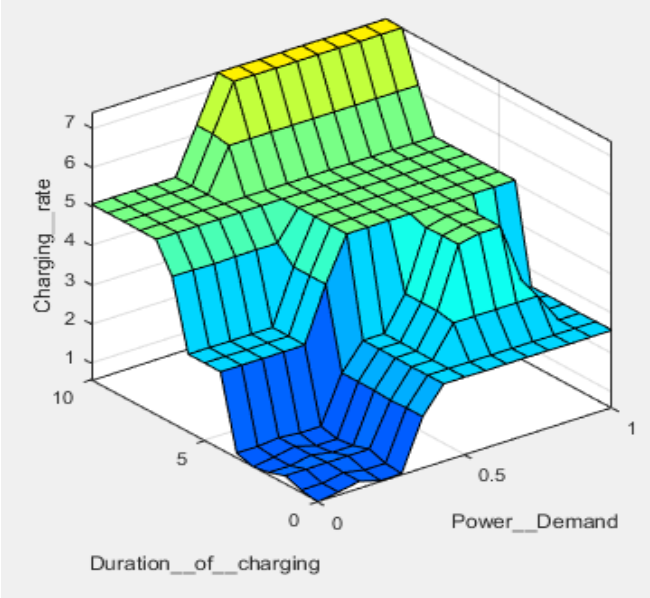
Electric Vehicle Charging Station optimization for minimizing charging cost is the main purpose of this research. In the present energy scenario of Bangladesh, charging electric vehicles leads to a huge consumption with increasing system loss. EV owners are also in a big trouble

for being less number of charging stations throughout the country. On the other hand, charging cost is very high when the electric vehicles are charged from the commercial line. Due to solve these problems with fuzzy optimization technique is employed in this research.

Fuzzy "if-then" rules based strategy is used in this optimization system. The membership functions are defined according to the data for the current battery electric vehicles of Bangladesh. The surface view of the charging rate, power demand, time of charging and duration of charging is shown in fig. 5 (a) & (b).



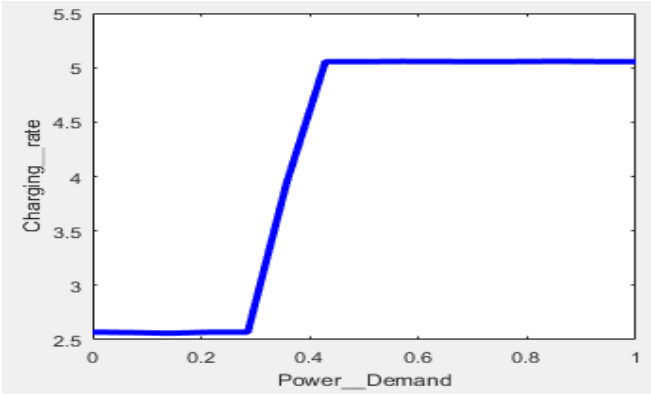
a)



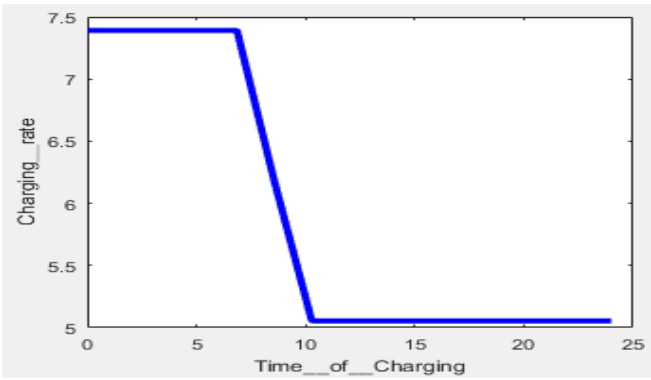
b)

Fig. 5 a) b): Surface view of charging rate, power demand time of charging and duration of charging.

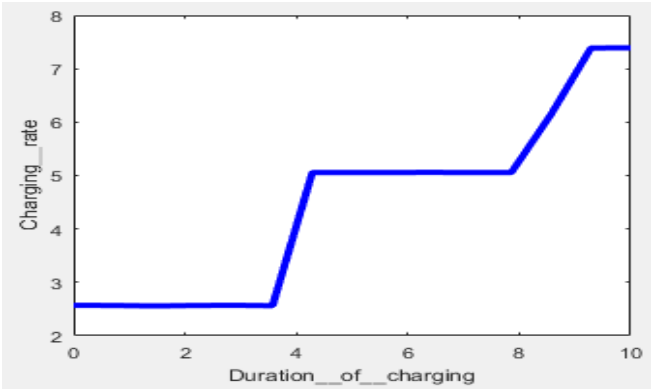
The variation of charging rate with power demand, time of charging and duration of charging are expressed in fig. 6 (a), (b) & (c).



a)



b)



c)

Fig. 6: Variation of charging rate with power demand (a), time of charging (b) and duration of charging (c)

The conventional EVCS takes on average BDT. 120 - 150 for charging an EV. In this proposed method, the charging cost is optimized which is shown in fig. 7. Conventional charging station has charging rate per kWh BDT. 7.70 whereas the fuzzy logic based EVCS offers BDT. 7.95 peak hour conditions and BDT.5.90 for off-peak hour conditions.

Using such optimization technique, EV owner can save extra cost for charging electric vehicles. It will reduce the system loss of power sector and reduce the extra burden on the utility grid by proper scheduling of EV charging coordination. This type of optimization will encourage the EV owner to recharge their batteries at scheduled period.

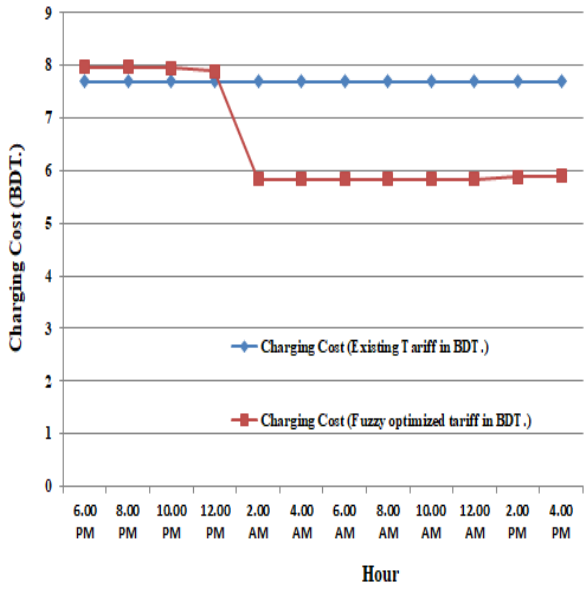


Fig. 7: Comparison of charging rate by fuzzy logic system with conventional electricity price.

VII. CONCLUSION

Electric vehicle penetration in Bangladesh is largely affected by the factors such as- lack of sufficient charging stations, high charging cost and low range. Rapid proliferation of electric vehicle consumes a large amount of electric power and the EV owner recharges their batteries from residence which leads to increase in the system loss. To minimize the charging cost with respect to power demand, charging period and charging duration, this fuzzy based EVCS model is proposed. During peak hour electricity rate is higher than the off-peak period irrespective of load connected in the charging station. However, in the proposed charging station, the EV owner can save the cost by choosing proper schedule. In case of low power demand of an EV, the charging rate is lower and higher for large power demand in the proposed charging station but in the existing system it is not possible. This type of charging station not only save the consumer income but also increase the chance of earning more revenue by choosing proper electricity charging rate.

REFERENCES

[1] Li, Wenbo, Ruyin Long, Hong Chen, and Jichao Geng. "A review of factors influencing consumer intentions to adopt battery electric vehicles." Renewable and Sustainable Energy Reviews 78 (2017): 318-328.

[2] AK Karmaker, MR Ahmed, MA Hossain, MM Sikder. "Feasibility assessment & design of hybrid renewable energy based electric vehicle charging station in Bangladesh." Sustainable Cities and Society. 2018 May 1;39:189-202.

[3] PDB for bringing battery-run vehicles under tariff regulations. Online:https://www.dhakatribune.com/bangladesh/powerenergy/2017/09/27/pdbbringingbatteryrunvehicles-tariff regulations/ (accessed on April 15, 2019).

[4] AK Karmaker, S Roy, MR Ahmed. "Analysis of the Impact of Electric Vehicle Charging Station on Power Quality Issues." In2019 International Conference on Electrical, Computer and Communication Engineering (ECCE) 2019 Feb 7 (pp. 1-6). IEEE.

[5] Durante, Larry, Matthew Nielsen, and Prasanta Ghosh. "Analysis of non-sinusoidal wave generation during electric vehicle charging and their impacts on the power system." International Journal of Process Systems Engineering 4, no. 2-3 (2017): 138-150.

[6] MR Ahmed, AK Karmaker. "Challenges for Electric Vehicle Adoption in Bangladesh." In 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE) 2019 Feb 7 (pp. 1-6). IEEE.

[7] Yao, Leehter, Zolboo Damiran, and Wei Hong Lim. "A fuzzy logic based charging scheme for electric vechicle parking station." In Environment and Electrical Engineering (EEEIC), 2016 IEEE 16th International Conference on, pp. 1-6. IEEE, 2016.

[8] Baloglu, Ulas Baran, and Yakup Demir. "A fuzzy queueing based model for controlling power demand of electric vehicle charging." (2017).

[9] García-Triviño P, Torreglosa JP, Fernández-Ramírez LM, Jurado F. "Decentralized fuzzy logic control of microgrid for electric vehicle charging station." IEEE Journal of Emerging and Selected Topics in Power Electronics. 2018 Jun; 6(2):726-37.

[10] Thomas, Polly, and Fossy Mary Chacko. "Electric vehicle integration to distribution grid ensuring quality power exchange." In Power Signals Control and Computations (EPSCICON), 2014 International Conference on, pp. 1-6. IEEE, 2014.

[11] Mohamed, Ahmed, Vahid Salehi, Tan Ma, and Osama Mohammed. "Real-time energy management algorithm for plug-in hybrid electric vehicle charging parks involving sustainable energy." IEEE Transactions on Sustainable Energy 5, no. 2 (2014): 577-586.

Modeling and Design of a Fuzzy Logic Based PID Controller for DC Motor Speed Control in Different Loading Condition for Enhanced Performance

Mirza Muntasir Nishat
Department of Electrical and Electronic Engineering
Islamic University of Technology
Dhaka, Bangladesh
mirzamuntasir@iut-dhaka.edu

Fahim Faisal
Department of Electrical and Electronic Engineering
Islamic University of Technology
Dhaka, Bangladesh
faisaleee@iut-dhaka.edu

Mahdi Rahman
Department of Electrical and Electronic Engineering
Islamic University of Technology
Dhaka, Bangladesh
mahdi@iut-dhaka.edu

Md Ashraful Hoque
Department of Electrical and Electronic Engineering
Islamic University of Technology
Dhaka, Bangladesh
mahoque@iut-dhaka.edu

***Abstract*— This paper proposes the design and simulation of a Fuzzy Logic based PID controller for DC motor speed control. The speed of a DC motor depends on armature voltage, field current and torque. In this paper, the armature voltage is changed in order to have the required speed output. The armature voltage is firstly controlled by a simple proportional controller and then controlled by PID constants and tuned by a Fuzzy Logic Controller (FLC). The FLC tunes the values of PID parameters based on the fuzzy rules fed into it and gives a satisfactory output wave shape for wide range of motor speed variation which is tested in MATLAB-Simulink environment achieving improved voltage regulation. The proposed closed loop implementation maintains constant output voltage despite changes in input voltage and significantly reduces overshoot thereby improving the efficiency of the motor. The output of this investigation has the potential to contribute in a significant way in electric vehicles, industry, communication and renewable energy sectors.**

Keywords—Dc Motor; Speed Control; Fuzzy Logic Controller (FLC); overshoot; PID

I. INTRODUCTION

DC drives are used extensively because of their simplicity, high starting torque, ease of application, high reliabilities, flexibilities and favorable cost. When it comes to power conversion, a DC-DC converter plays a significant role in widespread applications in cellular phones, laptop computers, LED Drivers, maximizing energy harvest for photovoltaic systems and for wind turbines, electric vehicles, hydro power plants and many more [1-4]. This widespread application demands that the converter should have good efficiency, less Total Harmonic Distortion (THD) and improved Power Factor (PF) at the load side. Moreover, reducing size and cost of the device and increasing availability are also some of major concerns in designing power converters for which researchers are thriving in recent years [5-9].

A high performance motor drive system must have good dynamic speed command tracking and load regulating response to carry out different tasks. In these applications, the motor should be precisely controlled to give the desired performance [10]. The regulated voltage sources used for DC motor speed control switching devices of power electronics such as MOSFET, IGBT and GTO [11]. However, the non-linearity arises due to armature current limitation, change in loads and drive inertia [12].

In addition, the parameters of the PID controller have to be fine-tuned which is difficult to model in a non-linear system. Hence to obtain better speed control, the conventional PID controllers combined with intelligent techniques such as Fuzzy logic are widely employed. Fuzzy logic controller deals with problems having uncertainties and use membership functions with values lying between 0 and 1, and uses rules to optimize the system performances [13]. This process is made particularly easy by adaptive techniques like ANFIS (Adaptive Neuro-Fuzzy Inference Systems), which are available in the Fuzzy Logic Toolbox [14]. In a DC-DC converter it is always desirable that a constant output voltage is achieved despite changes in the source voltage, the load current and variations in element values of the converter circuit [15-16]. Several control techniques have been proposed to ensure stability as well as fast transient response namely - Fuzz Logic controller [17], Artificial Neural Network (ANN) [18], PID controller [19] and PI controller [20].

In this paper voltage regulation of DC Motor is achieved employing PID controller which is tuned using trial and error method to find appropriate values for the proportional, integral and derivative gains, thereby improving converter performance. Section II of this paper deals with the modeling of conventional DC Motor followed by a brief idea about PID Controller in Section III. Section IV depicts the proposed self-tuned fuzzy logic controller and finally the simulation and results are presented by comparing the speed control with the proposed fuzzy rule based controller with PID controller in section V. The proposed circuit parameters, simulation and experimental results demonstrate the effectiveness and feasibility of the proposed scheme.

II. MATHEMATICAL MODELING OF DC MOTOR

DC Machines are of three kinds as well- Separately excited, Shunt excited & Series excited. Motor speed can be varied by controlling armature voltage, field current and torque demand.

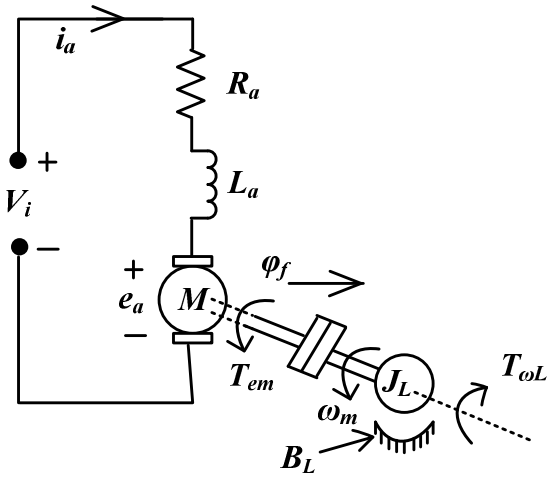


Fig. 1. Schematic Diagram of Electromechanical System

The schematic diagram of the electromechanical system is shown in Fig. 1. The corresponding parameters of the system are as follows:

V_i	Input voltage
i_a	Armature current
e_a	Back emf
R_a	Armature resistance
L_a	Armature inductance
T	Torque
J_L	Moment of inertia
B_L	Friction coefficient of motor
ω_m	Angular velocity of motor
K_b	Back emf Constant
K_m	Motor Torque Constant

The mathematical modeling of the electrical system and mechanical system are carried out by developing the following equations:

i. Electrical Modeling :

$$V_i(t) = R_a i_a(t) + L_a \frac{di_a(t)}{dt} + e_a(t) \tag{1}$$

$$\frac{di_a(t)}{dt} = \frac{1}{L_a}[V_i(t) - R_a i_a(t) - e_a(t)] \tag{2}$$

Therefore,

$$i_a(t) = \int [\frac{1}{L_a}\{V_i(t) - R_a i_a(t) - e_a(t)\}]dt \tag{3}$$

where,

$$e_a(t) = K_b \omega_m(t) \tag{4}$$

ii. Mechanical modeling :

$$T_{em}(t) = K_m i_a(t) \tag{5}$$

$$T_{em}(t) = J_L \frac{d\omega_m(t)}{dt} + B_L \omega_m(t) + T_{\omega L}(t) \tag{6}$$

$$\frac{d\omega_m(t)}{dt} = \frac{1}{J_L}[T_{em}(t) - B_L \omega_m(t) - T_{\omega L}(t)] \tag{7}$$

Therefore,

$$\omega_m(t) = \int [\frac{1}{J_L}\{T_{em}(t) - B_L \omega_m(t) - T_{\omega L}(t)\}]dt \tag{8}$$

III. Implementation of Conventional PID Controller in Simulink

One of the simplest and most widely used controllers for decades is the PID controller. PID stands for proportional (P), integral (I) and derivative (D) controller. Fig. 2 shows the block diagram of a typical PID controller.

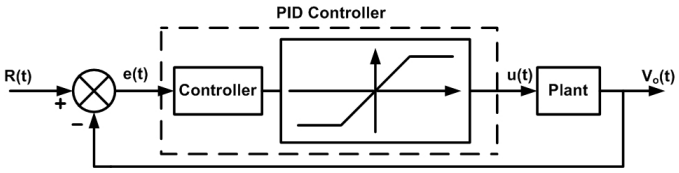


Fig. 2. Conventional PID control Structure [21]

The system under study is the plant to which necessary excitation is provided thereby achieving overall closed loop control effectively. PID controller can be expressed as –

$$C(s) = K_P + \frac{K_I}{s} + K_D s \tag{9}$$

Where, K_P, K_I and K_D are proportional, integral and derivative gains respectively. The signal e(t) as shown in Fig. 2 represents the tracking error obtained from the difference between the reference signal which serves as the input R(t) and the actual output signal $V_o(t)$. The tracking error is fed on to the PID controller which computes the derivative and integral of the signal provided. Simulink block diagram is evident in Fig. 3.

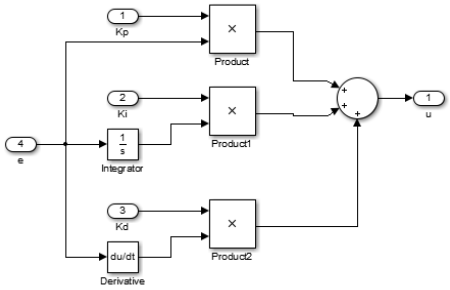


Fig. 3: Simulink Block diagram of PID Controller

Time domain representation of the signal u(t) fed to the plant is given below:

$$u(t) = K_P e(t) + K_I \int e(t)dt + K_D \frac{de(t)}{dt} \tag{10}$$

The plant on receiving the signal u(t) will generate a modified output $V_o(t)$ which will be again compared to the reference signal until the desired level is reached thereby forming a close loop system. Effect of proportional, integral and derivative control on close loop system is summarized in Table I [22].

TABLE I. EFFECT OF PID ON CLOSE LOOP SYSTEM

	Rise Time	Overshoot	Settling Time	Steady State Error
Proportional	Decrease	Increase	Small Change	Decrease
Integral	Decrease	Increase	Increase	Eliminate
Derivative	Small Change	Decrease	Decrease	Small Change

The overall system is implemented in Simulink which is evident from Fig. 4.

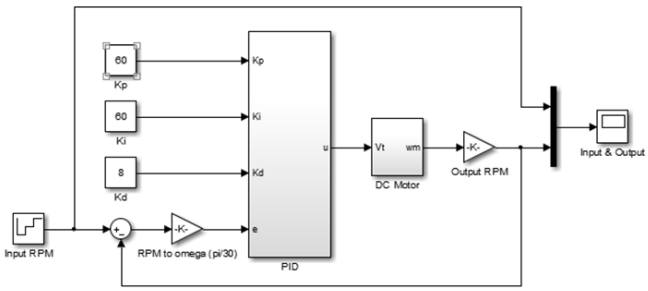


Fig. 4: Implementation of overall system in Simulink

IV. PROPOSED FUZZY LOGIC CONTROLLER

After implementing the overall system for conventional PID controller, the fuzzy controller is studied and employed in order to obtain better performance. Inputs which are fed to the Fuzzy Logic based PID Controller are speed error e(t) and change in speed error or rate de(t) which can be presented as:

$$e(t) = \omega_r(t) - \omega_a(t)$$

$$de(t) = e(t) - e(t-1)$$

PID parameters are auto-tuned using fuzzy control rules which is shown in Fig. 5. Incorporating a PID controller with the converter will improve the dynamic response and reduce the steady-state error.

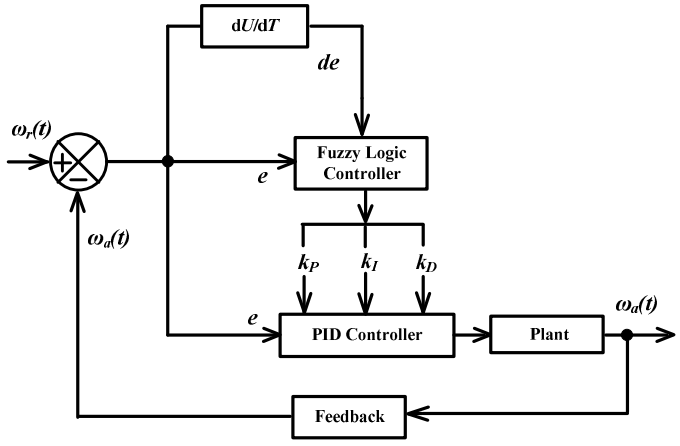


Fig. 5. Overall Block Diagram of Proposed Fuzzy Logic Based PID Controller.

The derivative controller (K_D) ameliorates the transient response and the integral controller (K_I) will reduce the steady state error of the system. The proposed system maintains an output of 200V when the input is in the range of 90V-110V which makes it quite feasible to apply in different industrial purposes.

Design of membership functions:

The parameters are expressed in the same format as they are written in the Membership Function Editor of the Fuzzy Logic Toolbox in MATLAB. All the membership functions of the input variables (speed error and change in speed error) are triangular in shape.

A. Fuzzy sets of speed error (e) variable:

Fuzzy Sets	Description	Parameters
Negative Large (NL)	Large speed difference in negative direction	[-20 -20 40]
Negative Small (NS)	Small speed difference in negative direction	[10 40 100]
Zero (ZE)	Speed difference is zero	[40 70 100]
Positive Small (PS)	Small speed difference in positive direction	[40 100 130]
Positive Large (PL)	Large speed difference in positive direction	[100 160 160]

B. Fuzzy sets of change in speed error (de) variable

Fuzzy Sets	Description	Parameters
Negative Large (NL)	Large speed difference in negative direction	[-1300 -1300 -800]
Negative Small (NS)	Small speed difference in negative direction	[-1050 -800 -300]
Zero (ZE)	Speed difference is zero	[-800 -550 -300]
Positive Small (PS)	Small speed difference in positive direction	[-800 -300 -50]
Positive Large (PL)	Large speed difference in positive direction	[-300 200 200]

C. Membership function of proportional gain K_P

Fuzzy Sets	Parameters
Positive Very Small (PVS)	[0 0 3.614 16.47]
Positive Small (PS)	[1.957 8.352 13.77 26.64]
Positive Medium Small (PMS)	[11.81 18.2 23.56 36.2]
Positive Medium (PM)	[21.8 28.2 31.8 38.2]
Positive Medium Large (PML)	[23.52 36.4 41.8 48.2]
Positive Large (PL)	[33.8 46.44 51.8 58.2]
Positive Very Large (PVL)	[43.68 56.4 60 60]

D. Membership function of integral gain K_I

Fuzzy Sets	Parameters
Positive Very Small (PVS)	[0 0 3.644 16.6]
Positive Small (PS)	[1.799 8.196 13.61 26.48]
Positive Medium Small (PMS)	[11.8 18.2 23.6 36.48]
Positive Medium (PM)	[21.8 28.2 31.8 38.2]
Positive Medium Large (PML)	[23.8 36.44 41.8 48.24]
Positive Large (PL)	[33.68 46.44 51.84 58.2]
Positive Very Large (PVL)	[43.68 56.4 60 60]

E. Membership function of derivative gain K_D

Fuzzy Sets	Parameters
Positive Very Small (PVS)	[0 0 0.09792 0.446]
Positive Small (PS)	[0.04804 0.2188 0.3632 0.7058]
Positive Medium Small (PMS)	[0.3148 0.4854 0.629 0.969]

Positive Medium (PM)	[0.5814 0.752 0.8484 1.02]
Positive Medium Large (PML)	[0.6346 0.9718 1.115 1.285]
Positive Large (PL)	[0.9012 1.238 1.381 1.552]
Positive Very Large (PVL)	[1.164 1.504 1.6 1.6]

F. Rule base for tuning K_P

de / e	*NL*	*NS*	*ZE*	*PS*	*PL*
NL	PVL	PVL	PVL	PVL	PVL
NS	PML	PML	PML	PL	PVL
ZE	PVS	PVS	PS	PMS	PMS
PS	PML	PML	PML	PL	PVL
PL	PVL	PVL	PVL	PVL	PVL

G. Rule base for tuning K_I

de / e	*NL*	*NS*	*ZE*	*PS*	*PL*
NL	PVL	PVL	PVL	PVL	PVL
NS	PML	PML	PML	PL	PVL
ZE	PVS	PVS	PS	PMS	PMS
PS	PML	PML	PML	PL	PVL
PL	PVL	PVL	PVL	PVL	PVL

H. Rule base for tuning K_D

de / e	*NL*	*NS*	*ZE*	*PS*	*PL*
NL	PVL	PVL	PVL	PVL	PVL
NS	PML	PML	PML	PL	PVL
ZE	PVS	PVS	PS	PMS	PMS
PS	PML	PML	PML	PL	PVL
PL	PVL	PVL	PVL	PVL	PVL

V. Result and Simulation

Simulations are carried out in MATLAB-Simulink environment. The parameters used for this simulation are given below:

TABLE II. Simulation Parameters

Parameter	Value
Armature resistance	0.5Ω
Armature Voltage	200 V
Armature inductance	0.02 H
Mechanical inertia	0.1 Kg.m^2
Friction Coefficient	0.008 N-m/rad/sec
Back emf constant	1.25 V/rad/sec
Rated Speed	3000 r.p.m
Motor torque constant	1 N-m/A

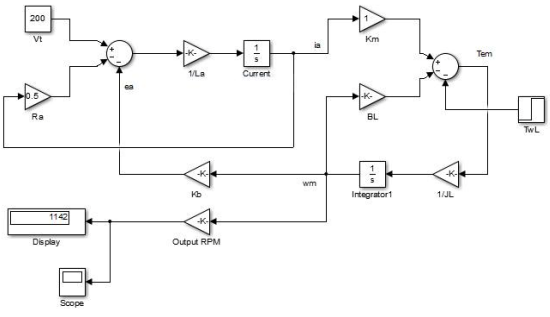
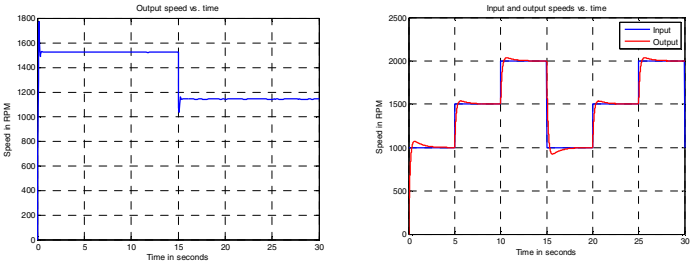


Fig. 6. Simulink Block Diagram of simple DC motor

Simulink block diagram of simple DC motor is shown in Fig. 6 and the Output speed vs. time curve is evident in Fig. 7(a). Then conventional PID controller is deployed to the system and the result is presented in Fig. 7(b).



(a) (b)

Fig. 7. (a) Speed vs. Time curve for DC motor (b) Input-Output Speed vs. Time curve for Conventional PID Controller

Initially motor runs at no-load condition. A load torque of 100 N-m is added after 15 seconds from start and remains till end of simulation. MATLAB Simulink model of Fuzzy Logic based PID controller is given in Fig. 8. Hence, speed vs. time curve is observed and shown in Fig. 9 which depicts that output follows input in a far better way with respect to conventional PID controller.

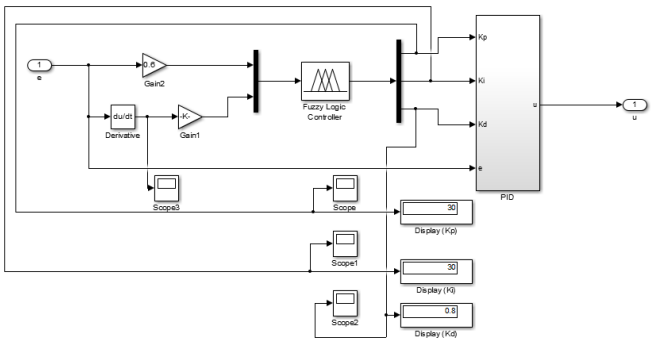


Fig. 8. MATLAB-Simulink model of fuzzy logic controller.

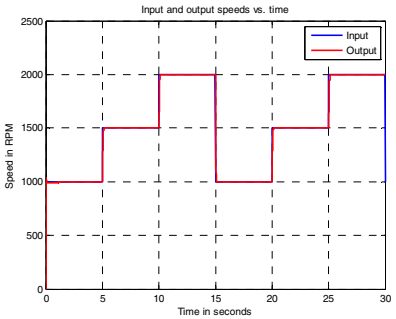
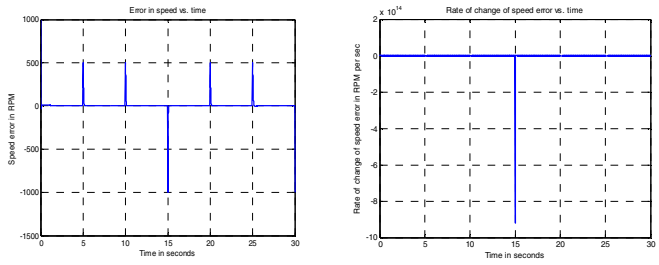


Fig. 9. Speed-time characteristics of DC motor controlled by self-tuned fuzzy PID controller.

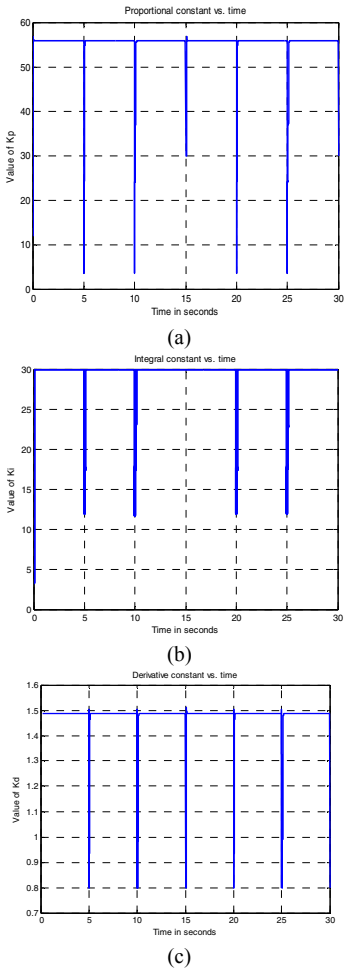
The same speeds that were fed into the controller during simulation of PID controlled DC motor as input are used here, i.e., 1000 r.p.m., 1500 r.p.m. and 2000 r.p.m. as a staircase function with sample time of 5 seconds. The simulation is run for 30 seconds, where the motor runs at no-load condition for the first 15 seconds, and at a load of 100 N-m from 15 seconds to the end of simulation. Now, error in speed vs. time, rate of change of error vs. time is shown in Fig. 10 (a) and (b). After

that, proportional, integral and derivative constant vs. time curve is sown in Fig. 11.



(a) (b)

Fig. 10. (a) Error in speed vs. time, (b) rate of change of error vs. time



(a)

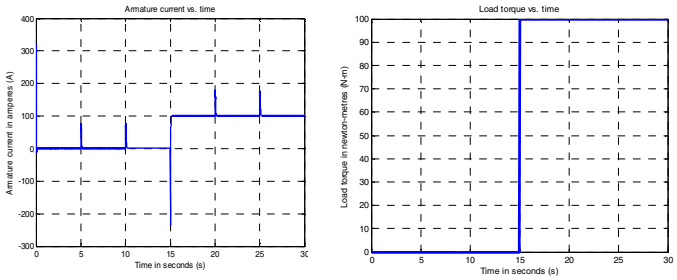
(b)

(c)

Fig. 11. (a) Proportional (b) Integral and (c) Derivative Constant vs. Time Curve

When simulation is started, high armature current flows as the motor just got started from rest. As the motor reaches steady speed, the armature current falls down to a very small value. This happens very abruptly as the rise time of the motor from rest to steady speed is extremely small. The armature current only increases at the instances when the output speed changes. This is shown by the positive spikes in Fig. 12. At 15 seconds, the speed of the motor decreases sharply.

Hence, the armature current flows in the opposite direction by a large amount. This is shown by the negative spike in Fig. 12. At the same time, a load torque is added to the motor. The current immediately increases in order for the motor to take the load, and retains a value much higher than no load condition.



(a) (b)

Fig. 12. Variation of (a) armature current (b) load torque vs. time

Rise times and fall times, Percentage maximum overshoots and undershoot, settling times and steady state errors of the output speed of DC motor controlled by proposed Fuzzy Logic based PID controller are tabulated in Table III, IV, V and VI respectively.

TABLE III. RISE TIME/FALL TIME BY SPEED VARIATION FOR UNLOADED AND LOADED CONDITION

Load condition	*Speed variation (r.p.m.)*	*Rise time/fall time (seconds)*
Unloaded	0 – 1000	0.0404
	1000 – 1500	0.0866
	1500 – 2000	0.0863
Loaded	2000 – 1000	0.0559
	1000 – 1500	0.0852
	1500 – 2000	0.0866

TABLE IV. PERCENTAGE OF MAXIMUM OVERSHOOT/UNDERSHOOT BY SPEED VARIATION FOR UNLOADED AND LOADED CONDITION

Load condition	*Speed variation (r.p.m.)*	*Percentage maximum overshoot/undershoot*
Unloaded	0 – 1000	3.55442
	1000 – 1500	0.03333
	1500 – 2000	0.07000
Loaded	2000 – 1000	0.04325
	1000 – 1500	0.07137
	1500 – 2000	0.06978

TABLE V. SETTLING TIME BY SPEED VARIATION FOR UNLOADED AND LOADED CONDITION

Load condition	*Speed variation (r.p.m.)*	*Settling time (seconds)*
Unloaded	0 – 1000	0.2307
	1000 – 1500	0.1050
	1500 – 2000	0.1008
Loaded	2000 – 1000	0.1520
	1000 – 1500	0.1069
	1500 – 2000	0.1016

TABLE VI. STEADY STATE ERROR BY SPEED VARIATION FOR UNLOADED AND LOADED CONDITION

Load condition	*Speed variation (r.p.m.)*	*Steady state error*
Unloaded	0 – 1000	0.71870
	1000 – 1500	-0.01823
	1500 – 2000	-0.03354
Loaded	2000 – 1000	0.01240
	1000 – 1500	-0.12005
	1500 - 2000	-0.11092

VI. CONCLUSION

The overall analysis illustrates that Fuzzy Logic based PID controller has very low rise time, fall time, overshoot, settling time and steady state error when employed in a DC motor under different load conditions. Both transient and steady-state responses of the system are quite satisfactory. Simulations at both unloaded and loaded conditions show the adaptability of the controller and consistency of the output response. The technique developed in this paper can be investigated for separately excited, series and shunt DC motor drives, and also different types of AC motor drives so that more robust and compact system can be developed.

REFERENCES

[1] Kumar, J. Sai, and Tikeshwar Gajpal. "A Multi Input DC-DC Converter for Renewable Energy Applications." (2016).

[2] Ortiz, G., J. Biela, D. Bortis, and J. W. Kolar. "1 Megawatt, 20 kHz, isolated, bidirectional 12kV to 1.2 kV DC-DC converter for renewable energy applications." In Power Electronics Conference (IPEC), 2010 International, pp. 3212-3219. IEEE, 2010.

[3] Li, Wuhua, Xiaodong Lv, Yan Deng, Jun Liu, and Xiangning He. "A review of non-isolated high step-up DC/DC converters in renewable energy applications." In Applied Power Electronics Conference and Exposition, 2009. APEC 2009. Twenty-Fourth Annual IEEE, pp. 364-369. IEEE, 2009.

[4] Koutroulis, Eftichios, and Kostas Kalaitzakis. "Design of a maximum power tracking system for wind-energy-conversion applications." IEEE transactions on industrial electronics 53, no. 2 (2006): 486-494.

[5] Faisal, Fahim, Mirza Muntasir Nishat, and Md Rasel Mia. "An Investigation on DC Motor Braking System by Implementing Electromagnetic Relay and Timer." In 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1-6. IEEE, 2019.

[6] Yang, Bo, Fred C. Lee, A. J. Zhang, and Guisong Huang. "LLC resonant converter for front end DC/DC conversion." In Applied Power Electronics Conference and Exposition, 2002. APEC 2002. Seventeenth Annual IEEE, vol. 2, pp. 1108-1112. IEEE, 2002.

[7] Wei, Huai, and Issa Batarseh. "Comparison of basic converter topologies for power factor correction." In Southeastcon'98. Proceedings. IEEE, pp. 348-353. IEEE, 1998.

[8] Nishat, M. M., F. Faisal and M. A. Hoque. "Modeling and Stability Analysis of a DC-DC SEPIC Converter by Employing Optimized PID Controller Using Genetic Algorithm" International Journal of Electrical & Computer Sciences IJECS-IJENS 19, no. 01(2019)

[9] Daniele, Matteo, Praveen K. Jain, and Geza Joos. "A single-stage power-factor-corrected AC/DC converter." *IEEE Transactions on Power Electronics 14*, no. 6 (1999): 1046-1055.

[10] Umesh Kumar Bansal and Rakesh Narvey, "Speed Control of DC Motor using Fuzzy PID Controller", *Advance in Electronic and Electric Engineering. ISSN 2231-1297,* Volume 3, Number 9 (2013), pp. 1209-1220.

[11] Khoei, A.. Hadidi, Kh., Microprocessor Based Closed-Loop Speed Control Systeni For DC Motor Using Power Mosfet. Electronics Circuits and Systems *IEEE International Conference ICECS'96*, Vol. 2, 1247-1250, 1996.

[12] O. Imoru, J. Tsado. "Modeling of an Electronically Commutated (Brushless DC) Motor Drives with Back-EMF Sensing", *Proceedings of the IEEE International Conference on Machine Design*, China, pp 828-831, June 2012.

[13] Md Mustafa kamal, Dr.(Mrs.)Lini Mathew, Dr. S. Chatterji, "Speed Control of Brushless DC Motor Using Fuzzy Based Controllers", *2014 IEEE Students' Conference on Electrical, Electronics and Computer Science*.

[14] Prahlad Kumar Sahoo, " Speed Control of Separately Excited DC Motor using Self Tuned Fuzzy PID Controller", 2010-2011 Department of Electrical Engineering, National Institute of Technology Rourkela.

[15] Dave, Mitulkumar R., and K. C. Dave. "Analysis of boost converter using pi control algorithms." International Journal of Engineering Trends and Technology 3.2 (2012): 71-73.

[16] M. B. Rashid, M. A. M. Oninda, F. Faisal, M. M. Nishat, G. Sarowar, and M. A. Hoque. "A Novel Topology of Single-Phase AC-DC Integrated Boost-SEPIC (IBS) Converter Using Common Part Sharing Method (CPSM) for High Step-Up Applications." Journal of Power and Energy Engineering 6, no. 06 (2018): 38.

[17] Raviraj, V. S. C., and Paresh C. Sen. "Comparative study of proportional-integral, sliding mode, and fuzzy logic controllers for power converters." IEEE Transactions on Industry Applications 33, no. 2 (1997): 518-524.

[18] Fekik, Arezki, Hakim Denoun, Ahmad Taher Azar, Mustapha Zaouia, Nabil Benyahia, Mohamed Lamine Hamida, Nacereddine Benamrouche, and Sundarapandian Vaidyanathan. "Artificial neural network for PWM rectifier direct power control and DC voltage control." In Advances in System Dynamics and Control, pp. 286-316. IGI Global, 2018.

[19] L. Mitra and N. Swain, "Closed loop control of solar powered boost converter with PID controller," *2014 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, Mumbai, 2014, pp. 1-5.

[20] M. M. Nishat, F. Faisal, M. A. M. Oninda and M. A. Hoque. "Modeling, Simulation and Performance Analysis of SEPIC Converter Using Hysteresis Current Control and PI Control Method." 2nd Int. Conf. on Innovations in Science, Engineering and Technology (ICISET), IEEE 2018, in press.

[21] V. Subramanian, and S. Manimaran. "Implementation of a sliding mode controller for single ended primary inductor converter." Journal of Power Electronics 15.1 (2015): 39-53

[22] G. J. Costa, "Tuning a PID controller." *Power Transmission Engineering* 5.2 (2011): 26-31.

A Systematic Approach to Design Optimal Controller of Single Phase VSC for Distributed and Renewable Energy Application

Md Khurshedul Islam
Electrical and Computer Engineering
Mississippi State University
Starkville, Mississippi, USA
khurshedeeeiiuc@gmail.com

Farhina Hoque
Electrical and Computer Engineering
Mississippi State University
Starkville, Mississippi, USA
farhina_meghna@yahoo.com

Naeemul Islam
Electrical and Electronic Engineering
International Islamic University Chittagong
Kumira, Chittagong, Bangladesh
naymnaeem08@yahoo.com

***Abstract*— This paper presents a systematic approach to design an optimal controller by using linear quadratic regulation (LQR) technique for single phase Full Bridge Voltage Source Converter (FB-VSC) for distributed and Renewable energy application. Traditionally, a Proportional (P) or Proportional-Integral (PI) controller is commonly used. These controllers are not robust enough to deal the converter and system uncertainties (For example, changing converter or filter parameter and disturbance in the input or output terminal of the converter). This paper presents 1) a systematic design method for both conventional PI and proposed optimal controller based on state space approach; and 2) a comparative study between these two controllers. Equivalent MATLAB Simulink simulation model for both controllers are developed to investigate their performance at various possible uncertainties. The proposed optimal controller exhibits more robustness and flexibility compared to the PI controller.**

Keywords*— *Voltage Source Converter (VSC), Linear Quadratic Regulation (LQR), PI controller, Optimal Controller, Distributed energy, and Renewable energy.

I. INTRODUCTION

The demand of electrical energy production is increasing day by day with the increasing population and it is difficult to meet this increasing demand through the existing means of power generation. Furthermore, most of our existing power generation is depleting energy source and causes environmental degradation [1]. To meet this rising power demand and to prevent the environmental degradation, expanded distributed and renewable energy resources need to be integrated into the existing utility grid system [2].

With the wide range of technological development, the installed capacity of distributed generation (DG) is increasing at a very faster rate [3, 4]. The DG technology includes solar photovoltaic (PV), wind turbines, fuel cells and battery energy storage (BES). Among these DG technologies, solar PV is becoming progressively more attractive and suitable in many countries because of the environmental and regional benefits [5]. Large- and small-scale DG units are generally connected to the utility grid at sub-transmission voltage level and at distribution voltage level respectively through a common DC bus [6, 7]. To regulate the common DC bus voltage properly a VSC is used between the DG unit and the common DC bus voltage [8]. This commonly used converter system is known as two-stage topology [9]. The first stage converter performs the power extraction and regulates the common DC bus voltage to provide a suitable voltage level for the second stage, whereas the second stage converter generates and injects stable ac power to the utility grid at a low harmonic level [10]. In the first stage of VSC converter, a close loop P or PI controller-based voltage or current controlling technique is widely used to regulate the power and voltage properly for the second stage [11]. The existing PI controller designs depend on the exact system parameters and assume that the system parameters are constant over the whole operation range and they do not perform satisfactorily against system uncertainties and disturbances.

Due to the load and temperature variation, the unmodeled dynamics such as resistance of the converter and filter elements can be changed over the time and usage, which is known as the converter parameter uncertainty. Intermittent nature of renewable energy resources (for instance, a sudden change of sunlight in solar PV and wind speed in wind turbine) can initiate a rapid disconnection of DG units. Furthermore, any unwanted fault in the utility grid can affect the dc bus voltage. These kinds of uncertainties can change the input or output voltage level of first stage converter, which is called as system uncertainty. To obtain benefits from the integration of wide-scale DG units to existing utility grids, it is expected that the first stage VSC will be robust enough to handle these converter and system uncertainties [12]. However, the conventionally used PI controller is very sensitive to the slight change of system parameters.

The purpose of this paper is to develop an efficient and systematic method to design the optimal controller and its gain to address the aforementioned controller issues. To this end, this paper first presents systematic design method for PI controller, then the optimal controller by using state-space approach. At the end, this paper shows the difference between the two controller's performance at various uncertainties.

II. CONTROL MODEL OF VSC

Voltage Source Converter (VSC) is an essential part of a variety of power electronic system and application including motor drive, integration of distributed, and renewable energy resources. Fig.1 shows a single-phase pulse-width modulation (PWM) full-bridge voltage-source converter (FB-VSC) with DC-side voltage of V_{dc}. Its output is connected to another DC supply of V_g through an inductor L. The inherent resistors of the inductor, of the line, and of the conducting-state of switches are aggregated and denoted

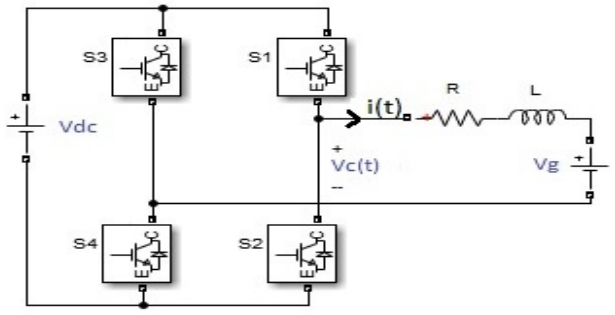


Fig. 1. Single phase FB-VSC.

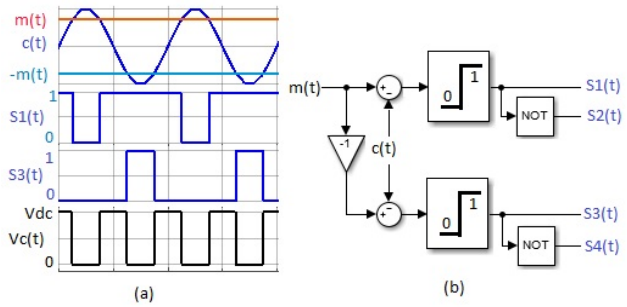


Fig. 2. Unipolar PWM for single phase FB-VSC

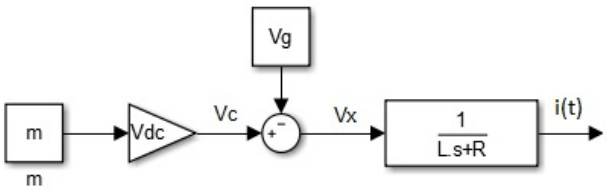


Fig. 3. Control-Average model of FB-VSC

By *R*. $V_c(t)$ is the converter output voltage and *i(t)* denotes the converter output current as well as line current. The differential equation of this FB-VSC can be expressed as

$$L\frac{di(t)}{dt} + Ri(t) = V_c(t) - V_g \qquad (1)$$

A unipolar PWM scheme is shown in Fig. 2, where *m(t)* and *c(t)* are modulation and carrier signal respectively produce the gating signals (S_1, S_2, S_3, S_4) for the IGBT's. The converter output voltage $V_c(t)$ is generated by the gate signals can be expressed as

$$V_c(t) = V_{dc}[S_1(t) - S_3(t)]$$

If switching duty cycle is D, then $S_1 = D$ and $S_3 = 1\text{-}D$

$$\bar{V}_c = V_{dc}(2D - 1) = mV_{dc} \qquad (2)$$

(2D-1) = m is Modulation index. From equation (1)

$$LsI(s) + RI(s) = V_c(s) - V_g \text{ [s is Laplace operator]}$$

$$I(s)(Ls + R) = V_x(s)$$

$$\text{Where, } V_x(s) = V_c(s) - V_g \qquad (3)$$

The open loop transfer function of the VSC converter is

$$G(s) = \frac{I(s)}{V_x(s)} = \frac{1}{Ls+R} \qquad (4)$$

From equation (2), (3), and (4), the control /average model of FB-VSC can be derived as shown in fig. 3.

III. Conventional PI Controller for FB-VSC

The open loop VSC control model is highly sensetive to a very small change of converter's input (V_{dc}) and output (V_g) terminal voltage, converter paramenters, and external disturbance. This characteristic is not desirbale at all for any kind of sensetive system for example, HVDC grid system, motor drive, PV connected grid system etc. Conventionally, to make this open loop control system robust, a close loop feedback control system with a P or PI controller is widely used [13]. Fig. 4, shows the simplified diagram of close loop feedback control FB-VSC with a PI controller. A current limiter is used to limit the reference current during grid low voltage transient fault. To solve the starting problem of VSC, the soft starter technique is used. The open loop transfer function of the PI controller is

$$G_{PI}(s) = \frac{K(s+z)}{s}$$

Here, z is the position of zero. $z = \frac{R}{L}$ (5)

Fig. 5 shows the close loop transfer function T(s) of the control system. The Close loop transfer function equation can be derived as:

$$T(s) = \frac{\frac{KV_{dc}}{Ls+R}}{1 + \frac{KV_{dc}}{Ls+R}} = \frac{KV_{dc}}{Ls + (R + KV_{dc})}$$

Close loop pole is $-\frac{R+KV_{dc}}{L}$. Assume, time constant = τ .

$$\tau = \frac{L}{R + KV_{dc}}$$

$$K = \frac{1}{V_{dc}}\left(\frac{L}{\tau} - R\right) \qquad (6)$$

For a proper system response, the maximum speed or the time constant of the system must be at least several times larger than the switching frequency (f_s). i.e $\tau > [\beta (1/f_s)]$, β should be equal or greater than 5 (Rule of thumb). This condition can be summarized as:

$$K < \frac{1}{V_{dc}\{(\beta / f_s L) - (1/R)\}}$$

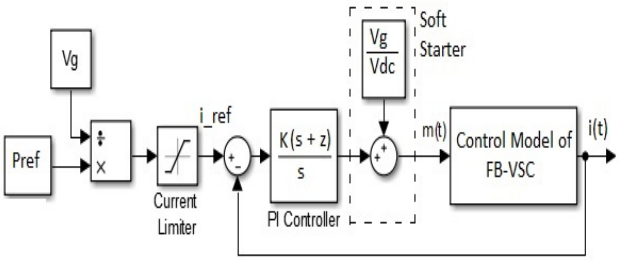


Fig. 4. FB-VSC with a PI controller.

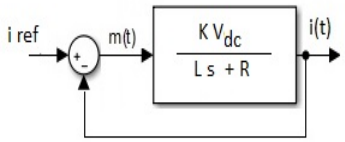


Fig. 5. Close loop Transfer function of FB-VSC with PI controller.

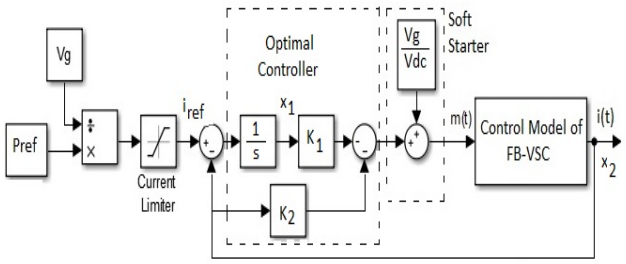


Fig. 6. FB-VSC with an optimal controller.

IV. Proposed Optimal Controller for FB-VSC

The optimal controller is designed by using state space full state feedback control approach. Fig. 6 shows the simplified diagram of close loop feedback control FB-VSC with an optimal controller.

A. State space model of proposed control structure

From fig. 6, two state variables of the control system are $x_1 = \int e$ (e=tracking error signal) and $x_2 = i(t)$ = converter output current.

$$x_2 = \frac{u V_{dc}}{Ls + R} \quad \text{or} \quad Lsx_2 + Rx_2 = uV_{dc}$$

[Where, feedback control signal $u = -k_1 x_1 - k_2 x_2$]

Derivatives of state variable $x_1 and\ x_2$ / (state equations of the system)

$$\dot{x_1} = e = i_{ref} - x_2 \tag{7}$$

$$\dot{x_2} = -\frac{R}{L}x_2 + \frac{uV_{dc}}{L} \tag{8}$$

To remove external input signals from the state equations and to convert the tracking problem to a regulation problem, apply $\Delta = \frac{d}{dt}$ transformation to both sides of equation (7 and 8) and then map **x to z=Δ(x).**

$$z_1 = \dot{x_1} \quad \text{and} \quad z_2 = \dot{x_2}$$

The control signal $v = \dot{u} = -k_1\dot{x_1} - k_2\dot{x_2} = -k_1\dot{z_1} - k_2\dot{z_2}$

The new state representation becomes

$$\dot{\boldsymbol{z}} = \boldsymbol{Az} + \boldsymbol{Bv} \tag{8}$$

Where, z_1 and z_2 are state variables, A is system matrix, B is input matrix, and v is the control signal. The New sate equations can be derived as: (assuming I_{ref}= constant)

$$\dot{z_1} = -\dot{x_2} = -z_2 \tag{9}$$

$$\dot{z_2} == -\frac{R}{L}z_2 + \frac{vV_{dc}}{L} \tag{10}$$

From equation (9 and 10), system matrix **A** and input matrix **B** can be expressed as:

$$A = \begin{pmatrix} 0 & -1 \\ 0 & -R/L \end{pmatrix} \quad B = \begin{pmatrix} 0 \\ V_{dc}/L \end{pmatrix}$$

Now LQR technique is applied to equation (8) to obtain the optimal controller's gains (k_1 and k_2).

B. Linear Quadratic Regulator (LQR) technique

LQR is a very useful and well-known concept in the optimal control theory. If a linear system has the state-space equation $\dot{\boldsymbol{z}} = \boldsymbol{Az} + \boldsymbol{Bv}$**,** then LQR is able to provide the full state feedback gains **k** in the control law $v = -kz$ which can minimize the following quadratic cost function:

$$J(v) = \int_0^\infty z^T Q z + v^T R v \tag{10}$$

In the cost function, v is the control signal, z is the system's space vector, Q is a positive semi-definite matrix, R is a positive definite matrix, and T stands for vector transposition [14]. The solution to finite-time LQR is obtained from a Differential Riccati Equation (DRE). The solution of the equation (6) is $K = R^{-1}B^T F$ can be obtained from the algebraic Riccati equation [13].

$$A^T F + FA + Q - FBR^{-1}B^T F = 0$$

Computational tools like MATLAB can be used to find the solution. These gains can provide a phase margin of greater than 60° to the close loop, which can make a system relatively good robustness to uncertainty and unmodeled dynamics.

For this FB-VSC system, **Q, R,** and **z** can be selected as:

$$Q = \begin{pmatrix} q1 & 0 \\ 0 & q2 \end{pmatrix} \quad R = 1 \quad Z = \begin{pmatrix} e \\ i \end{pmatrix}$$

Then the quadratic cost function can be written as:

$$J = \int_0^\infty (q_1 e^2 + q_2 i^2 + v^2)dt$$

By selecting q_1 and q_2 directly impact the tracking error (e) and the derivative of output current i(t). MATLAB control toolbox can be used to design Q, increase q_1 from an initial positive value and set all other coefficient to zero. Once q_1 reaches a definite desirable value representing a sufficient distance of the dominant pole from the imaginary axis, fix it to that value and then increase q_2 by same process. Bode diagram, system response, and close loop poles locations are observed during this process.

V. Simulation and Results

In this section, simulation of FB-VSC with PI controller and optimal controller are presented. Numerical Values of Table I are selected for the inverter with a simple L output filter. Fig. 7 shows the circuit diagram of FB-VSC. For both control system, same arrangement will be used, only different controller will be implemented in the controller block (Shown in the fig. 7).

TABLE I
SYSTEM PARAMETERS

Dc Link voltage at input side (V_{dc})	600 V
DC voltage at output side (V_g)	400 V
L Filter parameter (L)	700 uH
Total resistance (R)	0.005 ohm
Switching frequency (f_s)	10 KHz
Inverter Power rating (S)	640 KVA
Time constant (t)	5 ms

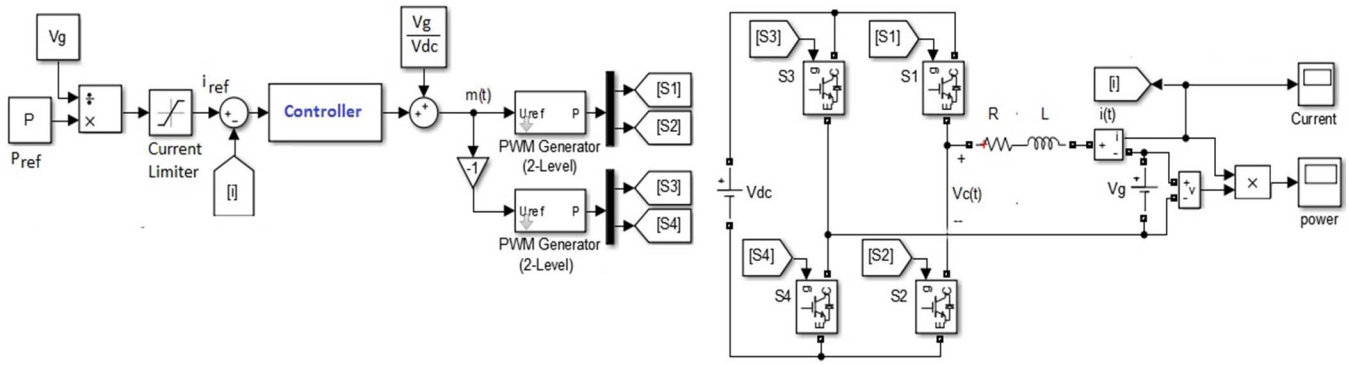


Fig. 7. Circuit arrangement of FB-VSC.

i) Open loop operation:

For a given power P,
The converter output current, $I = P/V_g = 640KW/400V = 1600A$.
The converter Voltage, $V_c = V_g + RI = (400 + 0.005*1600) = 408V$
The modulating signal, $m = V_c/V_{dc} = 408/600 = 0.68$.
And the duty cycle, $D = (m+1)/2 = 0.84$.

Simulation scenario (1):

Gating signal is applied at t=0 s, V_g jumps up 4% at t=1s, and V_{dc} jumps up 1% at t=2s.

Simulation scenario (2):

The value R is increased by 50%, 75%.

Fig. 8 & 9 show the robustness of the controller to the system and converter uncertainties. If V_g changes slightly, the direction of power flow changes. If V_{dc} and R change a little, inverter current changes by a large amount. And the system takes about 0.5 s to settle.

ii) FB-VSC with feedback loop using PI controller:

A PI controller (Designed in section III) is applied in the FB-VSC simulation circuit shown in fig. 7. Assume time constant τ =5ms. Then the value of z and k is calculated from equation (5) & (6) respectively.

$$z = \frac{R}{L} = \frac{0.005}{0.0007} = 7.143$$

$$K = \frac{1}{V_{dc}}\left(\frac{L}{\tau} - R\right) = \frac{1}{600}\left(\frac{0.0007}{0.005} - 0.005\right) = 0.000225$$

Simulation scenario (1):

i) Gating signal is applied at t=0 s. V_g jumps up 3% at t=0.05s, and V_{dc} jumps up 1% at t=0.07s.
ii) The value R is increased by 50% and100%.

Simulation scenario (2):

i) Reference power changes (P) from 640 KW to 600 KW at sec 0.05 s.

Fig.10 shows that if V_{dc}, V_g, and R change even by a very small value, the converter current (I) changes significantly. From Fig.11 it is also observed that when the reference power (P) changes from 640KW to 600KW at t = 0.05s, it takes approximately 0.02 s to track this slight change. From these observations, it is obvious that the PI controller is not robust enough to eleminate system uncertenities and disturbnce.

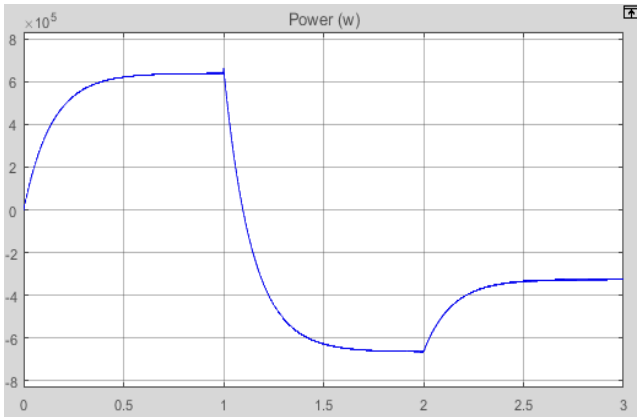


Fig. 8. Output Power of open loop FB-VSC.

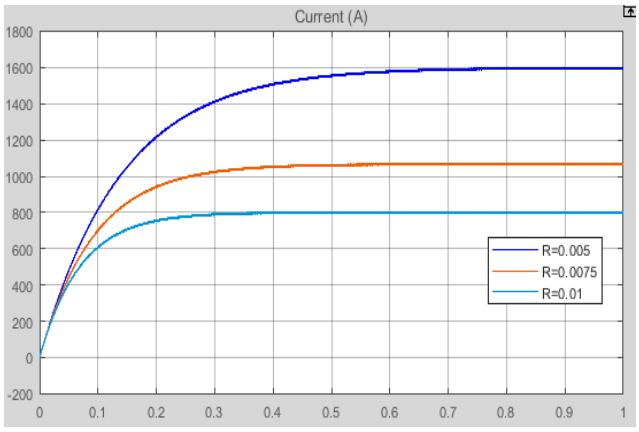


Fig. 9. Output Current of open loop FB-VSC.

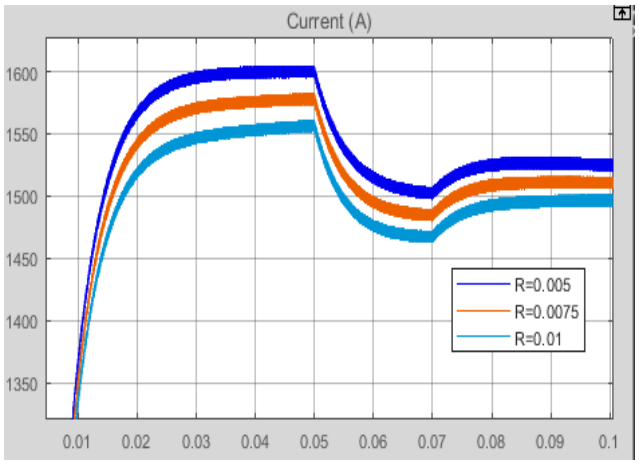


Fig. 10 Output Current of FB-VSC with PI controller (Zoomed)

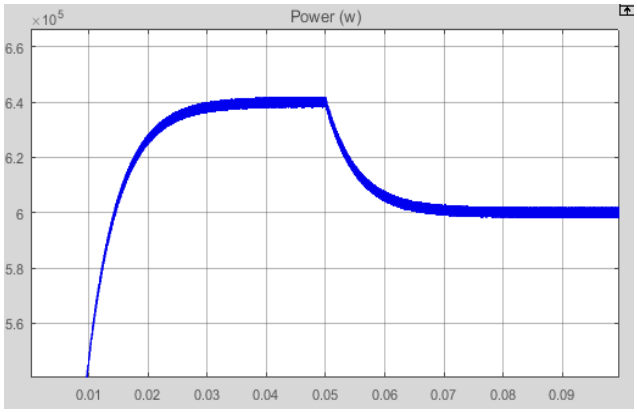


Fig. 11 Output power of FB-VSC with PI controller (Zoomed)

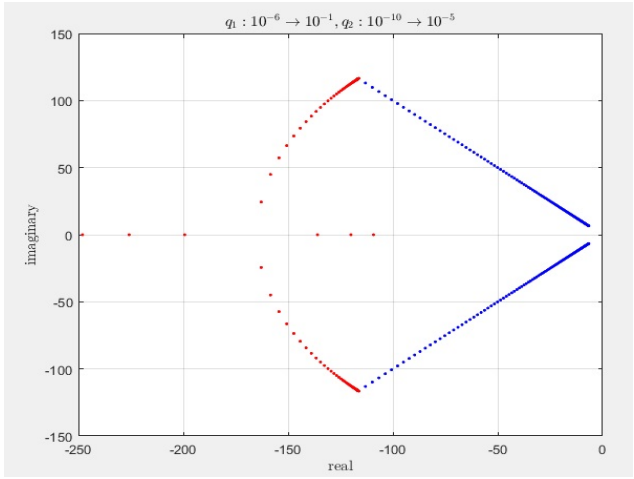


Fig. 12. root locus of close loop poles.

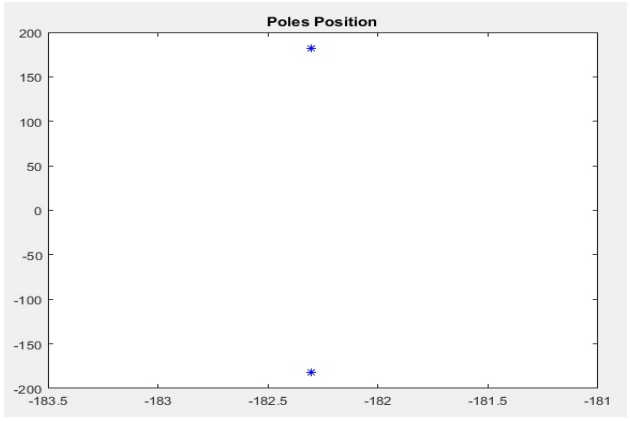


Fig. 13. Close loop pole position for k_1=-0.7071 & k_2=0.0036.

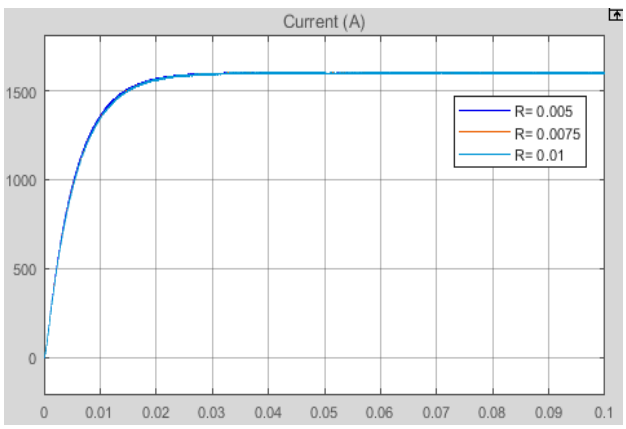


Fig. 14 Output current of FB-VSC with optimal controller (Zoomed)

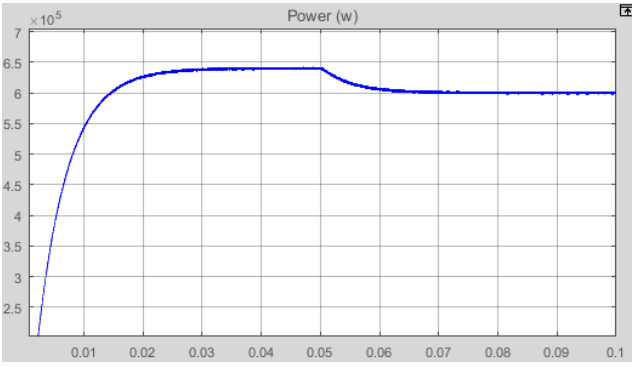


Fig. 15 Output power of FB-VSC with optimal controller (Zoomed)

iii) FB-VSC with feedback loop using optimal controller:

An optimal controller (designed in section IV) is applied in the FB-VSC simulation circuit shown in Fig.7. The value of k_1 and k_2 are obtained by using LQR technique. Calculate **A** and **B** using given parameter

$$A = \begin{pmatrix} 0 & -1 \\ 0 & -7.143 \end{pmatrix} \quad B = \begin{pmatrix} 0 \\ 1714285.714 \end{pmatrix}$$

$$Q = \begin{pmatrix} q_1 & 0 \\ 0 & q_2 \end{pmatrix} \quad Z = \begin{pmatrix} e \\ i \end{pmatrix} \; and \; R = 1$$

In this case, q_1 is varied from 10^{-5} to 10^{-1} and q_2 varied from 10^{-10} to 10^{-5}. To meet the time constant τ = 5ms, q_1 and q_2 are selected as $5e^{-1}$ and $2e^{-8}$ respectively. Fig.11 shows the root locus of close loop poles. Then k_1 and k_2 are obtained by using MATLAB command **lqr(A,B,Q,1).**

$$k_1 = -0.7071 \text{ and } k_2 = 0.0036$$

Simulation scenario (1):

i) Gating signal is applied at t=0 s. V_g jumps up 3% at t=0.05s, and V_{dc} jumps up 1% at t=0.07s.
ii) The value R is increased by 50% and100%.

Simulation scenario (2):

i) Reference power (P) changes from 640 KW to 600 KW at sec 0.05 s.

Fig.13 shows that in spite of changing V_{dc}, V_g, and R by the same value, the converter output current remains same. And from fig.11 it is also observed that when the reference power (P) changes from 640KW to 600KW at t=0.05 s, it takes only 0.015 s to track the change (less than 0.02 s takes by PI controller). So from these study, it is evident that the optimal controller is more robust to cancel various system uncertenities and disturbnce. Optimal controller also exhibits faster dynamics compare to PI controller.

VI. Conclusion

A Systematic design procedure and a new controller structure of single-phase full bridge VSC for renewable and distributed energy system have been addressed in this paper. The proposed controller has designed using the optimal control technique. A comparative study between this proposed optimal controller and conventional PI controller has also presented in this paper. The proposed optimal controller shows more robustness than PI controller to deal with external disturbance and system uncertainties.

Simulation results confirm effectiveness of the proposed optimal controller. In the future work, this proposed controller will be implemented in a prototype system.

REFERENCES

[1] Matallanas E., Castillo-Cagigal M., Gutiérrez A., Monasterio-Huelin F., Caamaño Martín E, Masa D., Jiménez-Leube J., "Neural network controller for Active DemandSide Management with PV energy in the residential sector," *Applied Energy*, Jun 2012. vol. 91, pp. 90-97,

[2] Dufo-López R., Zubi G, Fracastoro G. V., "Tecno-economic assessment of an off grid PV-powered community kitchen for developing regions," *Applied Energy*, Feb 2012. vol. 91, pp. 255-262

[3] P. Andea, A. V. Mnerie, F. Solomonesc, O. Pop and D. Cristian,"Conventional vs. alternative energy sources overview. Part II. European strategies," in *Proc. Int. Conf. Computational Cybernetics and Technical Informatics*, Timisoara, Romania, May 2010. pp. 601-606,

[4] M. K. Islam, M. M. Rahman and M. F. Rabbi, "Transformer less, lower THD and highly efficient inverter system," *2015 International Conference on Advances in Electrical Engineering (ICAEE)*, Dhaka, 2015, pp.293-296.

[5] Dufo-López R., Bernal-Agustín J. L., Yusta-Loyo J. M., Domínguez-Navarro J. A., Ramírez-Rosado I. J., Lujano J., Aso I., "Multi-objective optimization minimizing cost and life cycle emissions of stand-alone PV-wind-diesel systems with batteriesstorage," *Applied Energy,* May 2011. vol. 88, pp. 4033-4041.

[6] Wei Zhou W., Lou C., Li Z., Lu L., Yang H., "Current status of research on optimum sizing of stand-alone hybrid solar-wind power generation systems," *Applied Energy*, vol. 87, no.2, pp.380-389, Feb 2010.

[7] R. O'Gorman and M. A. Redfern, "Voltage control problems on modern distribution systems," *IEEE Power Engineering Society General Meeting, 2004.*, Denver, CO, 2004, pp. 662-667 Vol.1.

[8] Q. Li and P. Wolfs, "A Review of the Single-Phase Photovoltaic Module Integrated Converter Topologies With Three Different DC Link Configurations," in *IEEE Transactions on Power Electronics*, vol. 23, no. 3, pp. 1320-1333, May 2008.

[9] X. Yuan, F. Wang, D. Boroyevich, Y. Li and R. Burgos, "DC-link Voltage Control of a Full Power Converter for Wind Generator Operating in Weak-Grid Systems," in *IEEE Transactions on Power Electronics*, vol. 24, no. 9, pp. 2178-2192, Sept. 2009.

[10] M. Karimi-Ghartemani, S. A. Khajehoddin, P. Jain and A. Bakhshai, "A Systematic Approach to DC-Bus Control Design in Single-Phase Grid-Connected Renewable Converters," in *IEEE Transactions on Power Electronics*, vol. 28, no. 7, pp. 3158-3166, July 2013.

[11] Q. Li and P. Wolfs, "A Review of the Single Phase Photovoltaic Module Integrated Converter Topologies With Three Different DC Link Configurations," in *IEEE Transactions on Power Electronics*, vol. 23, no. 3, pp. 1320-1333, May 2008.

[12] N. Saxena, I. Hussain, B. Singh and A. L. Vyas, "Single phase multifunctional VSC interfaced with solar PV and bidirectional battery charger," *2016 7th India International Conference on Power Electronics (IICPE)*, Patiala, 2016, pp. 1-6.

[13] N. Krikelis, E. G. Papadopoulos, "An optimal design approach for tracking problems and its assessment against classical controllers", *Int. J. Control*, may 1982, vol. 36, no. 2, pp. 249-265.

[14] M.Karimi Ghartemani, S. Khajehoddin, P. Jain, A. Bakhshai, "Linear quadratic output tracking and disturbance rejection," International journal of Control, Jul 2011, Vol. 84, no. 8, pp. 1442-1449.

[15] B. Anderson, J. Moore, D. Naidu, Optimal Control: Linear Quadratic Methods, Englewood Cliffs, NJ, USA: Prentice-Hall, 1990.

Implementation of Finite Control Set Model Predictive Control on 7-level Flying Capacitor Multilevel Inverter Using Space Vector

Nasir Jubair Islam, Md. Selim Sarker*, N.M.Merazul Islam, Md.Shamim Reza
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka-1205,Bangladesh
*E-mail:selimsarker9497@gmail.com

***Abstract*—Multilevel inverter for industrial application has gained tremendous popularity in recent years. That's why an urge to develop simpler topologies and control techniques has drawn the attention of the researcher community. The flying capacitor multilevel inverter (FCMLI) is one such topology. FCMLI is an attractive alternative for medium voltage applications as it does not need complex transformers to obtain the dc-link voltage and also present good robustness characteristics. Different controlling techniques can be applied for modulation purpose in FCML inverters i.e. Pulse Width Modulation, Space Vector Modulation and Model Predictive Control. The standard procedure of predictive control is basically building a model of the system to predict the future values of the controlled variables. This information is utilized by the controller to achieve the optimal actuation, according to a predefined optimization criterion. This paper proposes the implementation of Finite Control Set Model Predictive Control with Space Vector technique applied to 7-level Flying Capacitor Multilevel Inverter. Our method controls output currents and voltages which allows one to increase the number of output voltage levels and reduce harmonic contents. It also ensure that output current contain minimum amount of harmonic. The feasibility of the proposed method operating with different loads and transient conditions has been verified by MATLAB/Simulink.**

Keywords—model predictive control, flying capacitor multilevel inverter, space vector, harmonic content

I. INTRODUCTION

Multilevel voltage source inverters has emerged as a new variety of power converter options for high power applications [1]. The multilevel voltage source inverters typically synthesize the staircase voltage wave from several levels of DC capacitor voltages. One of the major drawbacks of the multilevel inverters is it's inability to balance the voltage between different levels. The techniques to overcome the problem typically involve voltage clamping or capacitor charge controlling. There are multiple ways of achieving voltage balance in multilevel converters. The most popular multilevel converter topologies are the cascaded H-bridge converter [2], the modular multilevel converter [3], the neutral point clamped converter [4], and the flying capacitor Multilevel Inverter(FCMLI) [5].

The FCMLI is such a topology for high voltage applications that doesn't rely on any series connection of the power switches. It is a symmetric topology. It has some advantages over other types of converters like Diode Clamped Multilevel Inverter, Cascaded H-Bridge Multilevel Inverter. One of its significant advantages is that one can increase the number of voltage levels by simply adding basic cells in the load end of each leg. The prime challenge of implementing FCMLI is to maintain voltage balance on the different capacitors. To deal with the problem redundant switching states are used in each leg of the inverter with the help of adequate modulation scheme. Different connections of the flying capacitor is used to obtain different voltage levels of an FCMLI. It is necessary to ensure that all the flying capacitors ultimately reach a stable and constant voltage, so that the net charge variation become zero . Normally, FCMLI is implemented using Space Vector Modulation or Phase Disposition Pulse Width Modulation (PDPWM) techniques [6]-[7]. These techniques have good performance in controlling the three phase output voltage and reducing the harmonic content. But the output current has relatively high harmonic content (around 2 to 5 %). Also, in PDPWM technique the output current has a significant dc component in first few cycles. To deal with this problem we use Finite Control Set Model Predictive Control (FCS-MPC) in our FCMLI design to have smooth sinusoidal current and staircase voltage wave at load side with less harmonic.

FCS-MPC [8] (also known as direct MPC [9]) is considered to be the best predictive control for power inverters. In this method, the predictive control strategies precisely treat the power switches as control input of the system. In FCS-MPC a finite set of switch combinations is possible as power switches can have two particular states, 1 or 0. Therefore, a finite number of predictions can be obtained for the different electrical variables. The optimal solution is achieved by evaluating the cost function for each possible switching combination at each sampling interval. In this paper we use space vector technique to find the optimum voltage vector that will minimize the cost function (mainly, the difference between reference current and predicted current) in every time interval. Then the optimal switching combination is implemented in the inverter. As this method strictly ensures that the output current follows the predefined reference current, the output current is perfectly sinusoidal with less harmonic. The output current also has average value zero with no dc component. FCS-MPC can also deal with more complex target by including them in the cost function, e.g. reducing the common mode voltage [10] and switching loses [11]. Considering the performance, FCS-MPC allows one to achieve a fast transient response when load change or reference change occur compared to other modulation schemes. Also, this predictive technique allows the inverter to avoid any unwanted overcurrent condition,by adding safety limits for the electrical variables as constraints [12].

II. FLYING CAPACITOR MULTILEVEL INVERTER

The Flying Capacitor Multilevel Inverter (FCMLI) has been presented by Meynard and Foch [5]. It is a very efficient multilevel topology for medium voltage application because unlike cascade multicell converters these inverters do not rely on any isolated dc sources and, hence, big multi-secondary transformers.

An *m*-level FCMLI has *m*– 1 stages, where each stage is controlled with a switching function (s_j). The *j*th stage of a FCMLI consists of complementary switches with (*m*– *j*) capacitors connected in series. In real application, these switches are high speed semiconductor switches like MOSFET (Metal Oxide Semiconductor Field Effect Transistor), Thyristor or IGBT (Insulated Gate Bipolar Transistor) [13]-[14]. For generalized topology the voltage across each capacitor (V_c) maintains the same value. So, for an *m*-level converter,

$$V_c = \frac{V_{dc}}{m-1} \tag{1}$$

The generated leg voltage V_{aN} due to the collective action of the switching functions of each stage is given by

$$V_{aN} = \frac{V_{dc}}{m-1}\sum_{j=1}^{m-1} s_j \tag{2}$$

An *m*-level FCMC has (*m*– 1)(*m*– 2)/2 flying capacitors in each leg of the inverter with (*m*– 1) capacitors on the DC bus. These capacitors are named as flying capacitor as they are floating with respect to earth's potential and not connected to the ground. The number of capacitors in each stage can be reduced if the flying capacitors are rated for different voltages. Then, the (*m*– *j*) capacitors of each stage can be replaced by only one capacitor, and each leg will be consisting of only (*m*– 1) capacitors. In fig. 1 one leg of a seven level converter is shown where one capacitor is used in per stage. Here the ratio of voltage in different capacitors

$$v_{c_{6a}} : v_{c_{5a}} : v_{c_{4a}} : v_{c_{3a}} : v_{c_{2a}} : v_{c_{1a}} = 6:5:4:3:2:1 \tag{3}$$

FCMLI has many redundant states for which the output voltages are same. These redundant states play a vital role to maintain the average voltage on each flying capacitor. This condition ensures that the overcharge or undercharge in one state is removed or regained in the redundant state. In order to maintain constant average voltage on C_j ($j \in \{1,2,3,4,5,6\}$) by using those redundant states, it is necessary that the time span the capacitor is being charged and the time span capacitor is being discharged in one cycle should be same which ensures no net charge (ΔQC_j) along one cycle. So, the following criterion must be satisfied along a switching cycle T_c, in steady state:

$$\Delta Q_{c_j} = \int_{t_c} I_{c_j}.dt = 0 \tag{4}$$

To ensure the above relationship it is essential to know the direction of current flow and the voltage magnitude across the capacitor to understand whether it is charging or discharging.

III. FINITE CONTROL SET MODEL PREDICTIVE CONTROL

Model Predictive Control (MPC) is widely used control technique for various power inverters [15]-[17]. In this technique the model of system is used to accurately predict the future behavior of the controlled variables. The time horizon is divided into several interval and each interval, the optimum solution of controlled variables for next interval or

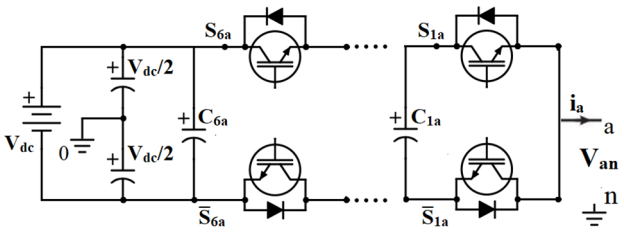


Fig. 1. One leg of a 7-level FCMLI

intervals is predicted using the value of controlled variable in present interval. A cost function is developed using the system model and the cost function is calculated for the predicted values of the variables to be controlled. Prediction of the future value of these variables is calculated for each possible switching state and then the state that minimizes the cost function is selected to be used in next time interval. In FCS-MPC (Finite Control Set Model Predictive Control) the number of combinations of controlled variables is a finite value and from this finite number of combinations the most preferable combination is found and implemented in next interval.

The control strategies we followed in this paper to implement FCS-MPC are

- Building an inverter model and its possible switching states
- Building a model of load for prediction
- Defining a cost function and finding appropriate switching state to minimize it

A. Inverter Model

The inverter used in this paper is a 7-level Flying Capacitor Inverter. It has six stages in each leg and each stage has two complementary switches with capacitor connected across them. The switching states and corresponding leg voltage is shown in Table I. We can see that 0 state can be achieved by two different combination. To deal with this without incorporating this constraint in cost function we measured the current flowing through the c_{3a} capacitor (the capacitor at third stage) $i_{c_{3a}}$ and voltage across the capacitor $v_{c_{3a}}$ and take appropriate action according to Table II. where, V_{ref} is the reference voltage need to be maintained across capacitor.

We know in balanced three phase connection three phase line current and three phase line to neutral voltages are balanced. So, we can write

$$i_a + i_b + i_c = 0 \tag{5}$$

Table I. SWITCHING COMBINATION

.

S_{6a}	S_{5a}	S_{4a}	S_{3a}	S_{2a}	S_{1a}	V_{aN}
1	1	1	1	1	1	$V_{dc}/2$
1	1	1	1	1	0	$V_{dc}/3$
1	1	1	1	0	0	$V_{dc}/6$
1	1	1	0	0	0	0
0	0	0	1	1	1	0
0	0	0	0	1	1	- $V_{dc}/6$
0	0	0	0	0	1	- $V_{dc}/3$
0	0	0	0	0	0	- $V_{dc}/2$

Table II. CONTORLLING CAPACITOR VOLTAGE

System State	***Switching state for one leg***
if ($i_{c_{3a}} \geq 0$) and ($v_{c_{3a}} \geq V_{ref}$) or if ($i_{c_{3a}} < 0$) and ($v_{c_{3a}} < V_{ref}$)	**S**= [0,0,0,1,1,1]
if ($i_{c_{3a}} \geq 0$) and ($v_{c_{3a}} < V_{ref}$) or if ($i_{c_{3a}} < 0$) and ($v_{c_{3a}} \geq V_{ref}$)	**S**= [1,1,1,0,0,0]

and also

$$V_{aN} + V_{bN} + V_{cN} = 0 \tag{6}$$

Now in this balanced condition as the quantities are not independent, the three phase components of a three phase quantity are determined by two components and therefore, can be described by a plane vector [18]. This is known as space vector. So, the three phase current can be converted into current vector and three phase voltage can be converted into voltage vector. In this paper, we used this approach to predict the future voltage and current of the inverter which enables us to avoid computational burden of three phase quantity. If the real and imaginary component of current vector is i_α and i_β, then

$$\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} = \frac{2}{3}\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}\begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} \tag{7}$$

B. Load Model:

The load we used in our simulation is a typical RL branch. If the voltage vector applied in the RL branch is **v,** the load current vector is **i** and then the dynamic operation of the load can be described by the differential equation

$$\mathbf{v} = R\mathbf{i} + L\frac{d\mathbf{i}}{dt} \tag{8}$$

We used discrete time model to predict the future value of current in different time interval. Equation (8) is discretized to implement FCS-MPC so that we can predicted the value of current in different interval of time horizon. The sampling time for the discretization process is expressed as T_s. At any sampling instant *n*th the discrete time model of the load is utilized to predict the future value of load current from applied voltages and measured currents. Different discretization principle can be implemented to obtain a suitable discrete time model for calculating future values. In this paper we used forward Euler approximation to replace $\frac{d\mathbf{i}}{dt}$ term by following equation,

$$\frac{d\mathbf{i}}{dt} \approx \frac{\mathbf{i}(n+1)-\mathbf{i}(n)}{T_s} \tag{9}$$

By using this approximation, we can predict the value of current vector of future interval *n+1* by using the value at *n*th interval. Then the predicted value of future current can be written as,

$$\mathbf{i^p}(n+1) = \left(1 - \frac{RT_s}{L}\right)\mathbf{i}(n) + \frac{T_s}{L}\mathbf{v}(n) \tag{10}$$

Where $\mathbf{i^p}$ is the predicted current vector and **v** is the applied voltage vector.

C. Design of cost function

The cost function is designed to minimize the difference between the measured load current and the predefined reference current. The expression of cost function of designed system is,

$$c = \left|i_\alpha^r(n+1) - i_\alpha^p(n+1)\right| + \left|i_\beta^r(n+1) - i_\beta^p(n+1)\right| \tag{11}$$

Here, $i_\alpha^p(n+1)$ and $i_\beta^p(n+1)$ are real and imaginary component of the predicted current vector and $i_\alpha^r(n+1)$ and $i_\beta^r(n+1)$ are real and imaginary component of reference current vector at $(n+1)$th instant for the applied voltage vector. To reduce complexity, we ignore the change of reference current in one time interval. So, we consider that

$$\mathbf{i^r}(n+1) = \mathbf{i^r}(n) \tag{12}$$

In each interval equation (11) is evaluated for each voltage vector and the voltage vector generates minimum error is applied in next time interval with its corresponding switching combination

IV. SIMULATION RESULT

The effectiveness of the proposed design is confirmed by doing several simulations in Matlab (Simulink) environment. The circuit parameters used for this simulation is given in Table III and Table IV. The reference current is predicted as 1.9A. The prediction is done by considering the fact that by applying a DC bus voltage of Vdc, we are able to get highest phase voltage of (2/3)*Vdc. We have applied 533V as DC bus voltage and consequently we have got highest phase voltage 355V in output. If the output voltage was a pure sine wave (with 355V peak value) then the peak value of output current would be 2.1A for the RL load of 10 ohm and 500 mH (in 50 Hz system). But as our output voltage is not a pure

Table III. INVERTER MODEL PARAMETER

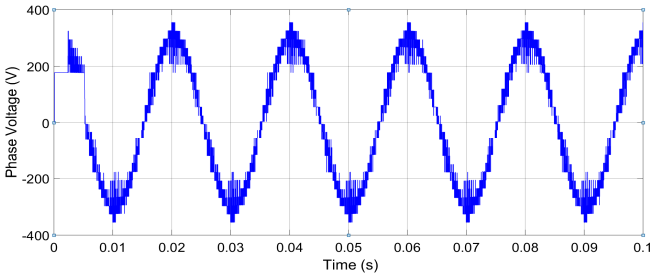
Type	*Value*
DC-Link Voltage V_{dc}	533V
Reference Current I_{ref}	1.9A
DC-Link Capacitor C_{dc}	22 mF
Flying Capacitor C_j	22 mF
Sampling Frequency F_s	14400 Hz
System Frequency f	50 Hz

Table IV. LOAD MODEL PARAMETER

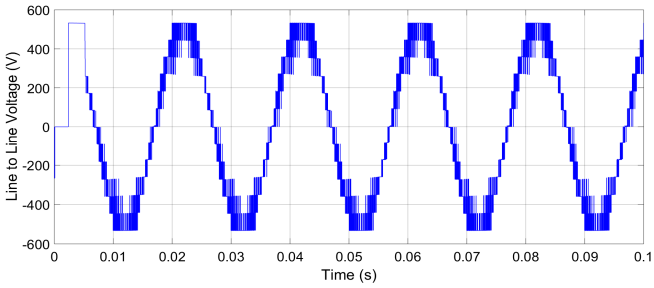
Type	*Value*
Resistance R	10Ω
Inductance L	500 mH

Table V. HARMONIC CONTENT

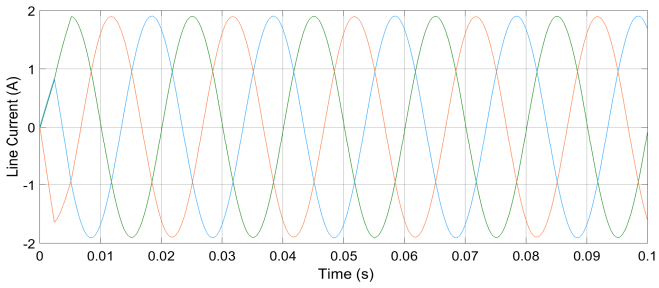
Voltage	*5-level FCMLI*	*7-FCMLI*
Line to Line	18.88%	16.65%
Phase	18.44%	16.09%



(a)

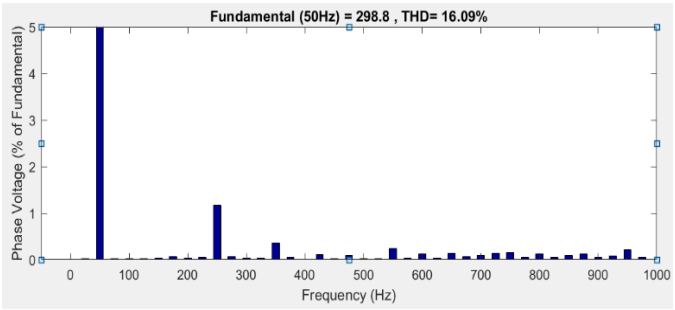


(b)

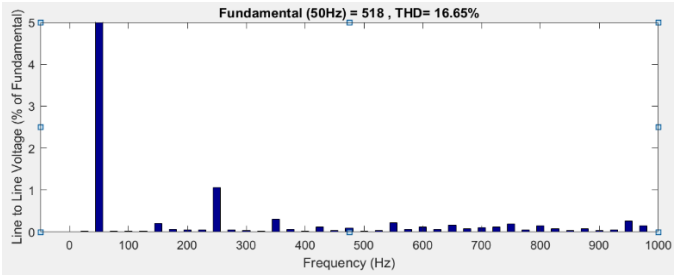


(c)

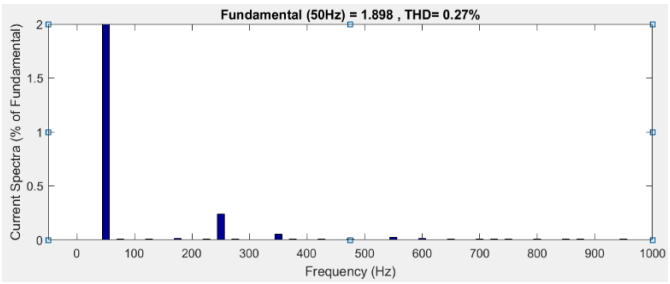
Fig. 2. Simulation result of 7-level FCMLI (a) Phase voltage (b) Line to Line Voltage (c) Line Current



(a)

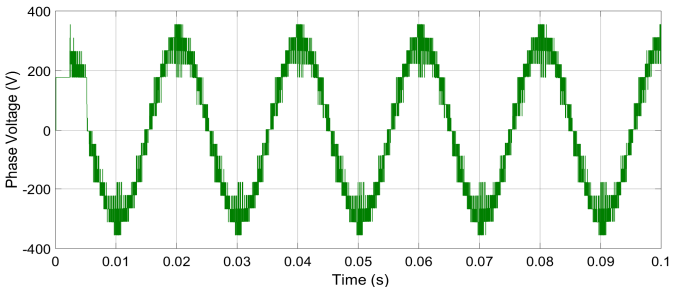


(b)

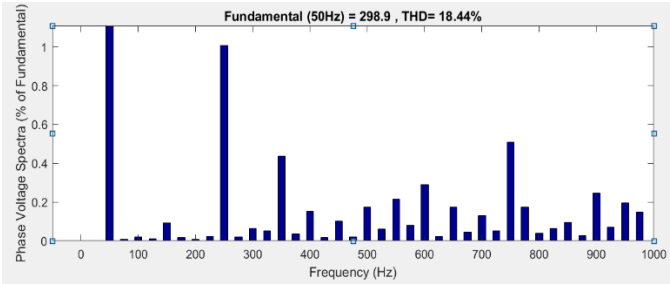


(c)

Fig. 3. Spectral characteristics of 7-level FCMLI (a) Phase voltage (b) Line to Line Voltage (c) Line Current

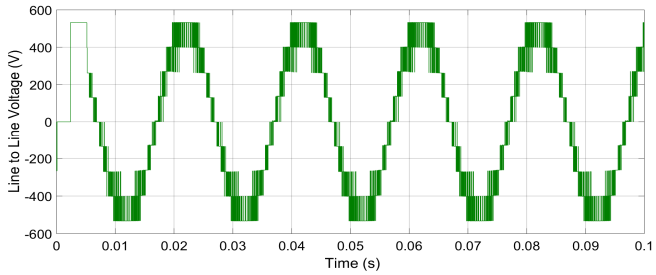


(a)

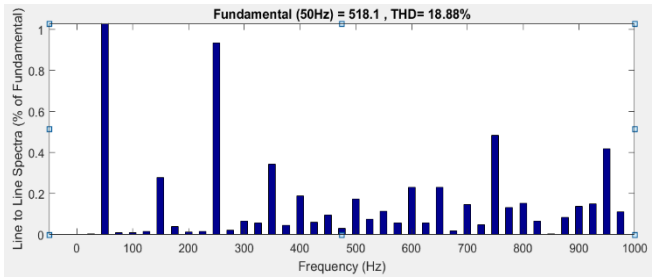


(b)

Fig. 4. Performance of 5-level FCMLI (a) Phase voltage (b) Spectral Characteristic

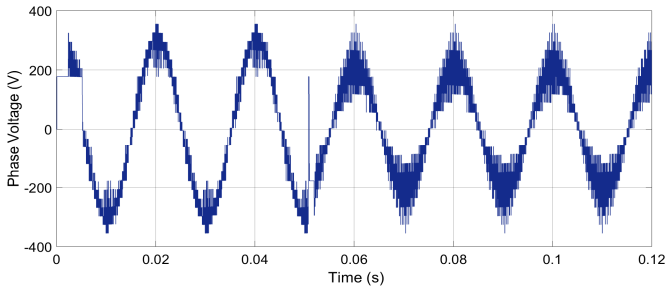


(a)

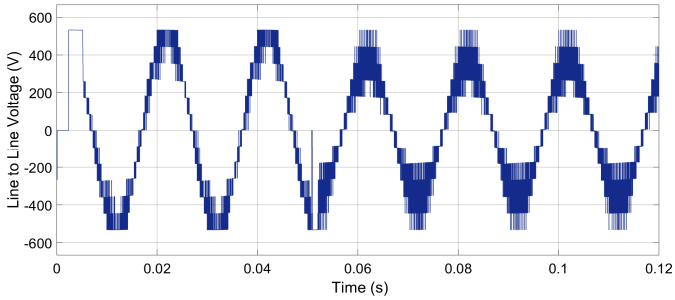


(b)

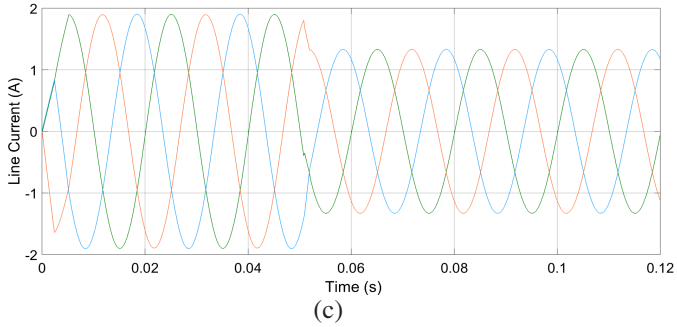
Fig. 5. Performance of 5-level FCMLI (a) Line to Line voltage (b) Spectral Characteristic



(a)



(b)



(c)

Fig. 6. Response of 7-level FCMLI on sudden change in reference current at t =50 ms (a) Phase voltage (b) Line to Line Voltage (c) Line Current

sine wave, the suitable value for predicting the peak output current is below 2.1A. By using trial and error we find 1.9A is the best value for predicting reference current magnitude In Fig. 2 the simulated value of Output Load Current, Line to Line voltage and Phase voltage are shown. From the figure we can see that initially the graphs are inconsistent due to transient effect of RL load. But the problem is overcome in first half cycle. Again, the output current has no dc component after first cycle which solve the problem other techniques where the output current has significant dc component. Fig.3 shows the harmonic content of current and voltage. This figure indicates that the harmonic content is almost zero (.27%) for output current. This is one of the important features of this method. As our technique ensures that the output current follows the reference current, the harmonic content is greatly reduced. The harmonic content of Line voltage and Phase voltage is high, but these values are much lower for 7-Level FCML compared to a 5-Level FCMLI which values are shown in Fig. 4 and Fig. 5. and Table V. This design can also efficiently react on any sudden reference change like change in reference current or sudden load change which is another benefit of this topology. Like in Fig. 6 at t = 50 ms the reference current I_{ref} is changed from 1.9A to 1.33A. The Output Load Current, Line to Line voltage and Phase voltage quickly cope with the changed reference with little discrepancy.

V. CONCLUSION

This paper presented an analysis on 7-level Flying Capacitor Multilevel Inverter using Model Predictive Control technique. It shows that we can ensure better output with less harmonic distortion if we increase the voltage level and design the system with appropriate parameter. It also have very good transient response. But the most significant characteristics of Model Predictive Control is that it can reduce the dc component of output current to zero with almost perfect sinusoidal current of minimum harmonic content. This method can be developed by adding more level or by using fixed switching frequency for every MOSFET which will further reduce the harmonic content of Line to Line voltage and Phase voltage.

REFERENCES

[1] Sergio Albarto Gonzalez, Santiago Andres Verne, and Maria Ines Valla, “Multilevel converters for industrial applications, ”1st Edition, CRC Press (2017)

[2] F. Z. Peng, J. S. Lai, J. W. McKeever, and J.VanCoevering, “A multilevel voltage-source inverter with separate DC sources for static VAr generation,” IEEE Transaction on Industrial Application, vol. 32, no. 5, pp. 1130-1138, Sep./Oct. 1996.

[3] A. Lesnicar and R. Marquardt, “An innovative modular multilevel converter topology suitable for a wide power range,” in Proc. IEEE Power Technology Conference, 23-26 Jun. 2003, vol. 3, pp. 6.

[4] A. Nabae, I. Takahashi, and H. Akagi, “A new neutral-point-clamped PWM inverter,” IEEE Transaction on Industrial Application, vol. IA-17, no. 5, pp. 518-523, Sep/Oct. 1981.

[5] T. A. Meynard and H. Foch, “Multi-level conversion: High voltage choppers and voltage-source inverters,” in Proc. 23rd Annual IEEE Power Electronics Specialists Conference, 29 Jun.-3 Jul. 1992, vol. 1, pp. 397-403.

[6] M.M Shobini, J. Kamala & R. Rathna, “Analysis and simulation of flying capacitor multilevel inverter using PDPWM strategy,”

International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), 2017.
[7] S. Choi & M. Saeedifard, "Capacitor Voltage Balancing of Flying Capacitor Multilevel Converters by Space Vector PWM," IEEE Transactions on Power Delivery, Vol. 27, issue 3, July 2012.
[8] S. Kouro, P. Cortés, R. Vargas, U. Ammann, and J. Rodriguez, "Model Predictive Control—A Simple and Powerful Method to Control Power Converters," IEEE Transactions on Industrial Electronics, vol. 56, no. 6, pp. 1826–1838, Jun. 2009.
[9] A.Linder and R. Kennel, "Direct model predictive control – a new direct predictive control strategy for electrical drives," in Power Electronics and Applications, 2005 European Conference on, 2005, p.10.
[10] R. Vargas, U. Ammann, J. Rodriguez, and J. Pontt, "Predictive Strategy to Control Common-Mode Voltage in Loads Fed by Matrix Converters,", IEEE Transactions on Industrial Electronics, vol. 55, no. 12, pp. 4372–4380, 2008.
[11] R. Vargas, U. Ammann, and J. Rodriguez, "Predictive Approach to Increase Efficiency and Reduce Switching Losses on Matrix Converters," IEEE Transactions on Power Electronics, vol. 24, no. 4, pp. 894–902, 2009.
[12] D. E. Quevedo, R. P. Aguilera, M. A. Pérez, P. Cortés, and R. Lizana, "Model Predictive Control of an AFE Rectifier with Dynamic References," IEEE Transactions on Power Electronics, vol. 27, no. 7, pp. 3128–3136, 2012.
[13] T. Jing, & Z. Chang-Hao, "The model of IGBT as a power switch," 2011 International Conference on Electronics, Communications and Control (ICECC)].
[14] C.W. Mueller, J. Hilibrand, "The "Thyristor"—A new high-speed switching transistor," IRE Transactions on Electron Devices, vol. 5, issue 1, Jan. 1958).
[15] Jose Rodriguez, Patricio Cortes, "Predictive control of power converters and electrical drives," 2012, John Wiley & Sons, Lt.
[16] Pablo Lezana, Ricardo Aguilera, and Daniel E. Quevedo, "Model predictive control of an asymmetric flying capacitor converter," IEEE Transaction on Industrial Electronics, Vol. 56, no. 6, June 2009.
[17] Males Tomlinson, Toit Mouton, Ralph Kennel and Peter Stolze, "Model predictive control with a fixed switching frequency for a 5-level flying capacitor converter," 2013 IEEE ECCE Asia Downunder, 3-6 June 2013
[18] Peter Vas, "Electrical machines and drives: A space-vector theory approach," Clarendon Press, Oxford (1992).

Design and Comparative Analysis of Dual-stage Servo controller in Hard Disk Drive

Alamgir Hossain* and Md. Arifur Rahman
Department of Electrical and Electronic Engineering,
Dhaka University of Engineering & Technology, Gazipur
Gazipur, Bangladesh
Email: eeealamgir@duet.ac.bd and marahman@duet.ac.bd

***Abstract*— In modern times, the design and optimization of different control systems for servo control of a hard disk drive (HDD) is popular interdisciplinary research area. Nowadays, dual-stage actuator system is widely used in HDD servo system, where voice coil motor (VCM) actuator and piezo-electric micro actuator are used as the primary actuator and secondary actuator, respectively. In this paper, different combinations of PI, PD, and PID are employed for primary and secondary actuators. The tuning method is applied on these combinations to obtain the best performance. After a rigorous comparative analysis, this paper proposes the optimized method among these different combinations by analyzing and comparing the results obtained through MATLAB/Simulink simulation.**

Keywords—hard disk drive (HDD); voice coil motor (VCM); piezo-electric (PZT); actuator;

I. INTRODUCTION

A hard disk drive (HDD), a very popular storage system used in computers and data-processing, is a magnetic storage device to record and recover digital information in magnetic disks. The disks are coated with magnetic materials and rotated at a specific speed. Information bits are recorded in circular rounds, known as tracks, on the disks. The read/write heads are installed in a slider and the slider is attached to the tip of an actuator arm. This actuator arm drives the read/write head in the horizontal direction over the disk surfaces and the motion of this head is controlled by head positioning servomechanism. This servomechanism has two primary actions in the disk drive systems; one is track seeking and another one is track following [1]. In a dual-stage HDD shown in Fig. 1, the mechanical and eletromechanical parts of the servo system consist of the VCM, the suspension, piezoelectric micro-actuator made from Lead-Zirconate-Titanate (PZT), and the slider. The magnetic read/write (R/W) head is fabricated with the slider [2]. Early age HDDs use only one actuator, a VCM actuator that plays a vital role to move the R/W head. To achieve the high precision positioning control, the dual-stage actuator system has been emerged recently. As a result, the data capacity of HDDs can be increased with the availability of increased servo bandwidth [3]. Different control techniques are designed for dual-stage system by many researchers over the previous years. Optimization and simplification of these control approaches is a growing demanding research work in order to obtain a suitable controller for the dual-stage system to make it acceptable for the mass production. Dual-stage HDD servo system is a dual-input single-output (DISO) system[4]-[5] and for designing a servo controller for this DISO system diverse methods are already designed and optimized that have been described in the revealed articles [6]-[9]. For dual-stage controller, there are three popular configurations: (a) parallel loop configuration, (b) master-slave configuration, and (c) decoupled master-slave configuration [10].



Fig. 1. A Hard Disk Drive.

This paper represents the design and comparative analysis of performances of PI, PD and PID controllers in different combinations as the controllers for primary VCM actuator and secondary PZT micro-actuator in dual-stage HDD servo system. By using MATLAB and Simulink simulation, the open-loop bandwidth, gain margin, phase margin, rise time, settling time, and percentage overshoot of each combination are obtained and then compared to achieve the optimized performance.

II. PLANT MODEL

Dual-input single-output (DISO) system of a HDD servo system shown in Fig. 2 uses two actuators- one is VCM actuator as primary and another one is PZT as secondary actuator. Control signals from the controllers are applied to the actuators and system's combined output is the displacement of the R/W head. In the dual-stage HDD servo system, bandwidth of the system increases which thereby increases the positioning speed than the single-stage actuation system.

A. VCM Model

Voice coil motor actuator is mathematically modeled as a double-integrator model which can be expressed as the following form [6]:

$$G_V(s) = \frac{K_V}{s^2} \times \sum_{i=1}^{N_V} H_{ri} \quad (1)$$

where K_V is the gain, N_V is the total number of mechanical resonant modes of VCM actuator. Each resonant mode is represented by three parameters (A_i, ω_{ri} and ζ_i) as follows:

$$H_{ri} = \frac{A_i}{s^2 + 2\zeta_i \omega_{ri} s + \omega_{ri}^2}$$

VCM model used in this work has the frequency response shown in the Fig. 3 with four mechanical resonant modes. These resonant modes decrease the performance of the HDD servo system since this force the designer to reduce the system bandwidth in order to keep the system stable. So effect of these resonant modes should be compensated by using resonance compensator.

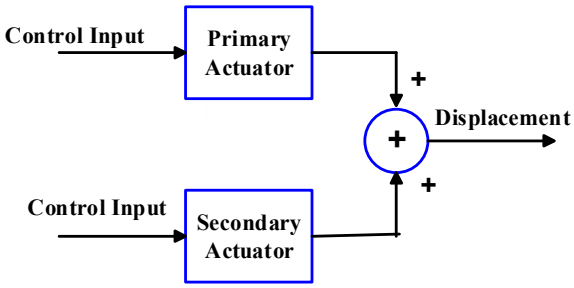


Fig. 2. Dual-Input Single-Output (DISO) system in HDD.

B. PZT Model

The secondary actuator PZT can be modeled in the following mathematical form [6]:

$$G_P(s) = K_P \times \sum_{j=1}^{N_P} R_j \quad (2)$$

where K_P is the gain, N_P is the number of the mechanical resonant modes of secondary actuator where each resonant mode can be expressed as,

$$R_j = \frac{a_i}{s^2 + 2\zeta_j \omega_{pj} s + \omega_{pj}^2}$$

The frequency response of secondary PZT actuator is shown in Fig. 4. Here, we deal with five mechanical resonant modes in the frequency response to further analysis. These resonant modes are also detrimental for the response of the plant output.

III. DUAL-INPUT SINGLE-OUTPUT CONTROLLER

A common design detailed description for controlling servo mechanism of dual-stage HDD is summarized in Table I. Due to the conflicting features of inertia, bandwidth and range of motion on both primary and secondary, our controller must be designed perfectly for these two actuators [10].

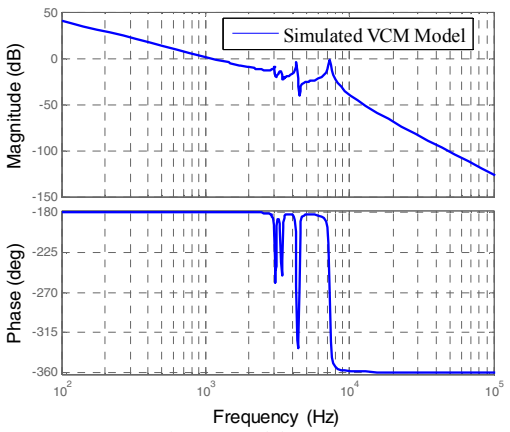


Fig. 3. Frequency response of VCM model.

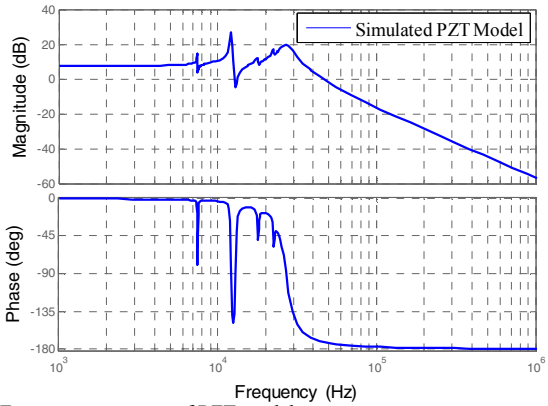


Fig. 4. Frequency response of PZT model.

TABLE I. DETAILED DESCRIPTION FOR A TYPICAL DUAL-STAGE CONTROLLER

Parameters	Required value
Open-loop bandwidth f_o :	>2000 Hz
Phase margin :	>40 degree
Gain margin :	>6 dB
Rise time :	<0.2 ms
Overshoot :	<20%

Fig. 5 shows the decoupled master-slave (DMS) controller architecture. In this configuration, the output of secondary controller is feed to input of the primary controller by adding with error signal. For this configuration, the transfer function

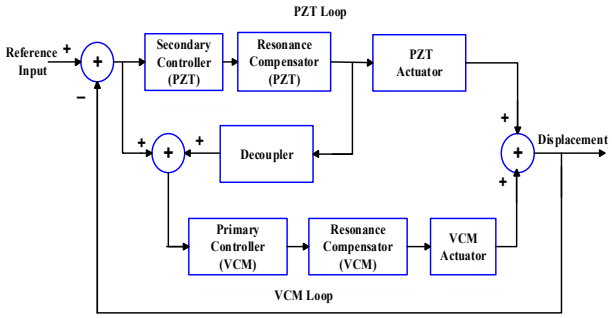


Fig. 5. Decoupled master-slave control architecture for dual-stage HDD.

of open-loop $L_{DMS}(s)$, closed-loop $T_{DMS}(s)$ and sensitivity $S_{DMS}(s)$ can be expressed as:

$$L_{DMS}(s) = (1 + C_M P_M) C_V P_V + C_M P_M,$$

$$T_{DMS}(s) = \frac{C_V P_V + C_M P_M + C_V P_V C_M P_M}{(1 + C_V P_V)(1 + C_M P_M)},$$

$$S_{DMS}(s) = \frac{1}{1 + C_V P_V} \frac{1}{1 + C_M P_M},$$

Where, C_V is the primary controller (VCM controller), C_M is the secondary controller (Micro actuator or PZT controller), P_V is the primary actuator (VCM actuator) and P_M is the secondary actuator (PZT actuator).

A. SIMULATION RESULTS

All the simulations of dual stage servo system are done by using MATLAB/Simulink. Fig. 6 shows the relation of output response among three configurations where PI controller used as primary controller and PI, PD, PID controller used as secondary controller respectively. In Fig. 6, using PI-PD controller, it is observed that, the steady-state error was not completely removed. On the other hand steady-state error can remove completely using both PI-PI and PI-PID.

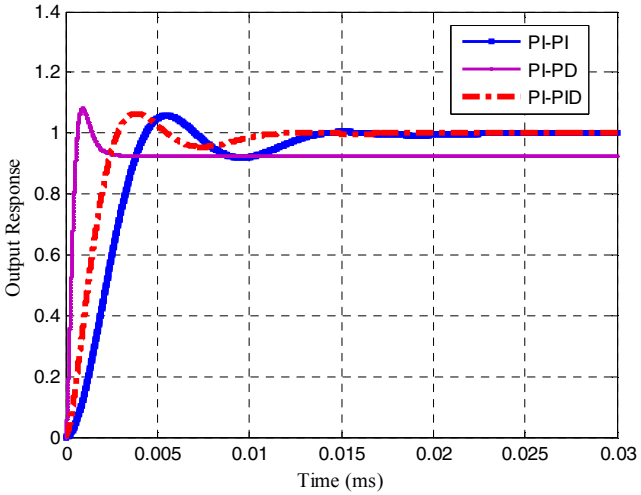


Fig. 6. Output Response of dual-stage controller where PI as primary controller.

Fig. 7 represents the comparison of output responses among the three configurations where PD controller is used as primary controller and PI, PD, PID controller are used as secondary controller individually. In Fig. 7, using PD-PD controller, it is observed that, the steady-state error was not completely removed. On the other hand steady-state error can remove completely using both PD-PI and PD-PID.

Fig. 8 shows the relation of output responses among three configurations where PID controller used as primary controller and PI, PD, PID controller used as secondary controller respectively. In Fig. 8, using PID-PD controller, it is observed that, the steady-state error was not completely removed. On the other hand steady-state error can remove completely using both PID-PI and PID-PID.

Fig. 9 and Fig. 10 presents the open-loop and closed-loop frequency response respectively where PI, PD and PID used as secondary controller. The rise time, settling time, percentage overshoot, open-loop bandwidth, phase margin and gain margin obtained from the output response and open-loop frequency response of different combination that are mentioned in Table III. PI, PD or PID controller in the primary loop results-in approximately same performance. However different types of controller in the secondary loop given different output responses. This can be observed from Fig. 6, Fig. 7, Fig. 8 and Table III.

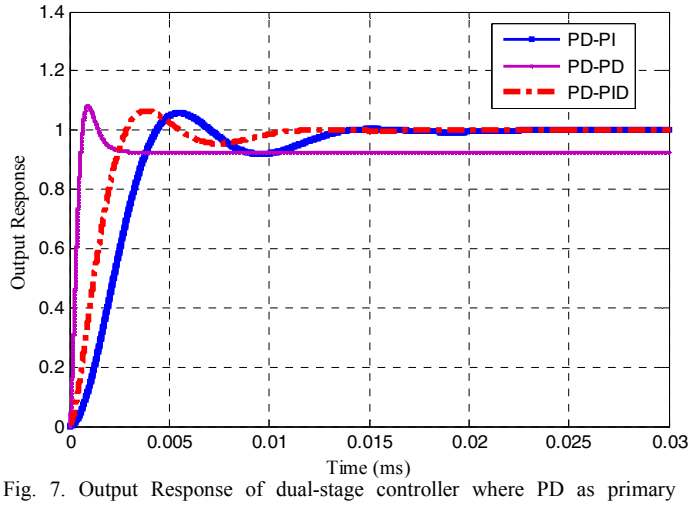


Fig. 7. Output Response of dual-stage controller where PD as primary controller.

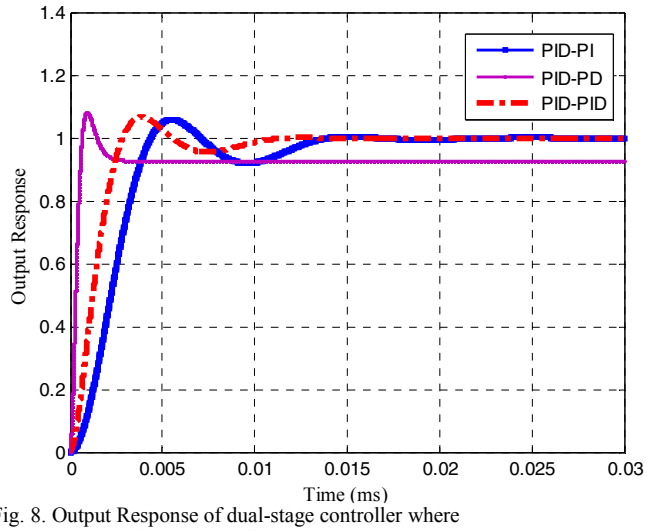


Fig. 8. Output Response of dual-stage controller where PID as primary controller.

Fig. 11 presents the error signal of three different configurations, where PI, PD and PID are used as the secondary controller individually. The standard deviation is calculated to determine the percentage improvement of each configuration (shown in Table II). It is observed that when PD

controller is used in secondary stage, 65.27% and 53.29% error can be improved than PI and PID controller. From the Table III it can be concluded that, PD controller in a secondary actuator controller has the best performance than others. It is observed from Table III, the bandwidth is comparatively high when PD used in secondary stage than that for other two controllers and phase margin is same.

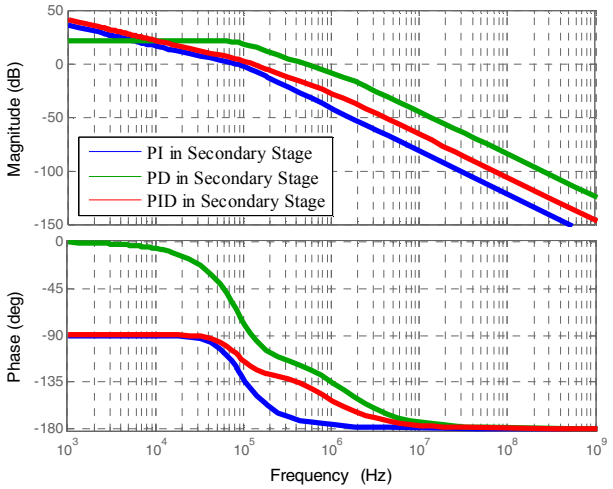


Fig. 9. Open loop frequency of different configuration.

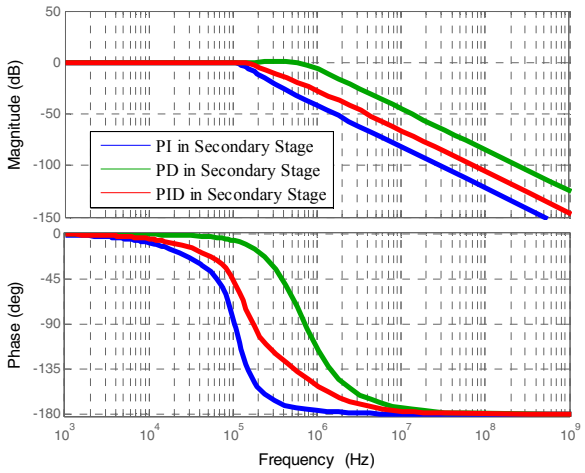


Fig. 10. Close loop frequency of different configuration.

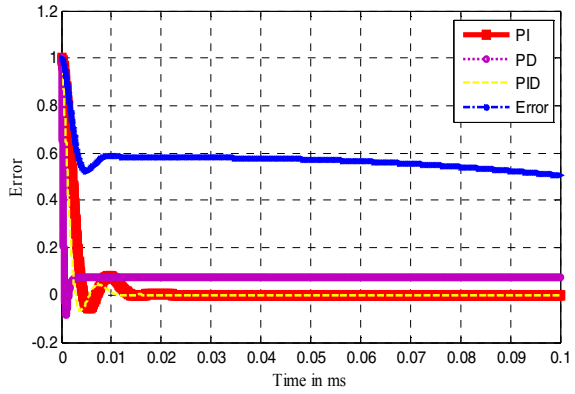


Fig. 11. Error signal.

TABLE II: STANDARD DEVIATION CALCULATION

Sl. No	Controller used as secondary	Standard deviation
01	PI	0.1267
02	PD	0.0440
03	PID	0.0942

TABLE III: SUMMARY OF RISE TIME, SETTLING TIME, PERCENTAGE OVERSHOOT, OPEN-LOOP BNADWIDTH, PHASE MARGIN AND GAIN MARGIN OF DIFFERENT CONFIGURATION

Sl. No.	Controller	Rise Time (µS)	Settling Time (µS)	%OS	Open-loop bandwidth f_0 (kHz)	Phase margin (Degree)	Gain margin (dB)
01	PI-PI	3.704	12.7	5.80	81.1	60	84.7
02	PI-PD	0.491	2.024	16.96	483	60	101
03	PI-PID	2.336	9.589	6.50	122	60	109
04	PD-PI	3.704	12.7	5.80	81.1	60	84.7
05	PD-PD	0.491	2.024	16.96	483	60	101
06	PD-PID	2.336	9.589	6.50	122	60	109
07	PID-PI	3.704	12.7	5.80	81.1	60	84.7
08	PID-PD	0.491	2.024	16.96	483	60	101
09	PID-PID	2.336	9.589	6.50	122	60	109

IV. CONCLUSION

The simulation results show the comparison among the different combination of PID controller used in HDD servosystem. When PD controller used in secondary stage, it shows the better and optimized performance than others. For this configuration we observed the rise time is 0.491 µs, settling time is 2.024 µs, open-loop bandwidth is 483 kHz, phase margin is 60 degree but percentage overshoot is 16.96% that is better than others and gain margin is 101 which is smaller than PID controller. Therefore, we can conclude that the PI-PD, PD-PD and PID-PD combinations can be the suitable controller configuration for the dual-stage servo system if the resonance is compensated properly.

REFERENCES

[1] Kemao Peng, Ben M. Chen, Tong H. Lee, V.Venkataramanan, "Design and implementation of a dual-stage actuated HDD servo system via composite nonlinear control approach", Mechatronics 14, pp. 965-988, 2004.

[2] Li, Yün-feng, "Dual-stage servo control and active vibration compensation in magnetic hard disk drives", PhD thesis, University of California, Berkeley, 2003.

[3] Huang, Xinghui, et al, "Design and analysis of a dual-stage disk drive servo system using an instrumented suspension", American Control Conference, Vol. 1. IEEE, pp. 535-540, 2004.

[4] Guo, Lin, Douglas Martin, and Don Brunnett, "Dual-stage actuator servo control for high density disk drives", Advanced Intelligent Mechatronics, pp. 132-137, 1999.

[5] Semba, T. et al., "Dual-stage servo controller for HDD using MEMS actuator", IEEE Transactions on Magnetics, Vol. 35, No. 5, pp. 2271-2273, 1999.

[6] Rahman, Md Arifur, Abdullah Al Mamun, and Kui Yao, "Analysis and modeling of hysteresis of piezoelectric micro-actuator used in high precision dual-stage servo system", Control Theory and Technology 13.2, pp. 184-203, 2015.

[7] Pang, Chee Khiang, et al, “Self-sensing actuation for nanopositioning and active-mode damping in dual-stage HDDs”, IEEE/ASME Transactions on Mechatronics 11.3, pp. 328-338, 2006.

[8] Guo, Guoxiao, Qi Hao, and Teck-Seng Low, “A dual-stage control design for high track per inch hard disk drives”, IEEE Transactions on Magnetics 37.2, pp. 860-865, 2001.

[9] Lee, Seung-Hi, Young-Hoon Kim, and Chung Choo Chung, “Dual-stage actuator disk drives for improved servo performance: Track follow, track seek, and settle”, IEEE transactions on magnetics 37.4, pp. 1887-1890, 2001.

[10] Abdullah Al Mamun, GuoXiao Guo and Chao Bi, Hard Disk Drive Mechatronics and Control, Vol. 23. CRC press, 2006.

Intelligent Speed Control of DC Motor using ANFIS D Controller

Tanjim Tarannum
Khulna University of Engineering and Technology,
Khulna-9203, Bangladesh.
tanjim729@gmail.com

***Abstract*—The main objective of this paper is to present the speed controlling method of a separately excited DC motor with Adaptive Neuro Fuzzy Inference System (ANFIS) based derivative controller. Efficient control of speed of DC motor has been achieved using by reducing the overshoot, undershoot, peak to peak ripple and steady state error by the proposed approach. A derivative controller has also been used to validate the performance improvement by comparison between the two methods. From the assessment, it clearly shows that, the suggested ANFIS D controller provides better performance including overshoot, undershoot, peak to peak ripple and steady state error the reduction and hence improving the speed control technique of separately excited DC motor.**

Keywords— Speed Control, DC Motor, ANFIS D controller, Derivative controller.

I. INTRODUCTION

The conversion of energy is done by electrical machines which are of different types. DC motor is a kind of electrical machine that takes electrical input and converts it into mechanical energy. Electric motors are mainly categorized into two types: direct current (DC) and alternating current (AC) motors. DC motors are utilized in thousands of regular devices used by in our daily life [1]. A DC motor as the name indicates is a motor initiated usually by direct current and is converted into mechanical energy according to the requirement. DC motors are ruling the world due to their extensive use in modern technologies and in almost every industry such as to operate steel rolling mills, electric screw drivers, sewing machines, hard disk drives, air compressors, reciprocating machine etc.[2]. Due to wide range of application, the speed control of DC motor is very important. The speed of DC motor stands various unexpected events such as overshoot, undershoot, peak to peak ripple, steady state error, delay in rise time and settling time, which need to be settled down for which controllers are used. The controllers can be proportional, proportional-integral (PI), proportional derivative (PD), proportional integral derivative (PID), artificial neural network based, Fuzzy Logic based, ANFIS based or any combination of these [3].

Many advancement has been done in Fuzzy PID as well as Fuzzy PI controller which provides better speed performance to the DC motor with more complex design pattern. Fuzzy D controller provides better performance to prevent the overshoot, undershoot, preshoot, ripple and steady state to a considerable amount [3]. In this paper, the FZ-D controller performance has been modified using ANFIS scheme, which enhances its performance over FZ-D controller in the case of overshoot, undershoot, steady state error and peak to peak ripple voltage. ANFIS is the combination of fuzzy logic and neural network. Neural network is supervised learning algorithms which uses a prior experience dataset for predicting future values. In fuzzy logic, the speed control signal of DC motor is generated from the rule base. This rule base is drawn on past experience data which is random. This means that controller's output is also random which may prevent optimal result. But if we use ANFIS, it can make the selection of the rule base more adaptive to the load condition on DC motor[4]. Thus a more robust speed controlling technique of DC motor is proposed in this paper.

II. METHODOLOGY

In this paper, we evaluated the improved method of speed control of a separately excited DC motor by making use of the ANFIS scheme in derivative type controller. This improvement has been ensured by the reduction in rise time, settling time and steady state error provided by the proposed controller. For the evaluation process, we used a step input as the dc input to a motor. We chose error speed and derivative of error speed as the input of ANFIS controller which are multiplexed to provide the input and the output of the controller is a voltage. The output voltage of the controller is added or subtracted to the error voltage depending on the polarity of output voltage of ANFIS controller, which is then applied to the plant. This is done to keep the control variable i.e. speed at a constant value. The same thing is done for Fuzzy logic controller except the fact that controller is provided with Mamdani type Fuzzy inference system with rules while Sugeno type inference system is used for ANFIS controller with training data. The whole process is presented in Fig 1.

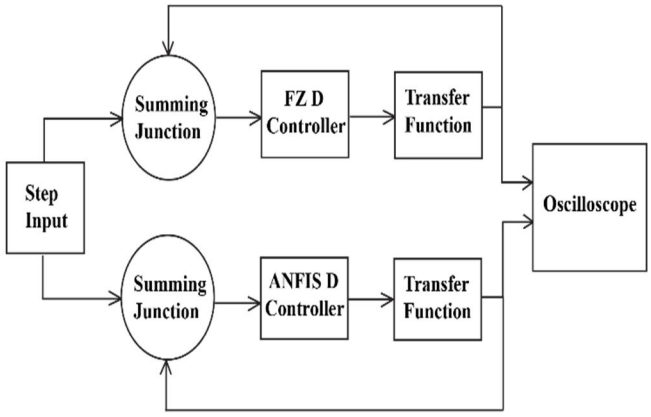


Fig 1. A block diagram employing FZ D and ANFIS D controller.

III. MATHEMATICAL MODELLING

A.Plant Model

DC motor stands a popular actuator in control system engineering. It provides rotational displacement in the direction prescribed by Fleming's left-hand rule and also the translational motion combined with wheels to deliver speed as a mechanical output. Fig 2 represents the equivalent circuit of a DC motor.

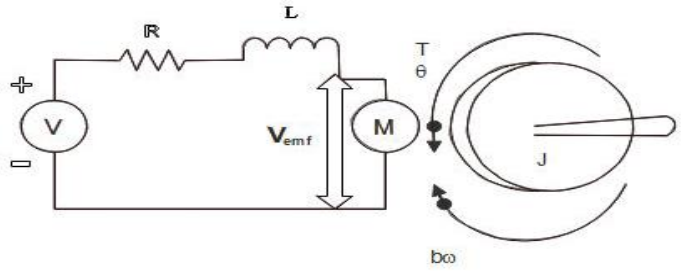


Fig 2. Equivalent circuit of DC motor

Here, V, R and L denotes source voltage (V), electrical resistance(Ω) and inductance(H) of armature circuit and J, b, T,θ, ω, i, V_{emf} denotes moment of inertia of the rotor (kgm), viscous friction coefficient (Nms/rad), motor torque (Nm), angular displacement (rad), output rotating speed (rad/s), armature current (A) and back emf (V) respectively. The mathematical expression of back emf and torque is provided via (1) and (2) respectively.

$$V_{emf}=K_e\omega \quad (1)$$

$$T=K_tI \quad (2)$$

Here, K_e and K_t denotes motor constant armature constant which are essentially equal in SI unit. Equation (1) points out that the back emf is proportional to the output rotating speed and (2) implies that the motor torque is proportional to the armature current. By Newton's law and Kirchhoff's law we obtain (3) and (4).

$$J\frac{d^2\theta}{dt^2}+b\frac{d\theta}{dt}=KI \quad (3)$$

$$L\frac{di}{dt}+Ri=V-K\frac{d\theta}{dt} \quad (4)$$

Here, K indicates electromotive force constant (Nm/A). Using Laplace transform on (3) and (4), (5) and (6) are achieved, and from (5) and (6) the transfer function of the DC motor is presented in (7) is being calculated. Lastly, to obtain its transfer function the value of system parameters of a DC motor is made use of.

$$Js^2\theta(s)+bs\theta(s)=KI(s) \quad (5)$$

$$LsI(s)+RI(s)=V(s)-Ks\theta(s) \quad (6)$$

$$G(s)=\frac{\theta(s)}{V(s)}=\frac{K}{(Ls+R)(Js+b)+K^2} \quad (7)$$

To ensure proper acclimatization of the controlled variable feedback is inveterately fed from the output. To avoid complexity, voltage mode unity feedback control loop is preferred to others.

The DC motor with the parameter is listed below in the table I [1].

TABLE I. VALUES OF DC MOTOR PARAMETERS

Parameter	Value
Armature resistance, R	11.2 Ω
Armature Inductance, L	0.1215H
Rotor inertia, J	0.02215 kgm
Viscous friction coefficient, b	0.002953Nms/rad
Torque constant, K_t	1.28 Nm/A
Back emf constant, K_b	1.28s/rad

B.D Controller Design

The PID is constituted with three different parameters connected in parallel namely the proportional, the integral, and the derivative. The proportional parameter is employed to meliorate the amendatory expedient to the coexisting error of a stable system but it fails to control an unstable system, and steady-state error takes place for the gain stands unchanged for varying frequencies. The integral parameter depreciates the steady state error to zero by constellating the former errors but fails in the regard for loop stability adding a pole at the origin. The derivative parameter presages the coming error and reacts akin to diminish it but fails to deal with constant errors. For a PID controller a weighting mechanism is handy to condition the outcome of the parameters. Equation (8) points the mathematical formula for designing a PID controller [3].

$$U(s)=[K_p+sK_d+\frac{K_i}{s}]E(s) \quad (8)$$

Here, U(S), K_p, K_d, K_i and E(S) denotes controlled variable, proportional gain, derivative gain, integral gain and error value respectively. The error value shows the deviation of the desired input from the actual output.

For a D controller, the equation would be,

$$U(s)=[sK_d]E(s) \quad (9)$$

Basing on rise time, preshoot, overshoot, undershoot, peak to peak ripple, and steady-state error the system performance is being assessed. The controllers can function solo or in conjunction with others. Should the count of parameters in a controller exceeds one, they are linked in parallel.

C.Fuzzy Logic Controller Design

Fuzzy logic allows the inclusion of partial truth to modify two valued logic which then becomes a modification to classical control theory that is more robust and immense [5]. It outreaches the binary logic incorporating the human ideology. It accounts all feasible values encompassing and within "yes" and "no". The fuzzy system consists of four main components;

fuzzification module, knowledge base, inference engine, and defuzzification module. Data is injected into the fuzzification module, and fuzzy sets are formed. In the knowledge base section required rules are put employing "If-Then" rules in general. Inference engine operates on data from fuzzification module based on the depiction of rules. Lastly, defuzzification module appropriates the fuzzy sets to traditional or crisp set.

To design the proposed FZ-D controller, "Fuzzy Logic Controller" block from MATLAB/Simulink is made use of. Input variables "Error" and "ChangeError", along with one output variable "Voltage", are defined through three suitable membership functions each. Utilizing Fuzzy Inference System (FIS) editor in MATLAB/Simulink the design procedure of the fuzzy logic controller, the definition of input and output variables, along the setup of the fuzzy operator, fuzzification module, inference engine, and defuzzification module are executed. The default Mamdani inference and aggregation method is used. A graphic view of the FIS editor.

Every input and output variable defined has seven membership functions; "BN", "MN", "SN", "Zero", "SP", "MP" and "BP". The befitting confines of the variables and apt parameters for the membership functions are picked. Values of the range and the parameters. We defined forty-nine rules making use of If-Then format to design the fuzzy logic controller. The input variables and the entire fuzzy set are respectively the input and output of the rule. Using Rule viewer, we displayed a road map of the fuzzy inference process.

D.ANFIS Controller Design

ANN provides strong learning capabilities by numerical training data while Fuzzy based system ensures better interpretation with the integration of expert level knowledge. The hybridization of these two, is assumed to include advantages of both of the systems. Hence ANFIS scheme comes as a new tool [2]. To improve its performance, the properties of fuzzy logic is tied in this ANFIS technique, by approximating a non-linear system through setting IF-THEN rules [4]. Using a given input/output data set, ANFIS establishes a Fuzzy Inference System (FIS), where the membership function for input and output are adjusted by a back-propagation algorithm or its combination with a least square method called the hybrid method [7, 8].

The ANFIS Editor Graphical User Interface of *MATLAB* window has four areas to support training-

- Load, Plot and Clear the Data
- Generate or Load the Initial FIS Structure
- Train the Fuzzy Inference System
- Validate the Trained Fuzzy Inference System [6]

In the proposed method, we took seven membership function for the two inputs "E" and "CE" which denotes speed error and change of speed error respectively. We chose Gaussian type membership function for the two inputs and there are 7 membership function in each of the inputs which forms 49 conditions for output combining the membership function of the two inputs. The membership function for the two inputs are shown in Fig 3 and Fig 4.

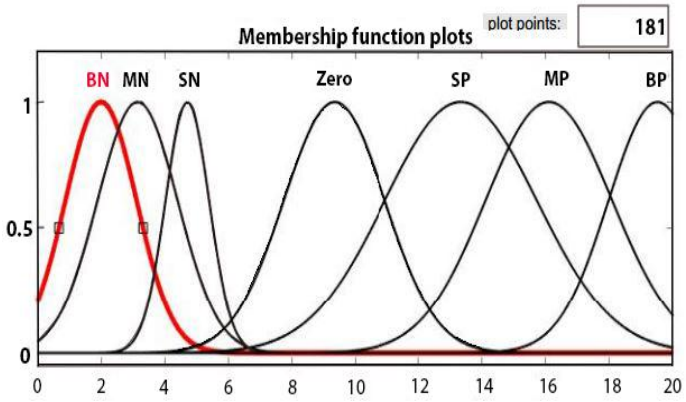


Fig 3. Membership functions of input "E"

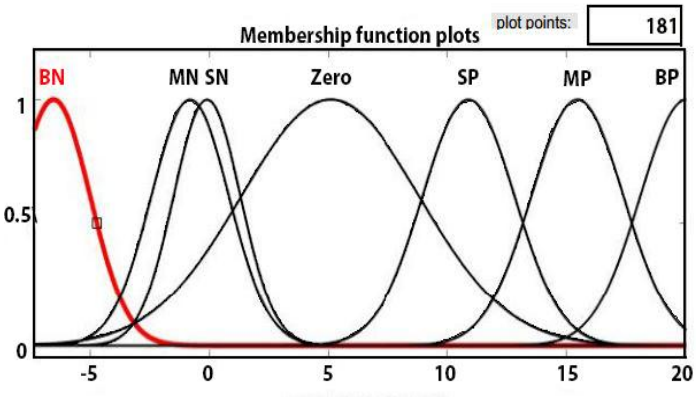


Fig 4. Membership functions of input "CE"

As the Fuzzy Inference system is Sugeno type, the output of each if-then rule a constant which is shown in Fig 5.

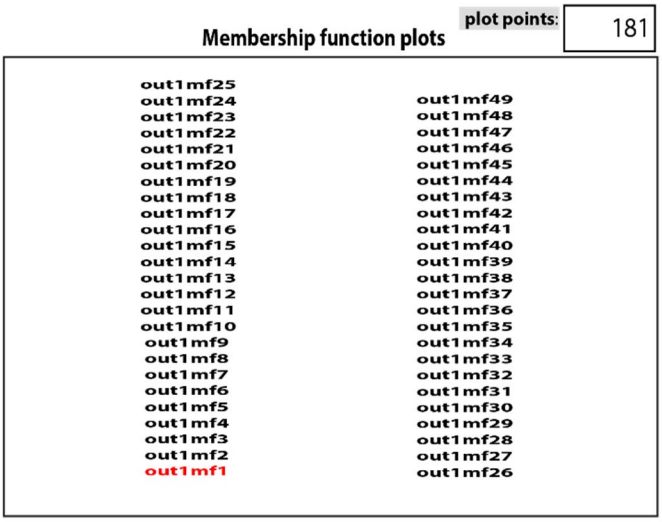


Fig 5. Membership functions of output.

In ANFIS, appropriate training data for inputs and outputs are provided which trains the ANFIS controller to perform in desired manner.

According to the input and output training data, the rules for ANFIS controller is created from the training. The rule base for training data is shown in table II.

TABLE II. RULES FOR ADAPTIVE NEURO FUZZY BASED D CONTROLLER

CE / E	BN	MN	NS	Zero	PS	PM	PB
BN	BN	BN	BN	BN	MN	SN	Zero
MN	BN	BN	BN	MN	SN	Zero	SP
SN	BN	BN	MN	SN	Zero	SP	MP
Zero	BN	MN	SN	Zero	SP	MP	BP
SP	MN	SN	Zero	SP	MP	BP	BP
MP	SN	Zero	SP	MP	BP	BP	BP
BP	Zero	SP	MP	BP	BP	BP	BP

The abbreviations for the membership function stand for

1) BN = Big Negative
2) MN = Medium Negative
3) SN = Small Negative
4) Zero
5) SP = Small Positive
6) MP = Medium Positive
7) BP = Big Positive

IV. SIMULATION

In this section, we evaluated the performance of FZ-D controller and ANFIS-D controller to ascertain the comparative study and performance improvement regarding the rise time, settling time, overshoot and steady state error. A step input is provided to the system which is eventually fed to the D-controller for both Fuzzy and ANFIS. The output of the controller is combined with the error signal to provide an input to dc motor which substantially performs the control mechanism. Thus the speed of DC motor is controlled. The schematics for the proposed speed control mechanism of DC motor is shown in Fig 6.

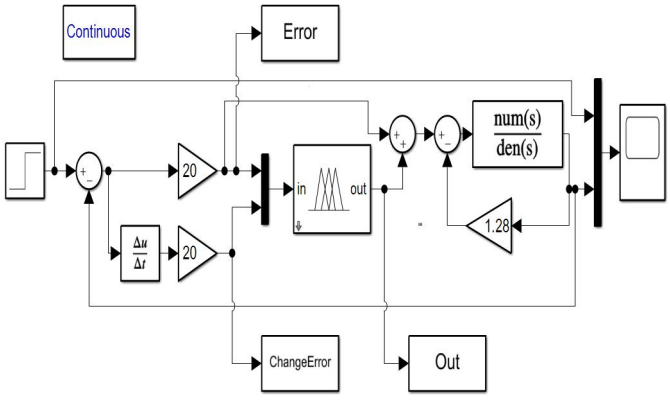


Fig 6. Schematics of speed control of DC motor using FZ D controller.

We have evaluated the performance characteristics for FZ-D controller which is shown in Fig 7.

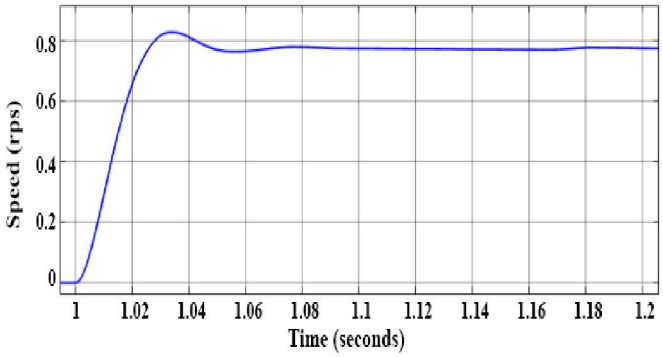


Fig 7. Speed of DC motor for FZ-D controller.

From the data, we have found the that the rise time, preshoot, overshoot, undershoot, peak to peak ripple and steady state error are 17.080ms, 0.513%, 6.989%, 1.972%, 8.273e-01 and 0.463% respectively.

We did similar work for ANFIS-D controller which is shown in Fig 8 to visualize the improvement of performance which is shown in Fig 9.

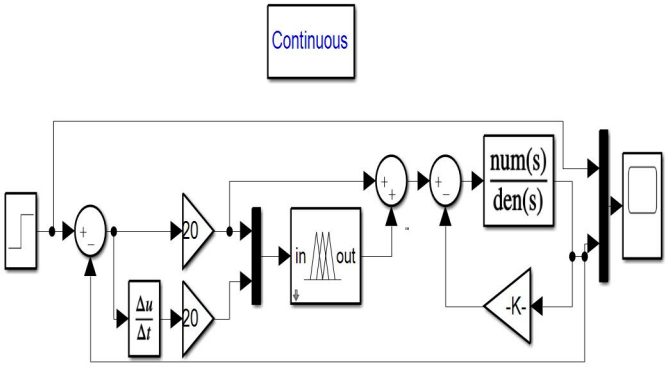


Fig 8. Schematics of speed control of DC motor using ANFIS D controller.

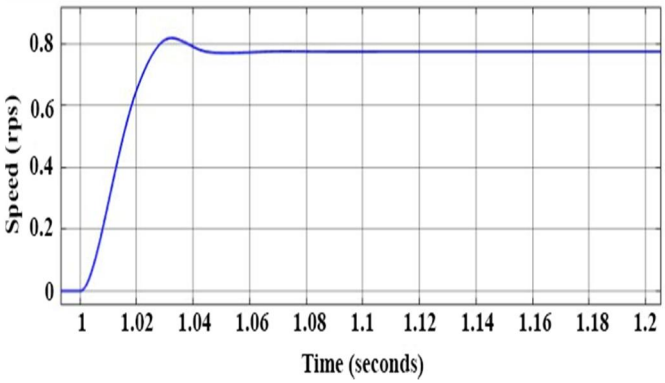


Fig 9. Speed of DC motor for ANFIS D controller.

The rise time, preshoot, overshoot, undershoot and peak to peak ripple and steady state error found in this case are 17.269ms, 0.532%, 5.851%, 1.926%, 8.183e-01 and 0.402% respectively.

V. Outcomes

The result obtained from the simulation has been tabulated in table III. From the comparative evaluation between the data obtained from FZ-D controller and ANFIS-D controller, we demonstrated that, the overshoot, undershoot, peak to peak ripple and steady state error has been reduced in case of ANFIS-D controller that ascertains its better performance compared to FZ-D controller. The negligible increase of rise time and preshoot in ANFIS-D controller has been compensated by its capability to reduce other demerits to a considerable amount.

TABLE III. Speed Controlling Performances of Different Controllers

Controller	FZ-D	ANFIS-D
Rise time	17.080ms	17.269ms
Preshoot	0.513%	0.532%
Overshoot	6.989%	5.851%
Undershoot	1.972%	1.926%
Steady State Error	0.463%	0.402%
Peak to Peak Ripple	8.273e-01	8.183e-01

VI. Conclusion

The main target of this paper is to ensure efficient speed control mechanism for DC motor which suffers from certain unwanted phenomenon. To serve this purpose, FZ-D controller has been modified by ANFIS scheme which has managed to improve its performance in manifold ways. The proposed ANFIS-D controller has caused the increase of rise time and preshoot to a small extent. In spite of these drawbacks, the proposed ANFIS-D controller has managed to reduce the overshoot, undershoot, peak to peak ripple and steady state error successfully and thus it serves the purpose of the proposed method discussed in this paper.

References

[1] Y. Ali Almatheel and A. Abdelrahman "Speed Control of DC Motor Using Fuzzy Logic Controller", *International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE)*, Khartoum, Sudan, vol.-15, pp.215-219, January 2017.

[2] H. Chaudhary, S. Khatoon and R. Singh, "ANFIS Based Speed Control of DC Motor", *2nd International Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH)*, vol.- pp.-63-67, November 2016.

[3] A. S. Ahmed, A. I. Kabir and M. Shahjahan, "Speed control of DC motor with FZ-D controller," *2017 3rd International Conference on Electrical Information and Communication Technology (EICT)*, Khulna, 2017, pp. 1-6, December 2017.

[4] R. Kavathe , J. O. Chandle , N. Patil , M. Kokare,"ANFIS Based Speed Control of BLDC Motor with Bidirectional DC-DC Converter", *International Journal of Research and Scientific Innovation (IJRSI),* vol.-5,pp.153-158, June 2018.

[5] Salim, J. Ohri and Naveen,"Speed Control of DC Motor using Fuzzy Logic based on LabVIEW", *International Journal of Scientific and Research Publications,* Volume 3,Issue 6, June 2013.

[6] Tripura, Pidikiti & Yadlapati, Srinivasa., "Intelligent speed control of DC motor using ANFIS". J*ournal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology*. 26. 223-227. 10.3233/IFS-120729, January 2014.

[7] L.A. Zadeh, Fuzzy Sets, Informal Control (1965), p.~01.8, p.338–353.

[8] L.A. Zadeh, "Outline of a new approach to the analysis complex systems and decision processes", *IEEE Transaction on System, Man and Cybernetics SMC-3(I973), 28–44*.

A New Topology of Input Switched Single Phase AC-DC Buck Boost Converter for High Performance

Khadiza Akter[1], Sairatun Nesa Soheli[2], Md. Nafiul Islam[3]
Department of Electrical and Electronic Engineering[1,2,3]
International University of Business Agriculture and Technology[1,2,3]
Dhaka, Bangladesh
*khadiza@iubat.edu

***Abstract* — In this paper, a new design of bridgeless single phase AC-DC buck boost converter has been proposed to enhance the overall performance of AC-DC converter. Use of single switch and bridgeless topology makes the proposed converter a suitable choice than other conventional AC-DC converters. The conventional AC-DC converters with bridge configuration supply a pulsating ripple signal to the input of the converter circuit which creates lots of disturbance in the overall power system. Due to absence of bridge rectifier such problems do not arise in the proposed converter and it has also reduced using high value filtering element in the input side. Comparisons are shown between proposed and conventional converters. Comparison analysis makes it clear that proposed AC-DC converter performed better than conventional converters. The circuit performances were observed based on software circuit simulation using PSIM 9.1.1.400 professional environments.**

Keywords— AC-DC Buck Boost Converter, Bridgeless Topology, THD, PFC

I. INTRODUCTION

The power quality of the AC system is a vital importance due to the increasing number of newly designed AC-DC converters. Conventional single phase AC-DC converters having full wave rectifiers in bridge configuration always suffer from high total harmonic distortion in the input current and low input power factor. Large input filter like inductor and capacitor can improve THD but power factor still remains low and practical feasibility of such converters is too poor [1, 2]. This problem can be solved by using switch mode regulators. Lots of research works have been done to improve these drawbacks. All techniques are targeting to achieve more precise results [3-7]. Researchers are using different method to improve overall power quality of AC-DC converter. All new methods are providing some advantages than the existing methods [8-11]. Newer research works are going in this field to bring new techniques too. One of the techniques is active PFC [12].

Active filtering requires smaller sized filter than passive filtering, but a DC-DC converter is required after it. The DC-DC converters use unidirectional switch for the switching purpose [13]. In conventional AC-DC converter the bridge rectifier converts AC to uncontrolled DC and the DC-DC converter provides controlled and regulated DC output [14].

In this paper a new design of bridgeless buck boost converter is proposed with high performance. Section II of the papers are discussed about principle of operation of the proposed converter, in section III deals with all the simulation result of the proposed converter circuit. The proposed circuit is advantageous for two reasons. The first is the use of a single switch composed of a unidirectional switch and proper positioning of diodes. The second is the input current switching. The second technique confirms high constant frequency switching of input AC which is almost in phase with input AC voltage without using any feedback control scheme [15-18]. This technique provides smooth input current by using a small filter. It would also results improved power factor with high efficiency.

II. PRINCIPLE OF OPERATION

Fig. 1 below depicts the circuit configuration of a conventional AC-DC buck boost converter.

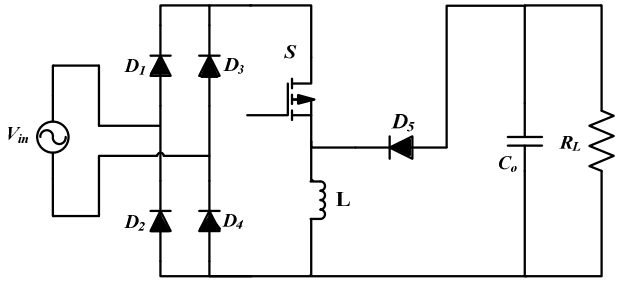


Fig. 1 Conventional AC-DC Buck Boost Converter

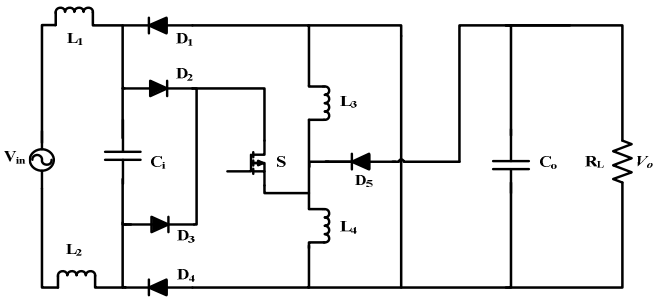


Fig. 2 Proposed Converter Circuit

In conventional buck boost AC-DC converter input resistance of inductor has sudden variety of input current. In the event that the switch S is ON, the inductor takes the energy from the input and it stored the energy of magnetic energy. When switch S is close it releases the energy. The capacitor of the output circuit is accepted as high adequate than the time constant of a RC circuit is high on the output. Primarily inductor L_1 is working as a source of current.

The circuit diagram of the proposed AC-DC converter is given in Fig. 1. This converter has thirteen components; five diodes (D_1, D_2, D_3, D_4 & D_5), two capacitors (C_i & C_o), four inductors (L_1, L_2, L_3 & L_4), one switch S and one load resistor R_L. The supply is given by AC input voltage and

from there can get DC output voltage. The proposed AC-DC buck boost converter can be operated in two conditions–

(i) Continuous conduction mode (CCM)
(ii) Discontinuous conduction mode (DCM)

Switch ON mode for positive half cycle:

When switch S is turned ON for positive half cycle, diode D_1, D_3 & D_5 are turned OFF, because of reverse biased and diode D_2 & D_4 are turned ON due to forward biased condition. In the positive half cycle, the flow of current creates a loop shows in Fig. 2(a).

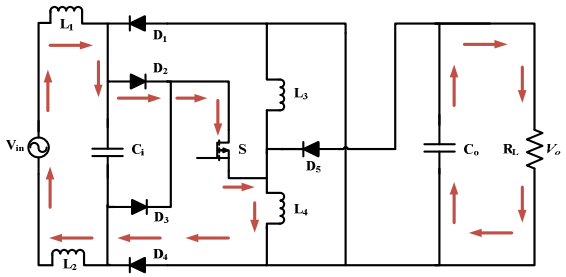


Fig. 2(a) Switch ON Mode (positive half cycle)

In this loop, the AC source V_{in} flows through the diode L_1, D_2 & switch S; connected with the circuit and the inductors L_3, diode D_4 & inductor L_2 for the time duration DT, also current flows to output capacitor C_o & load resistor R_L; from there can get output.

Switch OFF mode for positive half cycle:

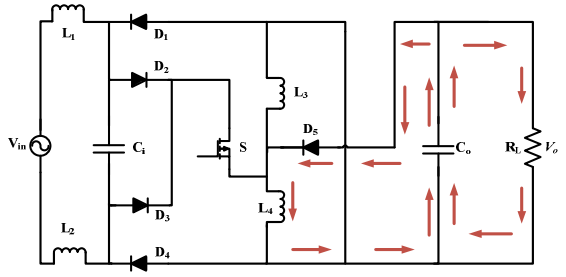


Fig. 2(b) Switch OFF Mode (positive half cycle)

When switch S is turned OFF for positive half cycle, diode D_1, D_2, D_3 & D_4 are turned OFF because of reverse biased. Inductor L_4 released the energy & inductor L_4, output capacitor C_o & diode D_5creates a loop, also current flows to output capacitor C_o & load resistor R_L for the time duration (1-D)T which shows in Fig. 2(b).

Switch ON mode for negative half cycle:

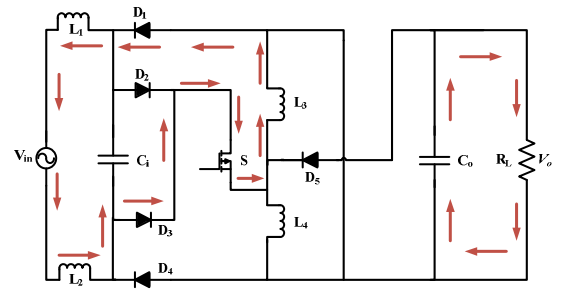


Fig. 2(c) Switch ON Mode (negative half cycle)

When switch S is turned ON in Fig. 2(c) for negative half cycle, diode D_4 & D_5 are turned OFF because of reverse biased and diodeD_1, D_2 & D_4 are turned ON due to forward biased. In the negative half cycle, the flow of current creates a loop. In this loop, the AC source V_{in} flows through the inductor L_2, diode D_3 & switch S; connected with the circuit and the inductor L_3 diode D_1 & inductor L_1 for the time duration DT, also current flows to output capacitor C_o and load resistor R_L; can get output value.

Switch OFF mode for negative half cycle:

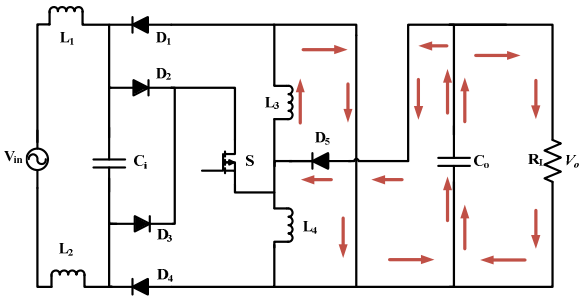


Fig. 2(d) Switch OFF Mode (negative half cycle)

When switch S is turned OFF for negative half cycle, diode D_1, D_2, D_3 & D_4 are turned OFF, because of reverse biased. During this time, Inductor L_3 realized the energy and inductor L_3, output capacitor C_o & diode D_5 creates a loop, also current flows to output capacitor C_o & load resistor R_L; shown in Fig. 2(d).

III. Simulation and Result

The simulation of the two conventional and proposed converter were performed using PSIM 9.1.1 professional version environment and also with Matlab simulink.

A. Values of elements of proposed and conventional circuits:

Simulation results for the proposed and conventional converters are given below to justify the practicability of the proposed circuit. Following values have been used for both of the proposed and conventional converters are given below in Table I.

Table I
Values of the elements used both in proposed and conventional circuit

Parameter	Symbol	Value
Input voltage	V_{in}	220 V, 50 Hz
Switching frequency	F_s	5 kHz
Inductor	L_1 , L_2	9 mH
Inductor	L_3 , L_4	1.5 mH
Capacitor	C_i	5 µF
Capacitor	C_o	220 µF
Load resistor	R_L	10 Ω

B. Open Loop Simulation Result:

For circuit simulation we used PSIM Professional Version 9.1.1.400. After drawing the circuit we used volt probe and current probe both in input and output circuit to view the input and output waveform. Then we set the simulation duration for 0.5 sec and ran it. We observed input voltage versus current and output voltage versus current waveform. PSIM software automatically calculates the power factor and power for voltage versus current both in input and output side. Simulation results of the proposed and conventional circuit in open loop operation are given in Table II.

Table II
Performance Analysis of Proposed and Conventional CUK Converter under Power Factor and Efficiency

Duty Cycle	Conventional AC-DC Buck Boost Converter		Proposed AC-DC Buck Boost Converter	
	Power Factor	Efficiency (%)	Power Factor	Efficiency (%)
0.1	0.29	80.51	0.67	94.57
0.2	0.40	91.11	0.90	98.50
0.3	0.49	95.19	0.96	99.18
0.4	0.58	97.28	0.96	99.22
0.5	0.66	98.41	0.87	98.95
0.6	0.74	98.95	0.73	98.48
0.7	0.74	99.01	0.50	97.42
0.8	0.80	98.78	0.25	94.95
0.9	0.85	96.93	0.09	85.96

Comparison analysis for the power factor and efficiency of the proposed and conventional circuit is given by the bar diagram in Fig. 3 and Fig. 4 respectively whereas Fig.5 shows the THD comparison. From graphical comparison it is visible that proposed circuit is showing better performance in all parameters in duty cycle variation. The conventional and proposed AC-DC converters are not performing well in extremely high and low duty cycle due to non-ideal characteristics MOSFET switch.

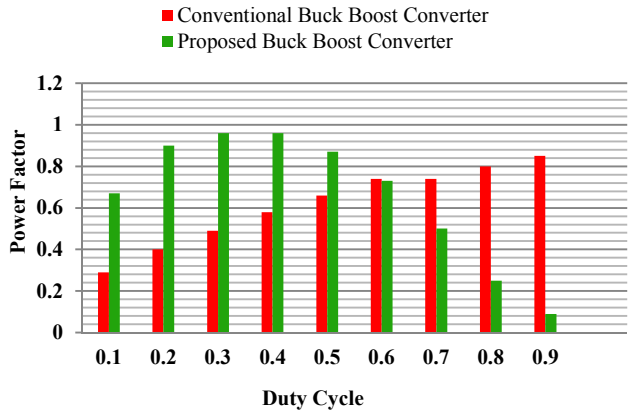


Fig. 3 Comparison of the Performance of Power Factor between Conventional and Proposed Converter with Duty Cycle Variation

Moreover, Table III shows the comparative analysis of the AC-DC buck converter for THD under duty cycle variation.

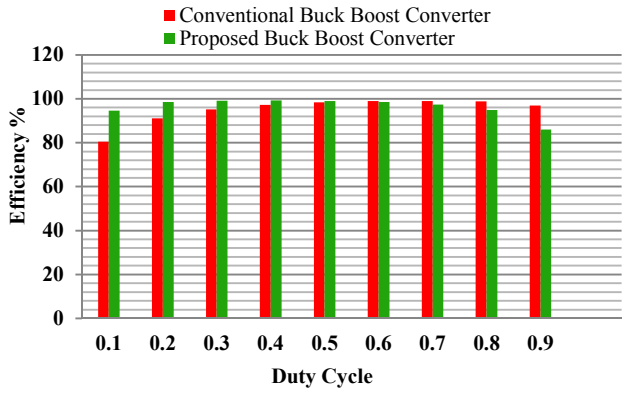


Fig. 4 Comparison of the Performance of Efficiency between Conventional and Proposed Converters with Duty Cycle

Both proposed and conventional AC-DC buck boost converters kept fixed at 5 kHz for switching frequency and the load variation is fixed at 10 Ω throughout the duty cycle 0.1 to 0.9. It is clear from the bar diagram of Fig. 4 that the efficiency of the proposed AC-DC buck boost converter is much better for maximum the range of efficiency than conventional AC-DC buck boost converter.

Table III
Performance analysis of proposed and conventional cuk converter under duty cycle and THD

Duty Cycle	THD	
	Conventional AC-DC Buck Boost Converter	Proposed AC-DC Buck Boost Converter
0.1	0.29	0.22
0.2	2.20	0.32
0.3	1.70	0.28
0.4	1.38	0.18
0.5	1.11	0.06
0.6	0.89	0.04
0.7	0.72	0.02
0.8	0.70	0.009
0.9	0.61	0.007

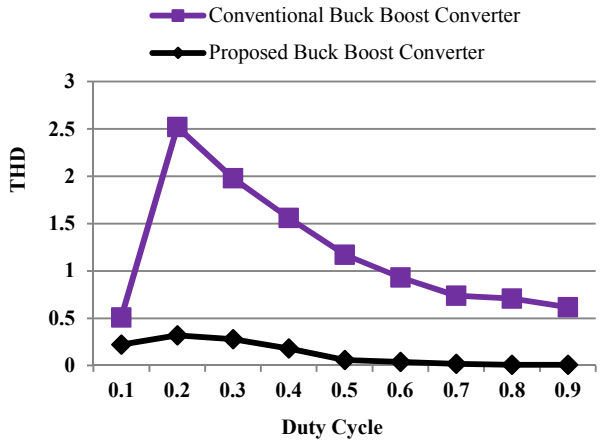


Fig. 5 THD comparison of Conventional and Proposed AC-DC Buck Boost Converters, throughout the Duty Cycle of 0.1 to 0.9

Furthermore, Fig. 6 shows the input current and voltage waveform of the proposed AC-DC buck boost converter, it is clear from the given wave diagram that the input voltage and current are not having any phase difference that means the power factor is almost unity at 40 to 50 percent of duty cycle variation which is mentioned in comparison Table II.

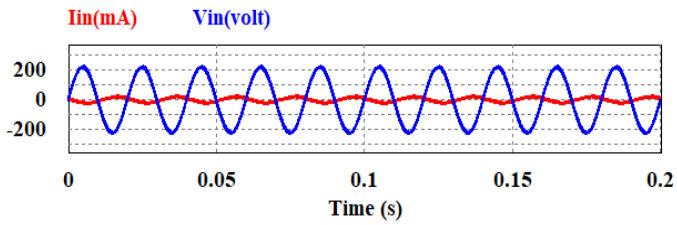


Fig. 6 Input Current and Voltage Waveform of Proposed AC-DC Converter

In the wave diagram the blue wave is representing the input voltage where the rms value of the voltage is 155 V whereas the red one is representing the input current holding the rms value of 14 mA.

IV. CONCLUSION

The proposed circuit significantly improves the performance of all parameter. It is evident from the performance analysis the proposed buck boost Converter without feedback controller scheme outperforms all the other converter circuits based on the individual simulation results and achieves superior power quality not only in terms of power factor and efficiency, but also the input current THD value has been brought to 0.02 which is within IEEE prescribed limit. The proposed Buck Boost converter circuit offer 99% conversion efficiency Thus, this converter circuit holds promise for future which can be practically realized through hardware implementations.

References

[1] H.-S. Cho, *et al.*, "Bridgeless half-bridge AC-DC converter with series-connected two transformers," in *2013 Twenty-Eighth Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*, 2013, pp. 3241-3245.

[2] M. Arias, *et al.*, "High-efficiency asymmetrical half-bridge converter without electrolytic capacitor for low-output-voltage AC–DC LED drivers," *IEEE transactions on Power Electronics,* vol. 28, pp. 2539-2550, 2013.

[3] M. S. Arifin and M. J. Alam, "Input switched high performance single phase single switch Ćuk AC-DC converter," in *2015 International Conference on Advances in Electrical Engineering (ICAEE)*, 2015, pp. 226-229.

[4] M. M. S. Khan, *et al.*, "Input switched high performance three phase Buck-Boost controlled rectifier," in *2013 IEEE International Conference on Industrial Technology (ICIT)*, 2013, pp. 557-562.

[5] M. R. T. Hossain, *et al.*, "A single phase SEPIC AC-DC converter with improved power factor and input current THD," in *2016 9th International Conference on Electrical and Computer Engineering (ICECE)*, 2016, pp. 373-376.

[6] U. Kamnarn and V. Chunkag, "Analysis and design of a modular three-phase AC-to-DC converter using CUK rectifier module with nearly unity power factor and fast dynamic response," *IEEE transactions on Power Electronics,* vol. 24, pp. 2000-2012, 2009.

[7] M. I. Hossain and M. J. Alam, "Cuk topology based power factor correction and output voltage regulation of AC-DC converter," in *2014 International Conference on Electrical Engineering and Information & Communication Technology*, 2014, pp. 1-6.

[8] J.-H. Kim, *et al.*, "Boost integrated flyback AC-DC converter with valley fill circuit for LED light bulb," in *Proceedings of The 7th International Power Electronics and Motion Control Conference*, 2012, pp. 457-462.

[9] M. Mahmood, *et al.*, "Isolated dual boost bridgeless power factor correction AC-DC converter," in *2013 1st International Future Energy Electronics Conference (IFEEC)*, 2013, pp. 465-470.

[10] G. Spiazzi and P. Mattavelli, "Design criteria for power factor preregulators based on SEPIC and Cuk converters in continuous conduction mode," in *Proceedings of 1994 IEEE Industry Applications Society Annual Meeting*, 1994, pp. 1084-1089.

[11] M. S. Arifin and M. J. Alam, "Input switched single phase SEPIC controlled rectifier with improved performances," in *2016 9th International Conference on Electrical and Computer Engineering (ICECE)*, 2016, pp. 38-41.

[12] N. A. Ahmed, "Modeling and simulation of ac–dc buck-boost converter fed dc motor with uniform PWM technique," *Electric power systems research,* vol. 73, pp. 363-372, 2005.

[13] B. Axelrod, *et al.*, "Switched-capacitor/switched-inductor structures for getting transformerless hybrid DC–DC PWM converters," *IEEE Transactions on Circuits and Systems I: Regular Papers,* vol. 55, pp. 687-696, 2008.

[14] Q. Lu, *et al.*, "DC-DC converters," ed: Google Patents, 1997.

[15] B. Singh, *et al.*, "A review of three-phase improved power quality AC-DC converters," *IEEE Transactions on Industrial Electronics,* vol. 51, pp. 641-660, 2004.

[16] M. A. Kabir, *et al.*, "Boost and Buck topology based single phase AC-DC converters with low THD and high power factor," in *2011 IEEE 33rd International Telecommunications Energy Conference (INTELEC)*, 2011, pp. 1-7.

[17] K. Akter, *et al.*, "Close Loop Analysis of a Single Phase High-Efficiency AC-DC CUK Converter with Low Input Current THD and Improved Power Factor," in *2018 10th International Conference on Electrical and Computer Engineering (ICECE)*, 2018, pp. 393-396.

[18] K. Akter, *et al.*, "Modeling and Simulation of Input Switched AC-DC SEPIC Converter with PFC Control for optimized Operation," in *2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT)*, 2018, pp. 240-245.

A Novel 36-Pulse Star Rectifier with Passive Injection Circuits at DC Link

Md. Abdul Malek
Department of Electrical and Electronic Engineering
Rajshahi University of Engineering & Technology (RUET)
Rajshahi-6204, Bangladesh
m.a.malekb88@gmail.com

Md. Abu Bakar Siddik
Department of Electrical and Electronic Engineering
Rajshahi University of Engineering & Technology (RUET)
Rajshahi-6204, Bangladesh
abubakarsiddik24@gmail.com

Md. Shoaib Akther
Department of Electrical and Electronic Engineering
Rajshahi University of Engineering & Technology (RUET)
Rajshahi-6204, Bangladesh
shoaibkiron@gmail.com

***Abstract*—This paper presents a new 36-pulse star connected diode rectifier. It takes intimate sinusoidal currents from utility with the elimination of lower than 35th order harmonics. The proposed rectifier consists of a parallel four-star 12-pulse diode rectifier and two passive injection circuits at DC links. This passive injection circuit can triple the number of pulses of parallel four-star 12 pulse rectifier system. The primary side of specially tapped interphase transformer (STIT) with two diodes constitutes the first injection circuit while the secondary winding of the STIT with a single phase full-wave rectifier (SPFR) constitutes the second passive injection circuit. This rectifier is compatible for large current rating application such as electroplating, plasma torches, dc arc furnaces etc. The simplicity of the proposed rectifier scheme makes it more economical compared to other existing designs of same performance. The proposed rectifier is analyzed and simulated in MATLAB environment. The simulation result presented in this paper confirms the IEEE-519 standards requirements at varying loads.**

Keywords—specially tapped interphase transformer (STIT), single phase full-wave rectifier (SPFR), total harmonic distortion (THD), passive injection circuit.

I. INTRODUCTION

Star rectifier is used for low voltage and high current applications that require minimum voltage drop due to source impedance, leakage reactance, source impedance and devices used in the path of load current. The star rectifiers have one device in the path of load current and preferred in large current rating applications such as induction heating, electrochemical processes, plasma torches, dc arc furnaces etc [1]. Unfortunately, rectifier affects the power quality by injecting harmonics into AC mains and can not meet IEEE-519 standard requirements [2]. Harmonics increase error of the measuring instruments, increase the losses, disrupt control systems, affect the torque of the motors, interfere with communication systems, degrade the operating characteristics of the protective devices and may cause undesirable resonances. To minimize the unwanted effects of harmonic distortion, IEEE Std. 519-1992 suggests that an individual harmonic distortion of voltage to be less than 3% of the fundamental component and the total harmonic distortion (THD) to be less than 5% of the fundamental component.

Passive filters and active filters have been used in the rectifiers to reduce the input line current harmonics [3]-[4]. But passive filter is bulky and susceptible to load switching transients while active filter suffers from complexity and high cost. Multi-pulse rectifiers are reliable for input line current harmonics mitigation. 12-pulse, 18-pulse and 24-pulse star rectifiers can not meet the IEEE-519 standard requirements [5]. More than 24-pulse rectifiers are required to meet the IEEE-519 standard requirements [6]-[7]. But these rectifiers require complex phase shifting transformer and increase complexity. In [8]-[9], active injection techniques have been proposed for rectifier. Active injection technique can reduce the input line current harmonic but it can not improve the output power quality.

One way of reducing the input line current harmonics and increasing the pulse number of the rectifier simultaneously without increasing significant complexity of the rectifier circuit is through the use of passive injection technique at the DC link [10]-[12]. In [10], pulse tripling technique has been proposed for parallel 12-pulse diode bridge rectifier and it can meet the IEEE-519 standard requirements. In [11], pulse doubling technique has been proposed for parallel four-star 12-pulse diode rectifier but it fails to meet the IEEE-519 standard requirements. In [12], a passive harmonic reduction method has been proposed for the double-star 6-pulse diode rectifier which fails to meet the IEEE-519 standard requirements.

This paper presents a new 36-pulse star connected diode rectifier that is based on the passive injection technique at DC link. In the proposed rectifier, STIT and its auxiliary circuits constitute the two passive injection circuits. When tap position and turns ratio of the STIT is designed optimally, the proposed rectifier operates as a 36-pulse rectifier. It eliminates 5th, 7th, 11th, 13th, 17th, 19th, 23rd, 24th, 29th and 31th order harmonics from utility line currents**.** After using passive injection circuits, THD of input line current is about 4.82%.

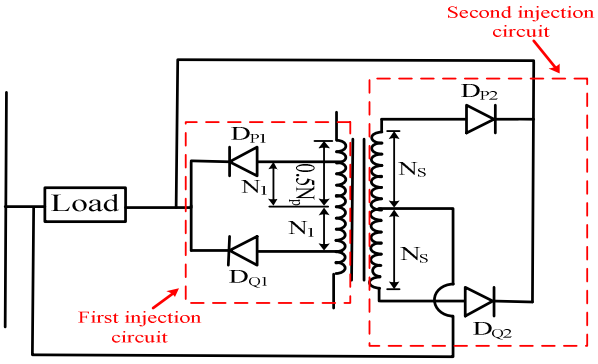


Fig. 1: Passive injection circuit.

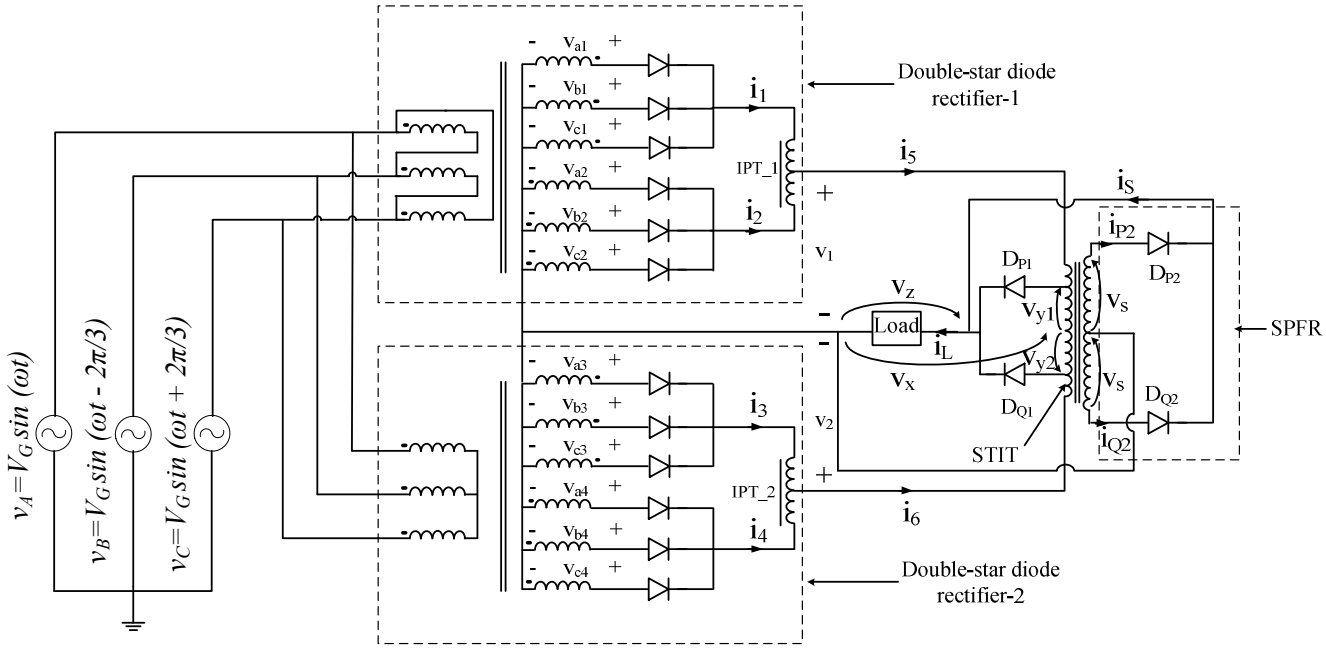


Fig. 2: Proposed 36-pulse star rectifier

II. Proposed 36-Pulse Scheme

A. Circuit Configuration

Fig. 2 shows the proposed 36-pulse uncontrolled star rectifier. It is similar to the conventional parallel four-star 12-pulse rectifier with the addition of two passive injection circuits at dc link. The proposed rectifier configuration uses delta-wye-wye and wye-wye-wye connected transformer for phase shifting from three-phase source to feed four three-phase star rectifier. The primary winding of the STIT and two diodes D_{P1} and D_{Q1} constitute the first injection circuit which is shown in Fig. 1. The primary winding of the STIT has two tap which are connected to the load through diodes D_{P1} and D_{Q1}. Taps on the primary winding of STIT are chosen such position that proposed rectifier produces asymmetrical 24-pulse output voltage across the load without second passive injection circuit which is shown in Fig. 4(d). The center-tapped secondary winding of the STIT with two diodes D_{P2} and D_{Q2} constitute the SPFR which is the second passive injection circuit. Output terminals of the SPFR are connected in parallel to the load such that the harmonic power drawn is being fed back to load thus improving the energy conversion efficiency. This second passive injection circuit extends asymmetrical 24-pulse operation to 36-pulse operation. Interphase transformers (IPT) are used in the proposed rectifier circuit to ensure the parallel operation of the star rectifier.

Following assumptions are taken into consideration.

N_p= Number of turns in primary winding of the STIT.

N_s=Number of turns in secondary winding of the STIT.

N_1=Number of turns between midpoint and tapped points of primary winding of the STIT.

m = Turns ratio = $\frac{N_s}{N_p}$

k = Tap ratio = $\frac{N_1}{N_p}$

t:1:1 = Turns ratio of the $\Delta - Y - Y$ transformer.

t: $\sqrt{3}$: $\sqrt{3}$ = Turns ratio of the $Y - Y - Y$ transformer.

B. Circuit Operation

Assume that load voltage is v_z, voltage across the primary winding of the STIT is v_m, voltage across the secondary winding is v_s and load current is i_L. According to relation among v_z, v_m and v_s, the proposed rectifier has four operation modes. Fig. 3 shows the operation modes of the proposed rectifier.

P1-Mode: Whenever $v_m > 0$ and $v_s < v_z$, the proposed rectifier operates under P1-mode which is shown in Fig. 3(a). In this mode diode D_{P1} is forward-biased and diode D_{Q1}, D_{P2} and D_{Q2} are reversed-biased. SPFR does not work and the output current i_{P2} and i_{Q2} are equal to zero. Double-star diode rectifier-1 and double-star diode rectifier-2 conduct and supply current i_5 and i_6 to the load respectively. Diode D_{P1} carries load current i_L. The MMF relationship of the STIT for the P1-mode is

$$i_5(0.5 - k)N_p = i_6(0.5 + k)N_p \quad (1)$$

Output current of the rectifier is

$$i_L = i_5 + i_6 \quad (2)$$

From (1) and (2), the output currents of two double-star rectifiers are obtained as

$$\begin{cases} i_5 = (0.5 + k)i_L \\ i_6 = (0.5 - k)i_L \end{cases} \quad (3)$$

Current of individual star rectifier is

$$\begin{cases} i_1 = i_2 = \frac{i_5}{2} \\ i_3 = i_4 = \frac{i_6}{2} \end{cases} \quad (4)$$

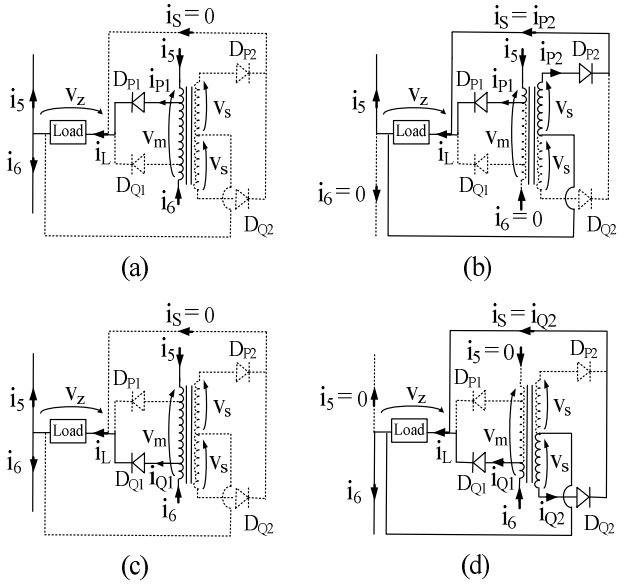


Fig. 3: (a) P1-mode (b) P2-mode (c) Q1-mode (d) Q2-mode.

P2- Mode: Whenever $v_m > 0$ and $v_s > v_z$, the proposed rectifier operates under P2-mode which is shown in Fig. 3(b). In this mode diode D_{P1} and D_{P2} are forward-biased and diode D_{Q1} and D_{Q2} are reversed-biased. SPFR works and current i_{P2} is injected to the load. The double-star diode rectifier-1 conducts and the double-star diode rectifier-2 does not conduct. Diode D_{P1} and D_{P2} carry load current i_L and relation among i_5, i_{P2}, i_S and i_L is

$$i_5 + i_{P2} = i_5 + i_S = i_L \tag{5}$$

The MMF relationship of the STIT for the P2-mode is

$$i_5\left(0.5N_p - kN_p\right) = i_{P2}N_s \tag{6}$$

From (5) and (6), the currents i_5, i_{P2}, and i_S are obtained as

$$\begin{cases} i_5 = \frac{m}{0.5-k+m} i_L \\ i_{P2} = i_S = \frac{0.5-k}{0.5-k+m} i_L \end{cases} \tag{7}$$

Current of individual star rectifier is

$$\begin{cases} i_1 = i_2 = \dfrac{i_5}{2} \\ i_3 = i_4 = 0 \end{cases} \tag{8}$$

Q1-Mode: Whenever $v_m < 0$ and $|- v_s| < v_z$, the proposed rectifier operated under Q1-mode which is shown in Fig. 3(c). In this mode diode D_{Q1} is forward-biased and diode D_{P1}, D_{P2} and D_{Q2} are reversed-biased. SPFR does not work and the output current i_{P2} and i_{Q2} are equal to zero. Double-star diode rectifier-1 and double-star diode rectifier-2 conduct and supply current i_5 and i_6 to the load respectively. Diode D_{Q1} carries load current i_L. The MMF relationship of the STIT for the Q1-mode is

$$i_5(0.5 + k)N_p = i_6(0.5 - k)N_p \tag{9}$$

Output current of the rectifier is

$$i_L = i_5 + i_6 \tag{10}$$

From (9) and (10), the output currents of the two double-star rectifier diode rectifiers are obtained as

$$\begin{cases} i_5 = (0.5 - k)i_L \\ i_6 = (0.5 + k)i_L \end{cases} \tag{11}$$

Current of individual star rectifier is

$$\begin{cases} i_1 = i_2 = \frac{i_5}{2} \\ i_3 = i_4 = \frac{i_6}{2} \end{cases} \tag{12}$$

Q2-Mode: Whenever $v_m < 0$ and $|- v_s| > v_z$, the proposed rectifier operates under Q2-mode which is shown in Fig. 3(d). In this mode diode D_{Q1} and D_{Q2} are forward-biased and diode D_{P1} and D_{P2} are reversed-biased. SPFR works and current i_{Q2} is injected to the load. The double-star diode rectifier-1 does not conduct and the double-star diode rectifier-2 conducts and supply current i_6 to the load. Diode D_{Q1} and D_{Q2} carry load current i_L and relation among i_6, i_{Q2}, i_S and i_L is

$$i_6 + i_{Q2} = i_6 + i_S = i_L \tag{13}$$

The MMF relationship of the STIT for the Q2-mode is

$$i_6\left(0.5N_p - kN_p\right) = i_{Q2}N_s \tag{14}$$

From (13) and (14), the currents i_6, i_{Q2} and i_S are obtained as

$$\begin{cases} i_6 = \frac{m}{0.5-k+m} i_L \\ i_{Q2} = i_S = \frac{0.5-k}{0.5-k+m} i_L \end{cases} \tag{15}$$

Current of individual star rectifier is

$$\begin{cases} i_1 = i_2 = 0 \\ i_3 = i_4 = \frac{i_6}{2} \end{cases} \tag{16}$$

III. Optimal Tap Ratio And Optimal Turns Ratio

A. Optimal Tap Ratio of the STIT

To find the optimal tap of the STIT, assume that second passive injection circuit is not connected to the proposed system and proposed rectifier produces asymmetrical 24-pulse output voltage across the load which is shown in Fig. 4(d). Consider that voltage across double-star rectifier-1 is v_1, voltage across double-star rectifier-2 is v_2, voltage across primary winding of STIT is v_m and variable voltage $\overrightarrow{V_y}$ is established between the taps point and the center point of the primary winding of the STIT. Variable voltage $\overrightarrow{V_x}$ is added to $\overrightarrow{V_y}$ and produces asymmetrical 24-pulses output voltage ($\overrightarrow{V_z}$). These voltage wave shapes are shown in the Fig. 4 and relations among them are

$$v_m = v_1 - v_2 \tag{17}$$

$$v_x = \frac{v_1 + v_2}{2} \tag{18}$$

$$v_y = v_{y1} = \frac{N_1}{N_p} v_m \text{ ,When diode } D_{P1} \text{ conducts} \tag{19}$$

$$v_z = v_{z1} = v_x + v_{y1}, \text{ When diode } D_{P1} \text{ conducts} \tag{20}$$

$$v_y = v_{y2} = -\frac{N_1}{N_p} v_m \text{ , When diode } D_{Q1} \text{ conducts} \tag{21}$$

$$v_z = v_{z2} = v_x + v_{y2}, \text{ When diode } D_{Q1} \text{ conducts} \tag{22}$$

$$\overrightarrow{V_z} = \overrightarrow{V_x} + \overrightarrow{V_y} \tag{23}$$

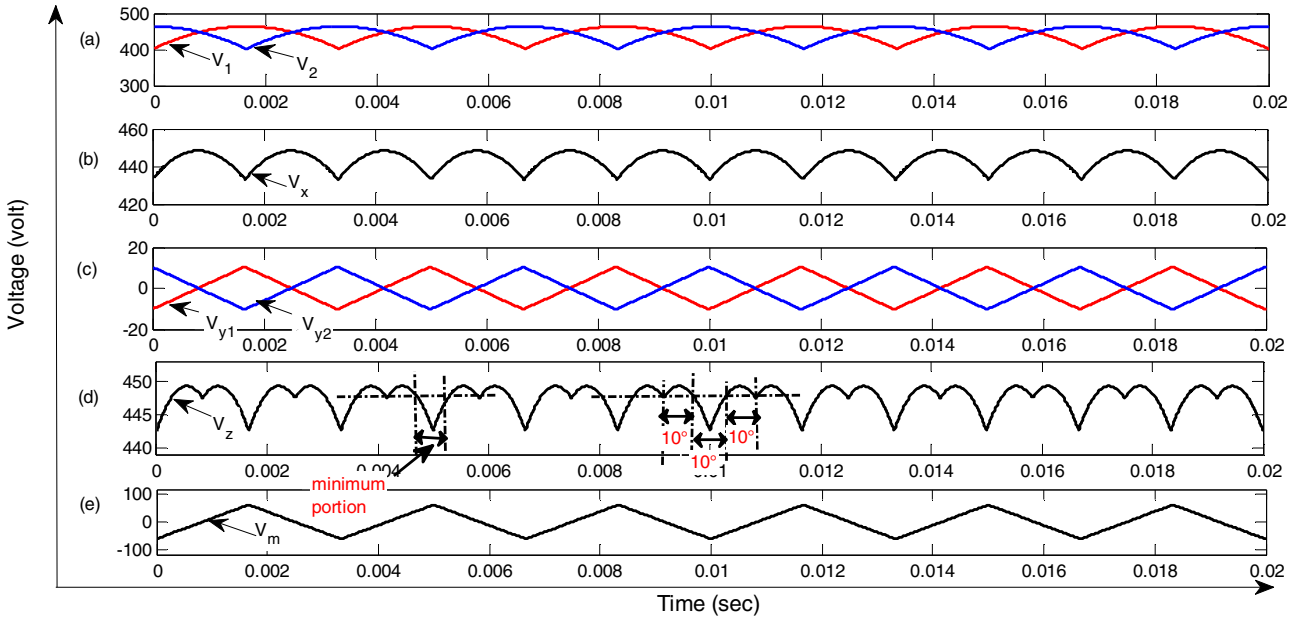


Fig. 4: Waveforms of (a) Output voltage of the rectifier-1 and rectifier-2 (b) Voltage v_x (c) Voltages v_{y1} and v_{y2} (d) Voltage v_z (e) Voltage v_m.

Tap ratio of the STIT is chosen such value that the output voltage v_z can be divided into 36 portions per power supply cycle and conduction period of each portion is 10° which is shown in the Fig. 4(d). The phasor diagram of Fig. 5 presents the relationship between voltage waveforms contained in Fig. 4. The following expression can be written from the Fig. 5.

$$\frac{N_1}{N_p} = \frac{V_y}{V_m} = \frac{V_1 cos\frac{\theta}{2}\, tan\frac{\theta}{6}}{2\, V_1\, sin\frac{\theta}{2}} \tag{24}$$

Output voltages of the two double-star 12-pulse rectifier are displaced by $30°$. Optimal tap ratio is obtained by substituting $\theta = 30°$ in (24).

$$k = \frac{N_1}{N_p} = 0.163 \tag{25}$$

B. Optimal Turns Ratio of the STIT

Output voltage of the double-star diode rectifier-1 and the double-star diode rectifier-2 without second passive injection circuit can be expressed as

$$v_1 = \frac{max(v_{a1}, v_{b1}, v_{c1}) + max(v_{a2}, v_{b2}, v_{c2})}{2} \tag{26}$$

$$v_2 = \frac{max(v_{a3}, v_{b3}, v_{c3}) + max(v_{a4}, v_{b4}, v_{c4})}{2} \tag{27}$$

Where v_{a1}, v_{b1}, v_{c1}, v_{a2}, v_{b2} and v_{c2} are the secondary side phase voltages of the $\Delta - Y - Y$ transformer and v_{a3}, v_{b3}, v_{c3}, v_{a4}, v_{b4}, v_{c4} are the secondary side phase voltages of the $Y - Y - Y$ transformer. The voltage across the load can be represent by the following expression

$$v_z = v_x + 0.163\,|v_m| \tag{28}$$

Expression of the voltages v_m and v_x are obtained by substituting (26) and (27) into (17) and (18). Expression of the voltages v_m and v_x are

$$v_m = \begin{cases} -\frac{3(\sqrt{6}-\sqrt{2})\,V_G}{4t}\, cos\left(\omega t + \frac{5\pi}{12}\right), & \omega t \,\epsilon \left[0, \frac{\pi}{6}\right] \\ \frac{3(\sqrt{6}-\sqrt{2})\,V_G}{4t}\, cos\left(\omega t + \frac{\pi}{4}\right), & \omega t \,\epsilon \left[\frac{\pi}{6}, \frac{\pi}{3}\right] \end{cases} \tag{29}$$

$$v_x = \begin{cases} \frac{3(\sqrt{6}+\sqrt{2})\,V_G}{8t}\, sin\left(\omega t + \frac{5\pi}{12}\right), & \omega t \,\epsilon \left[0, \frac{\pi}{6}\right] \\ \frac{3(\sqrt{6}+\sqrt{2})\,V_G}{8t}\, sin\left(\omega t + \frac{\pi}{4}\right), & \omega t \,\epsilon \left[\frac{\pi}{6}, \frac{\pi}{3}\right] \end{cases} \tag{30}$$

Now considering that second passive injection circuit is connected to the proposed rectifier system. The second passive injection circuit injects the harmonic current into the load in the 'minimum portion' of the asymmetrical 24-pulse output voltage which is shown in Fig. 4(d). The maximum amount of injected current by SPFR in P2-mode or Q2-mode must be equal to the maximum load current in P1-mode or Q1-mode minus the maximum current supplied by double-star rectifier-1 or double-star rectifier-2 in P2-mode or Q2-mode respectively. When the SPFR works, proposed rectifier operates in P2-mode or Q2-mode and only one double-star rectifier supplies the load current. In Q2-mode, the current i_6 is supplied by the rectifier-2 and expression of i_6 is

$$i_6 = \frac{P}{v_2} \tag{31}$$

Where P is the KVA rating of the $\Delta - Y - Y$ and $Y - Y - Y$ transformer. Assume that proposed rectifier operate under rated load condition. The expression of i_6 is obtained by substituting (27) into (31).

$$i_6 = \frac{Pt}{1.5\, V_G cos\, \omega t} \quad \omega t \,\epsilon\, [0, \frac{\pi}{36}] \tag{32}$$

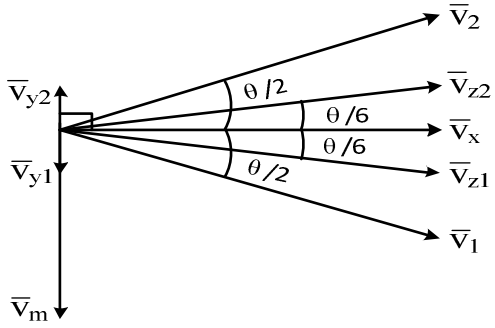


Fig. 5: Phasor diagram.

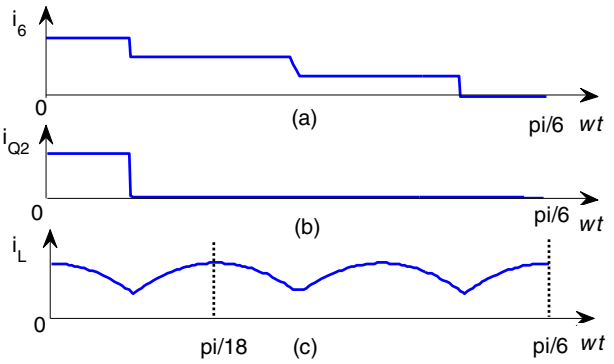


Fig. 6: Waveforms of (a) Current i_6 (b) Current i_{Q2} and (c) Current i_L.

Where V_G is amplitude of input phase voltages. Fig. 6(a) shows the maximum value of current i_6 is obtained at $\omega t = 0$ and expression is

$$i_{6max} = \frac{Pt}{1.5V_G} \tag{33}$$

The load current i_L in P1-mode and Q1-mode is

$$i_L = \frac{P}{v_z} \tag{34}$$

Fig. 6(c) shows that the maximum value of load current i_L is obtained in Q1-mode at $\omega t = \frac{\pi}{18}$ and expression is obtained by substituting (28) to (30) into (34). The maximum value of load current in P1-mode and Q1-mode is

$$i_{Lmax} = \frac{Pt}{1.4544\,V_G} \tag{35}$$

The maximum value of current should be injected by the SPFR in P2-mode and Q2-mode is

$$i_{Q2max} = i_{Lmax} - i_{6max} = \left(\frac{Pt}{1.4544\,V_G} - \frac{Pt}{1.5V_G}\right) \tag{36}$$

The optimal turns ratio is obtained by substituting (33) and (36) into (14) and this value is

$$m = \frac{N_s}{N_p} = 10.75 \tag{37}$$

IV. Simulation

MATLAB software is used for simulation to verify the theoretical analysis of the proposed rectifier. Following specifications are taken into consideration for simulation purpose.

TABLE I. SYSTEM SPECIFICATION

Parameter	Value
Input line-to-line voltage (rms)	400 V
Frequency	50 Hz
Turns ratio of the $\Delta - Y - Y$ transformer	1 : 1 : 1
Turns ratio of the $Y - Y - Y$ transformer	1: √3 : √3
Turns ratio of the STIT	1 : 10.75
Rated output power	50 KW

V. Results

Fig. 7 and Fig. 8 illustrate input line current and spectrum of the input line current respectively. It is concluded that input current has 36 steps approximately per cycle and THD of input line current is about 4.82% under rated load condition. When passive injection circuits are not used in the four-star diode rectifier, THD of the input line current is about 15.11%. TABLE II shows comparison of the input line current THD of the proposed rectifier and conventional four-star 12-pulse diode rectifier. It is concluded from the table that THD of the input line current is within the IEEE-519 standard requirements for wide load variation condition. Fig. 9 shows output voltage which contains 36 pulses per cycle of input voltage. The waveforms of current i_{P1} and i_{Q1} that flow through the diodes D_{P1} and D_{Q1} are reported in Fig. 10 (a) and 10 (b) respectively. Fig. 11(a) and 11(b) show the waveforms of current i_{P2} and i_{Q2} that flow through the diodes D_{P2} and D_{Q2} respectively. The graphical representation of output current of double-star rectifier-1 and double-star rectifier-2 are presented in Fig. 12(a) and 12(b) respectively.

TABLE II. PERFORMACE COMPARISON

% Rated Load	THD of the Input Line Current for Proposed Rectifier (%)	THD of the Input Line Current for Conventional 12-Pulse Star Rectifier (%)
30	4.97	14.70
40	4.94	14.81
50	4.88	14.90
60	4.83	14.96
70	4.83	15.02
80	4.82	15.07
90	4.82	15.07
100	4.82	15.11

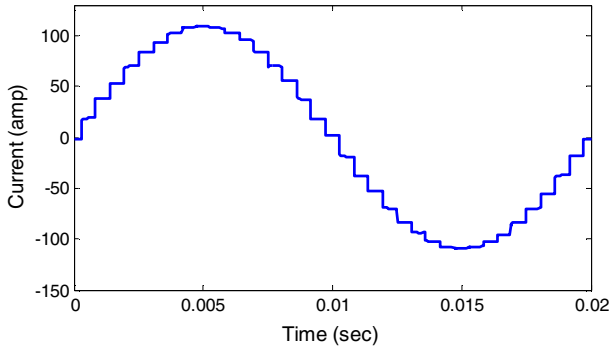


Fig. 7: Waveforms of input line current.

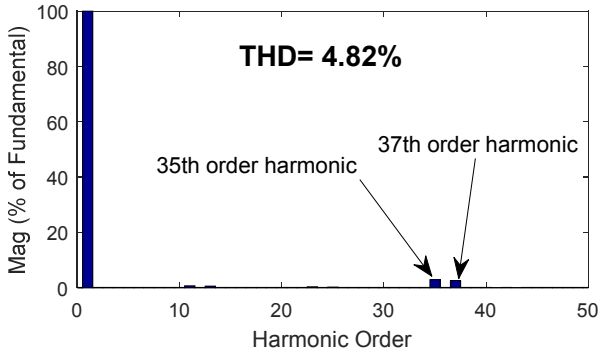


Fig. 8: Spectrum of input line current.

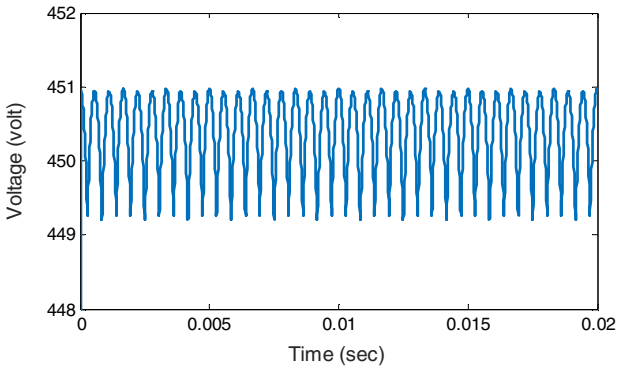


Fig. 9: Waveforms of output voltage.

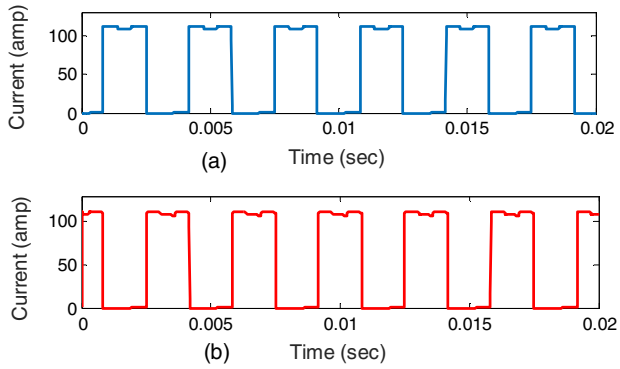


Fig.10: Waveforms of (a) Current i_{P1} (b) Current i_{Q1} of first injection circuit.

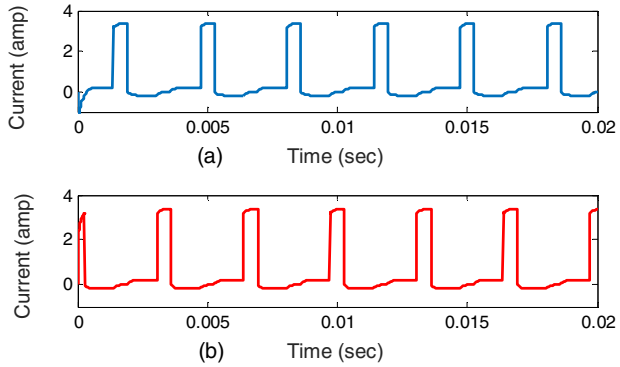


Fig. 11: Waveforms of (a) Current i_{P2} (b) Current i_{Q2} of second injection circuit.

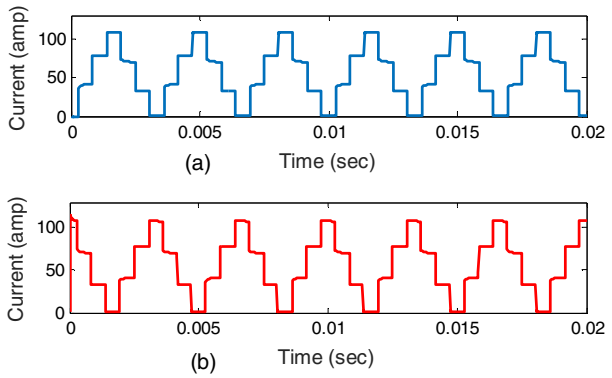


Fig. 12: Waveforms of (a) Current i_5 of double-star rectifier-1 (b) Current i_6 double-star rectifier-2.

VI. CONLUSION

In order to extend four-star 12-pulse rectifier operation to 36-pulse rectifier operation, passive injection method is introduced and presented in this paper. The proposed rectifier has been designed, analyzed and simulated in MATLAB environment. After using this method, the THD of the input line current is about 4.82% which is within the IEEE-519 standard requirements. The load voltage contains 36 symmetrical pulses and input line current contains 36 steps per cycle of power supply. The proposed rectifier works satisfactorily under wide load variation condition and preserves pulse tripling technique when STIT designed optimally. It suppresses lower than 35th order harmonics from input line currents. We will design a laboratory prototype in the next research.

REFERENCES

[1] J. R. Rodriguez *et al.*, "Large current rectifiers: State of the art and future trends," IEEE Trans. Ind. Electron., vol. 52, no. 3, pp. 738–746, Jun. 2005.

[2] IEEE Recommended Practices and Requirements for Harmonic Control in Electrical Power Systems, IEEE Std. 519-1992, 1992.

[3] B. Badrzadeh, K. S. Smith, and R. C. Wilson, "Designing passive harmonic filters for an aluminum smelting plant," IEEE Trans. Ind. Appl., vol. 47, no. 2, pp. 973–983, March/April. 2011.

[4] H. Akagi and K. Isozaki, "A hybrid active filter for a three-phase 12-pulse diode rectifier used as the front end of a medium-voltage motor drive," IEEE Trans. Power Electron., vol. 27, no. 1, pp. 69–77, Jan. 2012.

[5] B. Singh and S. Gairola, "Zigzag autotransformer based full-wave ac-dc converters," Asian Power Electronics Journal, vol. 2, no. 1, pp. 23–30, Apr. 2008.

[6] B. Singh and S. Gairola, "A 28-pulse AC–DC converter for line current harmonic reduction," IET Power Electron., vol. 1, no. 2, pp. 287–295, June 2008.

[7] R. Abdollahi and G. B. Gharehpetian, "Inclusive design and implementation of novel 40-pulse ac-dc converter for retrofit applications and harmonic mitigation," IEEE Trans. Ind. Electron., vol. 63, no. 2, pp. 667-677, Feb. 2016.

[8] F. Meng, W. Yang, S. Yang, and L. Gao, "Active harmonic reduction for 12-pulse diode bridge rectifier at dc side with two-stage auxiliary circuit," IEEE Trans. Ind. Inform., vol. 11, no. 1, pp. 64–73, Feb. 2015.

[9] F. Meng, X. Xu, L. Gao, and C. Cai, "Active harmonic reduction using dc-side current injection applied in a novel large current rectifier based on fork-connected autotransformer," IEEE Trans. Ind. Electron., vol. 64, no. 7, pp. 5250-5264, July. 2017.

[10] M. A. Malek and M. A. G. Khan, "A new 36-pulse diode rectifier with two passive injection circuits," in 2nd International Conference on Electrical & Electronic Engineering (ICEEE), pp. 1-4, Dec. 2017.

[11] S. Yang, J. Wang, and W. Yang, "A novel 24-pulse diode rectifier with an auxiliary single-phase full-wave rectifier at DC side," IEEE Trans. Power Electron., vol. 32, no. 3, pp. 1885–1893, March 2017.

[12] F. Meng, X. Xu, L. Gao, Z. Man, and X. Cai, "Dual passive harmonic reduction at dc link of the double-star uncontrolled rectifier," IEEE Trans. Ind. Electron., vol. 66, no. 4, pp. 3303–3309, April 2019.

Experimental Results of a locally developed BLDC Motor Controller for electric tricycle

Md. Fahim Bhuiyan, Nazmus Sakib, Mohammad Rejwan Uddin, and Khosru M Salim
Department of Electrical and Electronic Engineering
Independent University, Bangladesh
Dhaka, Bangladesh
fahim887@gmail.com; nazmussakibmustak@yahoo.com; rejwan.azad@gmail.com and khosru@iub.edu.bd

***Abstract*—The rapid growing battery-powered and plug-in hybrid electric vehicles will have a significant impact on power networks due to their zero carbon emission and mass deployment. DC motors, AC induction motors, and Brushless Direct Current (BLDC) motors are used worldwide in electric vehicles for years. Among them, BLDC motors are one of the electrical drives that are continually gaining popularity in motion control applications due to their high efficiency, good dynamic response and low maintenance. Unlike DC motors and Induction motors, BLDC motors need a controller as the motors run in 3 phase but take DC input. The controller converts DC power to 3 phase AC power suitable for the motor. This paper describes the design and development of a BLDC motor controller and its performance comparison with the existing BLDC motor controller imported from abroad. The controller circuit is simulated in Proteus software and the PCB is designed in ARES.**

Keywords*— *Electric vehicle, BLDC motor controller, No-load test, Performance comparison.

I. INTRODUCTION

In the urban area of Bangladesh, according to BRTA (Bangladesh Road Transport Authority), over 332,000 registered motorcycles were on the road in 2016 and now it has almost doubled [1]. The reason behind this massive increment is because motorcycle overcomes the traffic easily; it is much cheaper than four-wheelers and consumes less fuel. Both males and females will feel comfortable riding these electric tricycles/e-bikes as they are very easy to ride and there is no need to maintain balance like in cycles or bikes. They are the best solution for green transportation due to their zero greenhouse gas emissions. The vehicles run using electrical energy that is stored in a battery bank and requires a motor controller to control speed and some other functions.



Fig.1. A typical electric tricycle

The market of these e-bikes is getting vast in size day by day but yet there are no companies or facilities to manufacture or develop e-bikes and its controller locally. It is high time to design, develop and manufacture local electric vehicle and their motor controller which will be suitable for the ambient condition of Bangladesh.

For these E-bikes, BLDC motor is preferred over traditional DC brushed motors or AC induction motors [2]. They are more reliable and can be used in poor environmental conditions as the BLDC motor have no chance of sparking, unlike brushed DC motors, making them better suited to environments with volatile chemicals and fuels [3]. DC motors have been used in the industries for many years, due to precise control capability and high efficiency. Since the late 1980's new design concept of permanent magnet brushless direct current (BLDC) motors are gaining grounds in electrical vehicle sector.

These electric vehicles run using electrical energy that is stored in a 48V/60V/72V battery bank. This system voltage mainly depends on the BLDC motors and its controller. The BLDC motors are typically permanent magnet synchronous motors; they are well driven by dc voltage and have a trapezoidal back-EMF. They have a computation that is done mainly by electronics application [4].

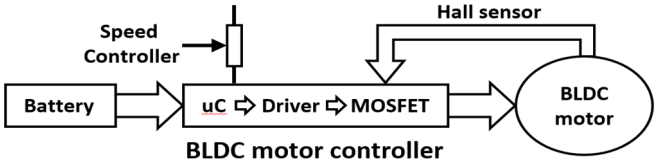


Fig.2. The functional block diagram of the electric tricycle

The paper will focus on designing and developing a BLDC motor controller and its performance comparison with the existing BLDC motor controller. The basic principle will be according to the block diagram of Fig.2. The controller will work according to the signals obtained from the BLDC motor's inbuilt hall sensors. This controller will invert the DC power from the battery to 3 phase AC power that will run the motor. The controller will be able to control the rotor speed and also the rotating direction. Lead acid batteries will be used in the battery bank.

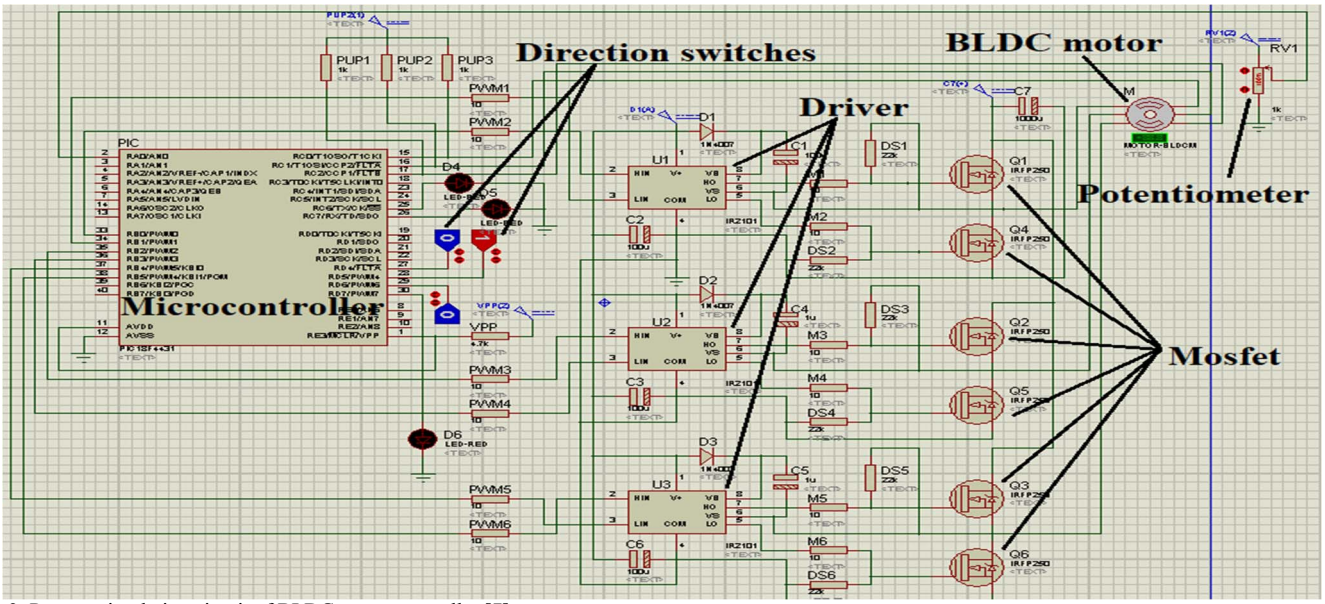


Fig.3. Proteus simulation circuit of BLDC motor controller [7]

II. Simulation of The Proposed BLDC Controller

Simulation of the BLDC motor controller is carried out by using Proteus ISIS Simulator software (Fig.3). Total 6 PWM signals of 20 KHz are generated by using PIC18f4431 microcontroller (PWM0 to PWM5, Pin B0 to B5). The duty cycle of the PWM signals is controlled by a reference input ADC signal (through potentiometer) into the microcontroller. A voltage divider circuit is then used to obtain the reference in the microcontroller to generate PWM signals and to control its duty cycle, output current and rotor speed [6]. Then 3 dual channels MOSFET driver (IR2101) is used to drive the MOSFETs with PWM signals. These drivers also initiate the dead time to the switches. Six IRFP250N N channel MOSFETs are used as switches to create the full bridge motor driver circuit.

The code was written in CCS PIC C compiler in C language for this project. The main task was to generate 6 different PWM signals according to different hall sensor combination. Six different combinations of hall sensors were taken as input in 3 analog pins (C0, C1, C2) of the PIC18F4413 microcontroller.

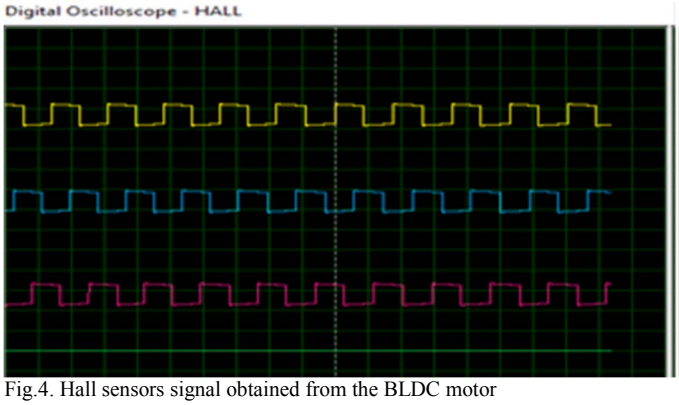


Fig.4. Hall sensors signal obtained from the BLDC motor

Hall sensors or optical sensors are used in BLDC motors to detect the position of the rotor (Fig.3). But it can be driven also without any position detection by sensors [5] using back-emf signals. The terminal voltages on the stator windings of each phase are controlled by the power electronic switches (MOSFETs). Each phase of the motor will connect with DC power source through two switches, one switch with positive terminal and other switch with negative terminal. To control terminal voltage of the three phases of the motor, 6 MOSFETs are used in this BLDC motor controller. After successfully circuit design in simulation software, desired output was achieved in 3 phases of the motor (Fig.5).

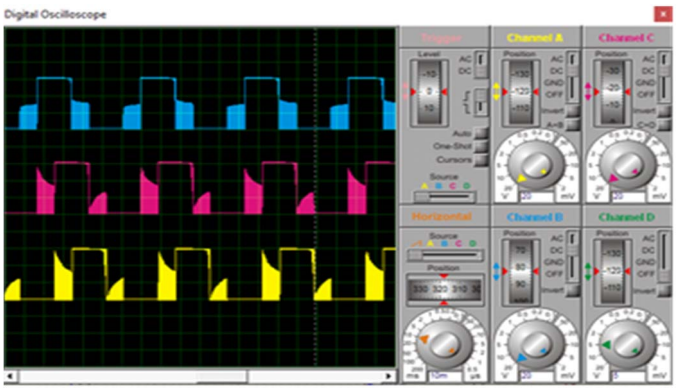


Fig.5. Three phase output waveform of the motor controller

III. System Development and Testing

Almost all power-electronics products have become microprocessor controlled in recent years. The main advantage of using a microcontroller is that it reduces the complex control hardware where the control functions can be changed as many times as it is required without extra

hardware implementation. Furthermore, it offers the possibility of improved reliability and flexibility.

After successful simulation, a PCB (printed circuit board) of the BLDC motor controller was made. The PCB was double-sided so components were placed on both sides. The microcontroller was placed on top side of the PCB which produces 6 distinctive PWM signals to drive the power converter full bridge circuit through a MOSFET driver circuit. The power converter is actually a three-phase power semiconductor bridge.

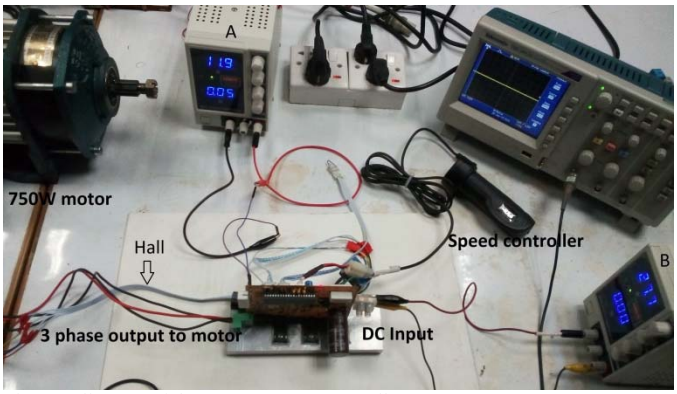


Fig.6. Full setup of the BLDC motor controller

This driver provides amplified gate signals at the same duty cycle to drive one high side and one low side MOSFET at a time. The bootstrap circuit is used to complete the driver circuit and amplify the signal coming from the microcontroller. The main function of the power converter is to transform power from the DC power to 3 phase AC power so the motor can convert electrical energy to mechanical energy. This converter can takes 48V to 60V DC power as input and gives a special kind of 3 phase AC power output (10V amplitude) to the motor according to the 3 Hall sensor's references. The complete setup can be seen in Fig.7 where the controller takes power from Lab bench power supply (A & B) and runs the motor according (upper left corner). The speed of the motor is controlled by using a throttle. This converter runs in 12V DC. The motor used in this project was 750W but the converter was designed to handle up-to 1000W of power. To meet up the criteria and availability, a components selection was done to get the possible optimized result.

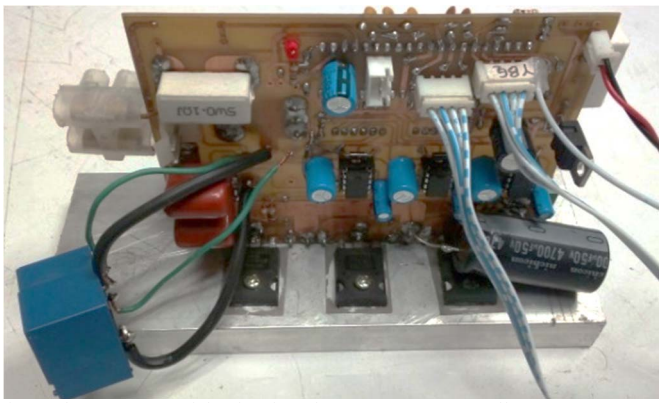


Fig.7. The prototype of the developed BLDC motor controller

IV. Results and Discussions

Once developed, the controller was tested using a 750W BLDC motor. The input of the controller was 48V DC which consists of 4 * 12V, 100Ah lead acid batteries. At full speed current consumed by the controller was 3A [Table 2]. The motor was well driven by the controller for hours and the motor running sound was quite low. Heat generation of the MOSFETs was tolerable so no extra cooling system was required. The three outputs or phase signals of the controller were considerably smooth and noise free (Fig.8).

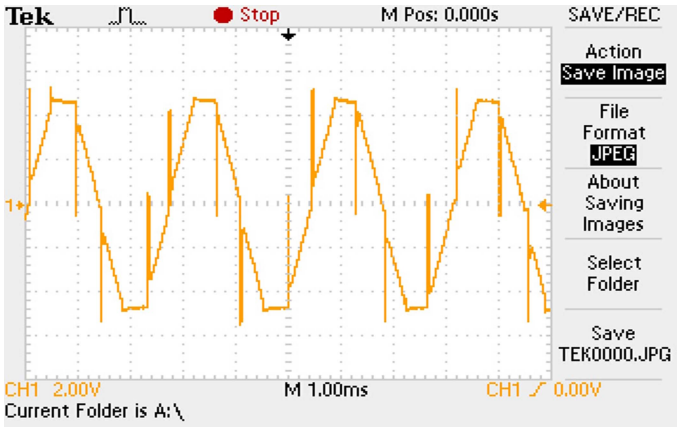


Fig.8. One of the three phases of the developed controller

Table 1 shows the experimental data obtained from the Existing BLDC motor controller for running 750W BLDC motor with no load.

TABLE 1. EXPERIMENTAL DATA OF THE EXISTING CONTROLLER

Speed	Input voltage (V)	Input current (A)	Input Power (W)
5%	48.6	0.6	29.16
10%	48.6	0.6	29.16
15%	48.6	0.7	34.02
20%	48.6	0.8	38.88
25%	48.5	0.9	43.65
30%	48.5	1.1	53.35
35%	48.4	1.3	62.92
40%	48.4	1.5	72.6
45%	48.4	1.8	87.12
50%	48.4	2.1	101.64
55%	48.3	2.3	111.09
60%	48.3	2.5	120.75
65%	48.3	2.7	130.41
70%	48.2	2.9	139.78
75%	48.2	3.1	149.42
80%	48.2	3.2	154.24
85%	48.2	3.2	154.24
90%	48.2	3.3	159.06
95%	48.2	3.4	163.88



Fig.9. Existing BLDC motor controller

Table 2 shows the experimental data obtained from the developed BLDC motor controller for running 750W BLDC motor with no load. The voltage of the battery bank with the corresponding input current was recorded as well.

TABLE 2. EXPERIMENTAL DATA OF THE DEVELOPED CONTROLLER

Speed	Input voltage (V)	Input current (A)	Input Power (W)
5%	48.6	0.3	14.58
10%	48.6	0.4	19.44
15%	48.6	0.5	24.3
20%	48.6	0.6	29.16
25%	48.6	0.8	38.88
30%	48.5	0.9	43.65
35%	48.5	1.1	53.35
40%	48.5	1.3	63.05
45%	48.5	1.5	72.75
50%	48.4	1.8	87.12
55%	48.4	1.9	91.96
60%	48.4	2.1	101.64
65%	48.4	2.2	106.48
70%	48.4	2.3	111.32
75%	48.3	2.4	115.92
80%	48.3	2.5	120.75
85%	48.3	2.7	130.41
90%	48.3	2.9	140.07
95%	48.3	3.0	144.9

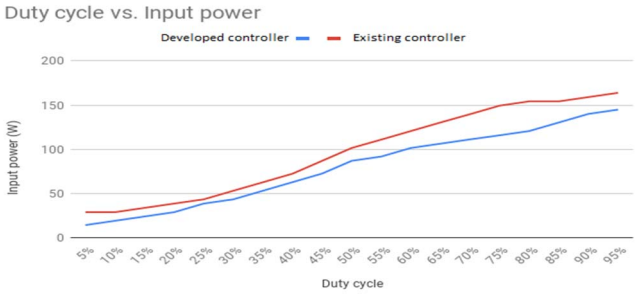


Fig.10. Duty cycle vs. Input power of developed controller and Existing controller

TABLE 3. PERFORMANCE COMPARISON OF EXISTING CONTROLLER AND DEVELOPED CONTROLLER (AT FULL SPEED / FULL THROTTLE)

Existing BLDC motor controller	Developed BLDC motor controller
Frequency: 16KHz	Frequency: 17KHz
Period: 62.5 us	Period: 59us
Input voltage: 48V	Input voltage: 48V
Input current: 3.4A	Input current: 3A
Input power: 163.2W	Input power: 144W

V. CONCLUSION

In this paper, a detailed description of the design and development of a more efficient and lightweight BLDC motor controller for electric bike/electric tricycle has been presented. The controller tested and experimented successfully. The experimental data of the no load test shows an excellent performance of the locally developed BLDC motor controller over the existing BLDC motor controller. The developed motor controller consumes less power compared to the existing motor controller and makes less noise. The developed motor controller has an improved performance over existing controller due to proper selection of the filtering component and for using high frequency switching. Future work of this project is to implement this motor controller in an electric vehicle.

REFERENCES

[1] S. M. Labib, Hossain Mohiuddin, Shahadat Hossain Shaki "Transport Sustainability of Dhaka: A Measure of Ecological Footprint and Means for Sustainable Transportation System" Journal of Bangladesh Institute of Planners ISSN 2075-9363 Vol. 6, December 2013.

[2] Mohammad Rejwan Uddin, Zaima Tasneem, Saila Ishrat Annie, and Khosru M Salim, "A High Capacity Synchronous Buck Converter for Highly Efficient and Lightweight Charger of Electric Easy Bikes" International Conference on Electrical, Computer and Communication Engineering (ICECCE 2017), Chittagong, Bangladesh.

[3] Mohammad Iqbal, Rakibulhasan and Tanvir Ehsan "study on merits and demerits of two transport systems: battery operated easy bike with motor operated rickshaw at Sylhet city in Bangladesh" International Conference On Mechanical Engineering And Renewable Energy 2015 (ICMERE2015) 26-29 November 2015.

[4] Jinka Prathima, Embedded Control Of Brushless Dc Motor Using Fuzzy Logic Controller, International Journal Of Innovative Research In Advanced Engineering (Ijirae) Issn: 2349-2163, Volume 1 Issue 10 (November 2014)

[5] N.Z. Yahaya, M.K. Lee, K.M. Begam & M. Awan "The Analysis of Dead Time on Switching Loss in High and Low Side MOSFETs of ZVS Synchronous Buck Converter" ACEEE Int. J. on Control System and Instrumentation, Vol. 02, No. 02, June 2011.

[6] Khosru M. Salim, Md. Jasim Uddin, M. Ishtiaque Rahman, Mohammad Rejwan Uddin "Design, Construction and Implementation of a Highly Efficient, Lightweight and Cost Effective Charger for Electric Easy Bikes" ICDRET 2016

[7] Md. Fahim Bhuiyan, Mohammad Rejwan Uddin, Zaima Tasneem, Mahady Hasan, and Khosru M Salim "Design, Code Generation and Simulation of a BLDC Motor Controller Using PIC Microcontroller" ICRIEECE 2018

HOMER Optimized Off-grid Hybrid Energy System: A Case Study on Rohingya Relocation Center in Bangladesh

Ferdous S. Azad
Department of EEE
Islamic University of Technology (IUT)
Dhaka, Bangladesh
ferdousr@iut-dhaka.edu

Istiak Ahmed
Department of EEE
Islamic University of Technology (IUT)
Dhaka, Bangladesh
tuhamahmed@gmail.com

Syed Raihan Hossain
Department of EEE
Islamic University of Technology (IUT)
Dhaka, Bangladesh
raihan93@iut-dhaka.edu

Rashedul Amin Tuhin
Department of CSE
East West University
Dhaka, Bangladesh
mcctuhin@ewubd.edu

***Abstract*— The use of fossil fuels in generating energy is not only depleting these fixed resources at a higher pace but also has enormous negative impact on the environment. In this scenario, using renewable energy to design hybrid generation systems is both cost optimizing and environment-friendly. This paper carries a detailed study of a renewable energy-based electrification scheme for a Rohingya Relocation Center, an isolated island named Bhasan Char with sensitivity analysis of the proposed system. Using HOMER (Hybrid Optimization Model for Electric Renewables), a hybrid system is designed which contains a combination of the wind turbine, solar PV modules, microturbine generators, converter, and battery. The focus of this study is to achieve low NPC (net present cost) and low LCOE (Levelized cost of energy) with reduced carbon emissions.**

***Keywords*—** *wind energy, PV modules, hybrid system, microturbine generator, etc.*

I. INTRODUCTION

Energy plays a vital role in our day to day activities and is a crucial tool to meet up the future challenges of human civilization. Electricity is a preferred form of energy for its numerous advantages and applications. Currently, an estimated 16 percent of the world's population that is 1.2 billion people have very little or no access to electricity [1]. Shocking comparability in this number is the urban-rural ratio of energy use, more than 80 percent of victims of energy-poverty live in rural areas. About 65% of the total population of Bangladesh lives in rural areas[2]. The fact is that Bangladesh faces a deficit in adequate electricity generation capacity and grid-connected electricity distribution for the whole country[3]. For this reason, off-grid generation schemes utilizing various forms of renewable energy sources has become a recent trend to minimize the energy crisis. There is no rebutting of the fact that both developed and developing countries are looking forward to the novel and advanced types of energy solutions for their economic sustainability. Bangladesh has set a target to achieve 10% of total energy from renewable resources by 2020. However, the achievement from these forms of energy is only 3.5% until 2015 [4]. So, it is imperative in the context of our country to design and implement microgrids incorporating renewable energy resources on a broader scale. The advantage of the microgrid is that it fulfills the requirement of improved quality and reliability more economically and efficiently. This advantage can be used as a solution to provide electricity in isolated islands like Bhasan Char (a Rohingya relocation center). Some studies have been conducted previously for areas like St. Martin's island [5], Nijhum Dwip [6], Sandip-para [7] and also for Bangladesh as a whole [8]. However, this is the first initiative of a detailed study on Bhasan Char. A comprehensive load calculation is done based on the information collected from different sources.

In our study, HOMER has been used for finding the best hybrid electric system for Bhasan Char with low NPC (net present cost) and low COE (cost of energy). HOMER is the most frequently used software tools by the researchers worldwide than other similar software like HYBRID2, SOMES, Dymola, iHOGA, HOGA, RETScreen, etc. [9] [10].

II. JUSTIFICATION OF SITE SELECTION

Bhasan Char, previously known as Char Piya and Thengar Char, is an island in Hatiya Upazila, Chittagong district (Coordinates: 22.3872° N, 91.4145° E). The island is 30 km away from the mainland. It has become a place for national and international interest because the government of Bangladesh has decided to move the Rohingya refugees to this isolated island. Nearly 700,000 of Rohingya Muslims have crossed the Rakhine state, Myanmar since the end of August 2017. They joined about 300,000 Rohingya already staying in Bangladesh, one of the world's most crowded nations, who had fled the previous spell of brutality. The government has decided to turn this island into a home for 100,000 Rohingyas to reduce the pressure on Cox' Bazar [11] [12]. As such, the design and implementation of a suitable energy system which will provide low NPC and low LCOE have become the recent challenge for national and international bodies. Our study provides an appropriate hybrid system with sensitivity analysis to show the pathway in handling this challenge.



Fig. 1: Google earth top view of Bhasan Char (Char Piya)

III. LOAD ESTIMATION

To calculate the load we categorized the whole load into three main parts:

i) Primary Load 1
ii) Primary Load 2
iii) Deferrable Load

The total load distribution scenario is presented in table 1. Primary load 1 is mainly a household load. Mobile tower, police station, medical center, school, etc. are considered in primary load 2.

Load varies in different seasons in Bangladesh. For simplicity, we have divided them into two main categories, i.e. summer and winter. Scaled monthly average load graph is shown in figure 1 and 2 for primary load 1 and primary load 2 respectively. The deferrable load is considered fixed throughout the year. In our study, the deferrable load is 2.79 kWh/day. The overall average annual load is 814.8 kWh/day whereas peak load is 106.45 kW.

TABLE I. ESTIMATED LOAD SCENARIO FOR THE HYBRID SYSTEM

Category	Load Type	Sample considered of load type	Description of loads
Primary load 1	Household	100	3 CFL bulbs (20W each), 2 fans (40W each), 1 socket for mobile charge
Primary load 2	Mobile tower	1	Total 5 kW
	School	1	15 CFL bulbs (20W each), 15 Fans (40W each), 10 Computers (100W each)
	Shop	1	2 CFL bulbs (20w each), 1 Fan (40W), 1 refrigerator (200W)
	Hospital	1	5 CFL bulbs (20W each), 5 fans (40W each), 2 refrigerators (200W each), 4 computers (100W each), 1 printer (60W), lab equipment (100kw)
	Police Station	1	3 CFL bulbs (20W each), 2 fans (40W each)
Deferrable load	Water pump	3	water pumps (150W each)

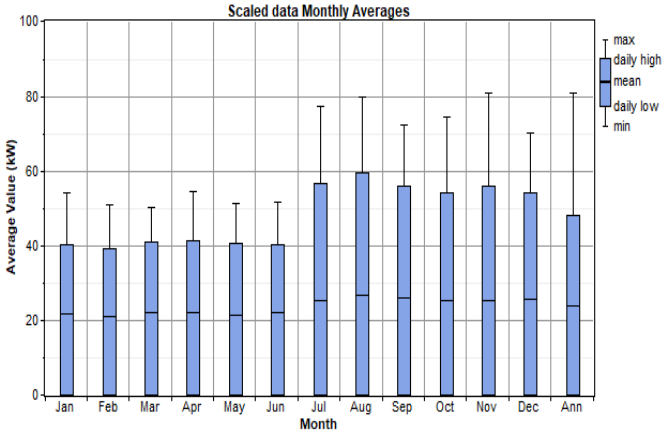


Fig. 2: Monthly average scaled load for primary load 1

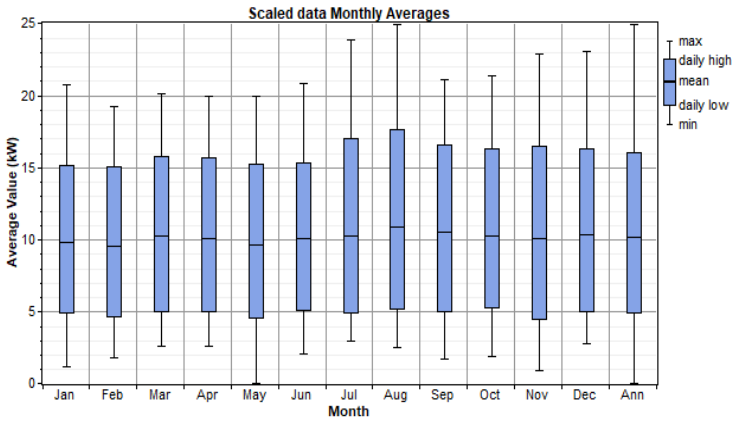


Fig. 3: Monthly average scaled load for primary load 2

IV. RENEWABLE ENERGY RESOURCE

The main renewable energy source in our selected site is Solar and Wind. The area is shiny and has promising wind speed in some of the months of the year.

A. Solar Energy:

Monthly solar irradiation data of Bhasan Char has been taken from NASA website. The average daily radiation is found to be 4.606 kWh/m^2/d with an average annual clearance index of 0.498. The maximum solar irradiation is found in February, April and May.

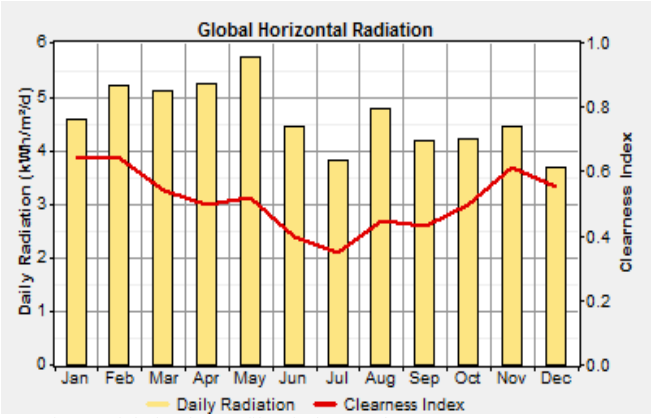


Fig. 4: Global Solar profile of Bhasan Char

B. Wind Resource

The average annual wind speed of Bhasan Char is 5.231 m/s which is quite promising. Wind speed is found maximum on the site at April, June, July, and August.

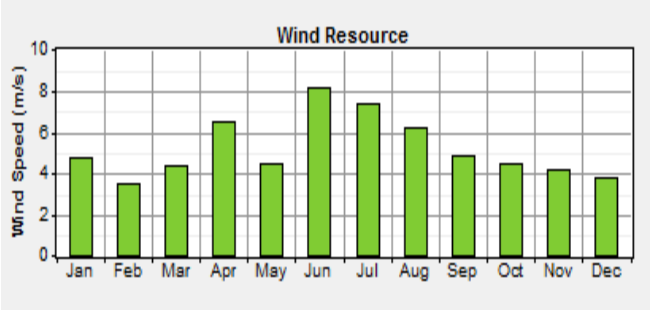


Fig. 5: Monthly wind resource of Bhasan Char

TABLE II. MONTHLY SOLAR AND WIND DATA FOR BHASAN CHAR

Month	Clearness index	Daily Radiation (kWh/m²/d)	Wind Speed (m/s)
January	0.648	4.560	4.750
February	0.643	5.210	3.500
March	0.542	5.100	4.360
April	0.500	5.230	6.490
May	0.521	5.730	4.510
June	0.399	4.440	8.140
July	0.345	3.810	7.380
August	0.445	4.770	6.220
September	0.428	4.170	4.900
October	0.496	4.200	4.470
November	0.611	4.430	4.200
December	0.551	3.680	3.770
Average	**0.498**	**4.606**	**5.231**

V. COMPONENTS OF THE HYBRID SYSTEM

Main components of this hybrid energy system are the wind turbine, PV panel, microturbine generator, converter and battery as shown in figure 6. To design the system, HOMER needs different parameters such as the number of units used, capital cost, replacement cost, operation and maintenance (O&M) cost of the system, lifetime and other parameters.

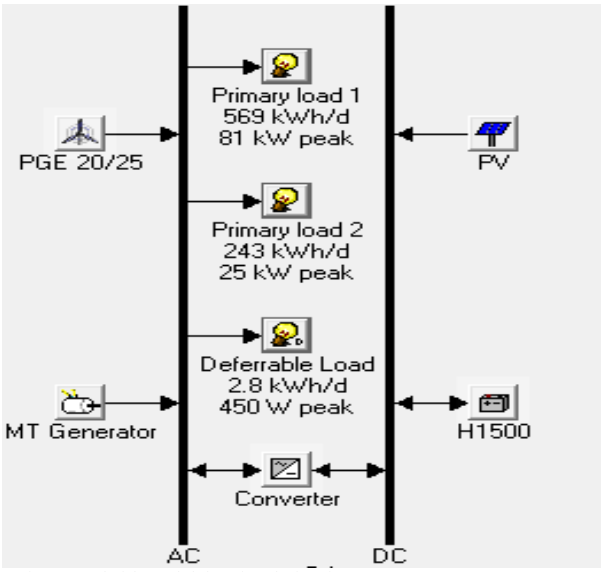


Fig. 6: Hybrid model for simulation in HOMER

A. Wind Turbine:

For the hybrid system, we used the PGE 20/25 wind turbine that has a rated capacity of 25 kW. It has rated wind speed of 12 m/sec and lifetime considered is 20 years.

TABLE III. PARAMETERS CONSIDERED FOR WIND TURBINE (PGE 20/25)

Parameter	Unit	Value
Power (rated)	kW	25
Starting Wind speed	m/sec	4
Rated Wind speed	m/sec	12
Cut off Wind speed	m/sec	25
Lifetime	Years	20
Capital Cost	BDT/turbine	9000000
Replacement Cost	BDT/turbine	7000000
O&M cost	BDT/year	20000

B. Solar PV modules

For solar modules, the capital cost is considered 1 lac BDT for one kW and replacement cost is 80 thousand BDT. Lifetime is 25 years.

TABLE IV. PARAMETERS CONSIDERED FOR PV MODULE

Parameter	Unit	Value
Capital Cost	BDT/kW	100000
Replacement Cost	BDT/kW	80000
O&M cost	BDT/kW/year	5000
Lifetime	Years	25
Derating factor	Percentage	92
Ground Reflectance	Percentage	20

C. Microturbine (MT) Generator

The fuel used in the MT generator was LNG. Bangladesh has started importing LNG in 2018 [13], and the price is as low as 33 BDT per cubic meter. Adding transportation and other costs, the LNG price is considered 35 BDT. In the sensitivity analysis, the LNG price is varied from 35 to 45 BDT. A 30 kW MT generator was considered in the system.

TABLE V. PARAMETERS CONSIDERED FOR MICROTURBINE GENERATOR

Parameter	Unit	Value
Capital Cost	BDT/kW	73333
Replacement Cost	BDT/kW	60000
O&M cost	BDT/hr (30kW)	30
Operational Lifetime	Hours	45000
Minimum load ratio	Percentage	20
Fuel Price (LNG)	BDT/m³	35
Lower Heating Value of fuel	MJ/kg	45
Fuel Density	Kg/m³	0.41

D. Converter:

For the purpose of converting AC-DC or DC-AC, the hybrid system needs converters.

TABLE VI. PARAMETERS CONSIDERED FOR CONVERTER

Parameter	Unit	Value
Capital Cost	BDT/W	15
Replacement Cost	BDT/W	10
Lifetime	Years	10
Efficiency	Percentage	90

E. Battery:

To store energy produced by PV panels, battery model of Hoppecke 12 OPzS 1500 are used. It is vented lead-acid, tubular-plate, deep-cycle battery with a nominal capacity of 1500 Ah.

TABLE VII. PARAMETERS CONSIDERED FOR BATTERY

Parameter	Unit	Value
Nominal Voltage	Volt	2
Maximum Charge current	A	306
Nominal Capacity	Ah	1500
Minimum Battery life	Years	10
Capital Cost	BDT/battery	100000
Replacement Cost	BDT/battery	87500
O&M cost	BDT/Year	50

VI. Optimization Results and Discussion

The Hybrid system comprising of MT generator, PV and wind is simulated using HOMER for an optimized system. HOMER shows the result of the optimized system in descending order with parameters such as Net present cost (NPC) and Cost of Energy (COE) are considered for economic analysis and hybrid system grading.

After setting suitable system control parameters and constraints, the simulation is carried out. In our simulation of MT-PV-Wind system, the optimized result is shown in fig. 7 with minimum COE of 16.88 BDT/kWh and minimum NPC of 82,089,760 BDT. 60 kW of PV, 2 wind turbine and 30 kW microturbine generator is found to be the most cost-effective and economically feasible solution. The unmet load is only 91 kWh/yr and capacity shortage is 0.1%.

PV (kW)	PGE25	Label (kW)	H1500	Conv. (kW)	Initial Capital	Operating Cost ($/yr)	Total NPC	COE ($/kWh)	Ren. Frac.	LNG (m3)	Label (hrs)	Batt. Lf. (yr)
60	2	30	130	50	$ 39,950,000	2,577,129	$ 82,089,760	16.886	0.68	54,056	4,039	16.9
60	2	30	140	50	$ 40,950,000	2,521,277	$ 82,176,496	16.904	0.68	53,497	3,994	17.9
50	2	30	150	50	$ 40,950,000	2,528,110	$ 82,288,216	16.928	0.65	55,932	4,168	19.0
50	2	30	160	50	$ 41,950,000	2,469,106	$ 82,323,416	16.934	0.65	55,445	4,130	20.0
60	2	30	150	50	$ 41,950,000	2,470,352	$ 82,343,800	16.938	0.68	53,200	3,970	19.0
70	2	30	110	50	$ 38,950,000	2,658,594	$ 82,421,824	16.956	0.69	53,291	3,998	14.9
60	2	30	160	50	$ 42,950,000	2,415,497	$ 82,446,840	16.959	0.68	52,851	3,941	20.0
70	2	30	120	50	$ 39,950,000	2,603,458	$ 82,520,264	16.975	0.70	52,711	3,950	16.0
70	2	30	130	50	$ 40,950,000	2,544,866	$ 82,562,208	16.982	0.70	52,171	3,907	17.1
70	2	30	140	50	$ 41,950,000	2,487,889	$ 82,630,552	16.996	0.70	51,566	3,856	18.1
70	2	30	150	50	$ 42,950,000	2,428,470	$ 82,658,968	17.002	0.70	51,075	3,817	19.1
70	2	30	160	50	$ 43,950,000	2,381,707	$ 82,894,328	17.050	0.70	50,749	3,791	20.0
60	2	30	100	100	$ 37,700,000	2,768,910	$ 82,975,648	17.071	0.67	55,652	4,166	13.6
60	2	30	110	100	$ 38,700,000	2,714,746	$ 83,089,992	17.091	0.67	55,169	4,124	14.8
50	2	30	120	100	$ 38,700,000	2,715,520	$ 83,102,648	17.098	0.65	57,004	4,253	15.6
80	2	30	110	50	$ 39,950,000	2,641,688	$ 83,145,392	17.102	0.71	51,782	3,891	15.1
80	2	30	100	50	$ 38,950,000	2,702,998	$ 83,147,888	17.103	0.71	52,642	3,963	14.1
60	2	30	120	100	$ 39,700,000	2,658,725	$ 83,173,968	17.108	0.67	54,442	4,064	15.7
50	2	30	130	100	$ 39,700,000	2,659,789	$ 83,191,368	17.115	0.65	56,613	4,222	16.8
50	3	30	120	50	$ 46,950,000	2,216,764	$ 83,197,272	17.118	0.75	47,055	3,523	18.0
80	2	30	120	50	$ 40,950,000	2,587,389	$ 83,257,520	17.125	0.71	51,295	3,851	16.2

Fig. 7: Optimized result of hybrid system simulated in HOMER **(NB: all currency values are in BDT, not in USD)**

This optimized system has renewable fraction around 68% and primary energy saving 75%. Production of electricity of each component is shown in table 8. Monthly average electricity production is shown in fig. 8. During June and July, wind turbine is supplying most of the electricity to the system as wind speed is higher in that time of the year.

The hybrid system in this study shows very less emission of pollutants as the renewable fraction is higher. CO_2 emission is found to be 60,290 kg/yr whereas CO is 351 kg/yr.

TABLE VIII. POWER GENERATION DATA OF PV-MT-WIND

Production	kWh/yr	Percentage (%)
PV panels	100,375	27.25
Wind Turbine	148,678	40.36
MT generator	119,290	32.38
Total	**368,343**	**100**

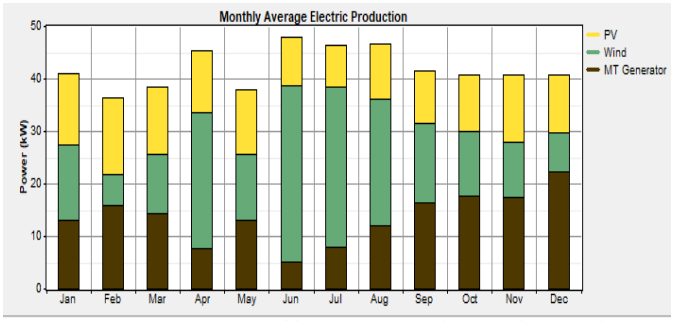


Fig. 8: Average monthly electricity production of PV-wind-MT

The result found in our study is compared with other relevant works, and it is very promising in both technical and economic perspectives. The optimized scheme in our study is PV-MT-Wind that has COE of 16.886 BDT. Some of the previous work [7] [14] [5] in this field of hybrid system had COE more than this.

VII. Sensitivity Analysis

Variation of wind speed, solar irradiation and price of LNG are considered for sensitivity study of the hybrid system. Fig. 10 shows sensitivity results in terms of LNG price and solar irradiance and fig. 9 shows the results in terms of LNG price and wind speed. In both cases, Levelized COE is superimposed and the maximum annual capacity shortage is 0%.

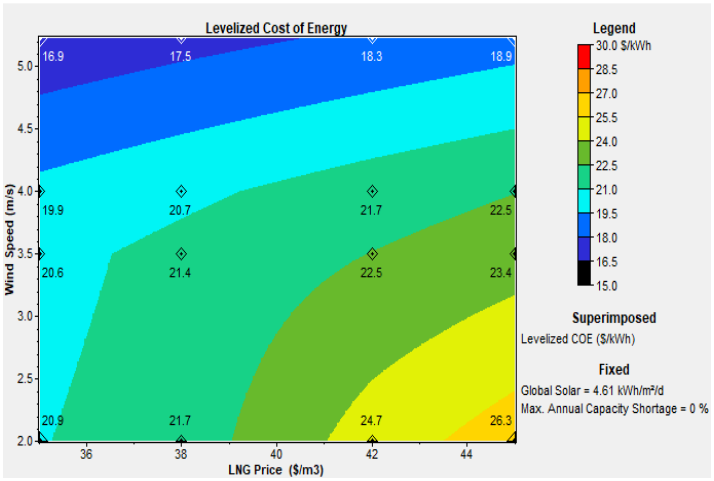


Fig. 9: Variation of LCOE in terms of LNG price and wind speed with fixed global solar o 4.61 kWh/m^2/d

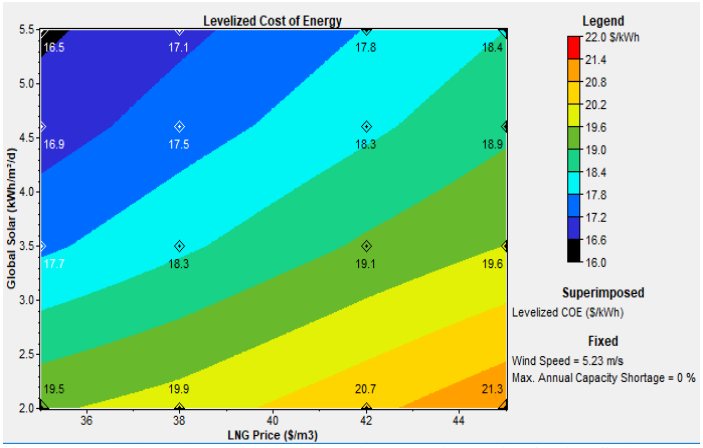


Fig. 10: Variation of LCOE in terms of LNG price and solar irradiation with a fixed wind speed of 5.23 m/s

It can be seen from fig 8 and 9; the LCOE is minimum when LNG price is minimum, 35 BDT. As wind speed and global solar is increased, the COE starts to decrease. If LNG price is increased COE also increases with fixed wind and solar profile.

The sensitivity analysis for optimal system type for fixed global solar of 4.61 kWh/m^2/d and a maximum annual capacity shortage of 0% is shown fig. 10. PV-MT system is feasible in terms of LCOE if the wind speed is below 3.3 m/s for any price of LNG. However, when the wind speed increases beyond this value, PV-MT-Wind system is economically more feasible.

Emission of CO_2, CO and other pollutants are less in the proposed system. The variation of this pollutants with other parameters can be seen from fig. 12. Wind-PV-MT system has lower carbon-di-oxide pollution than PV-MT system as the renewable fraction is high in the former system. Again for PV-MT system, CO_2 emission grows with the rise of fuel price. But for PV-MT-Wind system, CO_2 emission reduces with LNG price.

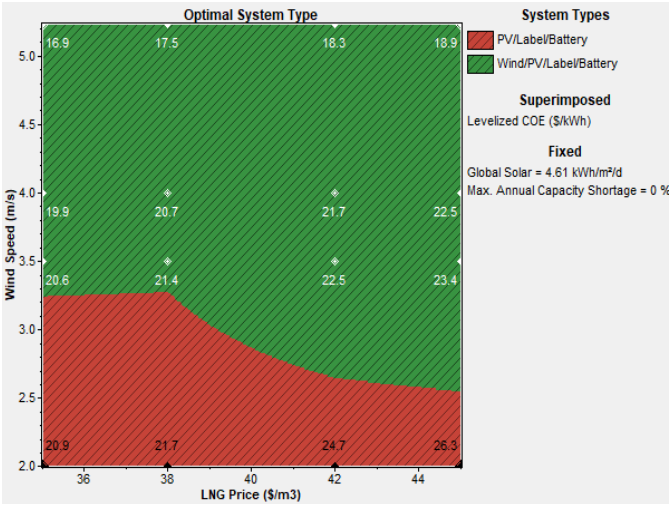


Fig. 11: Optimal system type in terms of LNG price and wind speed with a fixed global solar of 4.61 kWh/m^2/d

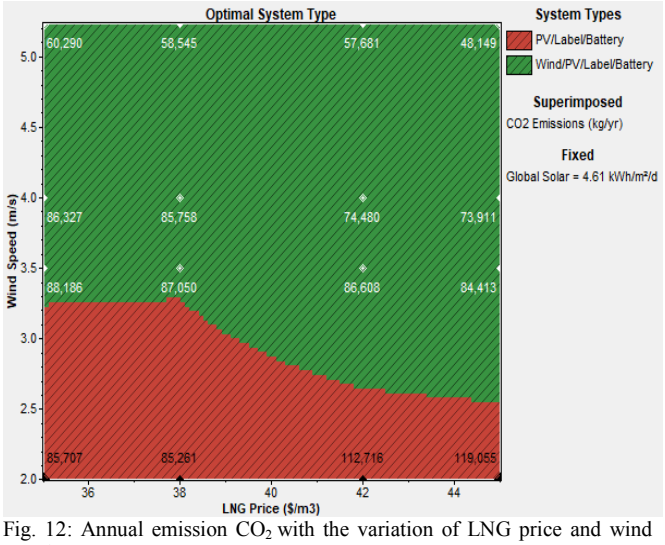


Fig. 12: Annual emission CO_2 with the variation of LNG price and wind speed

VIII. Conclusion

Hybrid systems are very important for electrification in areas like Bhasan Char that are far away from the mainland. This study of wind turbine, PV module and microturbine generator has found its optimized scheme with 60 kW from solar, 50 kW from wind and 30 kW from MT generator. A high fraction of renewable sources (around 67%) means the proposed scheme has fewer pollutants and thus more environment-friendly. LCOE is found to be less than 17 BDT per unit for the estimated load of 814 kWh/day. Similar studies of hybrid systems can be undertaken for other areas of Bangladesh to mitigate electricity problem and also to find alternative clean energy.

References

[1] N. Katsoulakos, L. Papada, and D. Kaliampakos, "The problem of energy poverty in mountainous areas," in *IISA*, 2014, pp. 222-226.

[2] "Bangladesh Rural Population Percent Of Total Population", 2019.[Online].Available:https://tradingeconomics.com/bangladesh/rural-population-percent-of-total-population-wb-data.html. [Accessed: 04- Feb- 2019].

[3] S. N. Uddin and R. Taplin, "A sustainable energy future in Bangladesh: current situation and need for effective strategies," in *Proceedings of the second joint international conference on sustainable energy and environment, SEE*, 2006, p. e3.

[4] M. A. H. Baky, M. M. Rahman, and A. S. Islam, "Development of renewable energy sector in Bangladesh: Current status and future potentials," *Renewable and Sustainable Energy Reviews,* vol. 73, pp. 1184-1197, 2017.

[5] A. S. Islam, M. M. Rahman, M. A. H. Mondal, and F. Alam, "Hybrid energy system for St. Martin Island, Bangladesh: an optimized model," *Procedia Engineering,* vol. 49, pp. 179-188, 2012.

[6] M. M. Hossain, S. Barua, and M. Matin, "A pre-feasibility study for electrification in Nijhum Dwip using hybrid renewable technology," in *Electrical & Electronic Engineering (ICEEE), 2015 International Conference on*, 2015, pp. 225-228.

[7] M. A. M. Bhuiyan, A. Deb, and A. Nasir, "Optimum planning of hybrid energy system using HOMER for rural electrification," *International Journal of Computer Applications,* vol. 66, 2013.

[8] M. S. H. Lipu, M. S. Uddin, and M. A. R. Miah, "A feasibility study of solar-wind-diesel hybrid system in rural and remote areas of Bangladesh," *International Journal of Renewable Energy Research (IJRER),* vol. 3, pp. 892-900, 2013.

[9] S. Sinha and S. Chandel, "Review of software tools for hybrid renewable energy systems," *Renewable and Sustainable Energy Reviews,* vol. 32, pp. 192-205, 2014.

[10] D. Kaur and P. Cheema, "Software tools for analyzing the hybrid renewable energy sources:-a review," in *2017 International Conference on Inventive Systems and Control (ICISC)*, 2017, pp. 1-4.

[11] I. Idris, "Rohingya refugee crisis: impact on Bangladeshi politics," 2017.

[12] "Rohingya relocation at Thengar Char: Minister", The Daily Star,2017.[Online].Available:https://www.thedailystar.net/country/rohingya-relocation-thengar-char-minister-1460200. [Accessed: 17- Feb- 2019].

[13] "Bangladesh receives 1st delivery of imported LNG, Dhaka Tribune,2018.[Online].Available:https://www.dhakatribune.com/bangladesh/power-energy/2018/04/24/bangladesh-receives-1st-delivery-imported-lng-supply-expected-mid-may.[Accessed: 11-Feb- 2019].

[14] S. Salehin, M. M. Ehsan, S. Noor, and A. Sadrul Islam, "Modeling of an optimized hybrid energy system for Kutubdia Island, Bangladesh," in *Applied Mechanics and Materials*, 2016, pp. 518-522.

A Design of High Differential Strobe Terminal Based Smart Power Factor Metering System

Shahriar Rahman Fahim[1], Sakib Shahrear Avro[2], Subrato K. Sarker[2], Md. Rokunuzzaman[3], and Sajal K. Das[2]
[1]Department of Electrical & Electronic Engineering
[2]Department of Mechatronics Engineering
[3]Department of Mechanical Engineering
Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
Email: fahim.fact@gmail.com; sakibavro.ruet@gmail.com; skshuvo138008@gmail.com;
rzaman_me.ruet@yahoo.com; das.k.sajal@gmail.com

***Abstract*—This paper represents a metering system which will measure the power factor of a load and frequency of a main-line and plot a real-time AC power triangle in MATLAB environment. The system uses Arduino as a controller. It reads the load voltage, current, their phase difference, and main-line frequency. A preliminary circuit arrangement, including a potential transformer (PT), a current transformer (CT), comparator and logic gate transforms the high value of voltage and current to the Arduino readable form. A serial communication between the Arduino and the MATLAB is done for plotting the power triangle. The simulation and experimental results are compared with the real data observed in the oscilloscope and frequency counter.**

***Index Terms*—Power triangle, MATLAB, high differential comparator, X-OR gate, power factor, Proteus and Arduino.**

I. Introduction

Electrical energy is the popular form of energy because of its easy transportation and reasonable cost. The cost of electrical energy increases due to power losses in transformer, transmission line and loads. Most of the loads are inductive where the current waveforms lagging the voltage. It consumes the reactive power which causes the real power to be less than the apparent power. The lower value of real power produces the poor power factor that can be used to measure the efficiency affecting by the phase lag [1]. A load with less power factor draws more current from the power line which is responsible for more losses.

The power factor directly relates to the system efficiency and performance. A large amount of load current results in lower voltage regulation of the system and provides a large joule loss in the power lines [2]. The lower voltage level generates heat in the motor windings that reduces the lifetime, reliability and performance of the machine. The reduced system performance and efficiency generates poor value of power factor that results in increasing utility charge. The economic returns of the power factor can be easily noticeable in the lower utility charges. Another visible benefit of power factor is the released system capacity, decreasing line losses and improved voltage regulation. Though having lots of benefits of power factor, the power factor measuring device is not available with the range of consumers. The system introduces in this paper measures the power factor and plots a power diagram for further analysis.

Fig.1 shows the basic block diagram of the proposed system. The system first converts the high value of load voltage and current to the proportional lower value. The conversion is done by using a PT of 20:1 trans ratio and a CT of 1:24 turns ratio. The output of CT and PT is fed into a comparator circuit designed with lm311 IC for detecting the zero crossing. Its strobe terminal operation and low voltage handling capacity makes it possible to respond to the lower value of CT output. As a result, no additional amplifier circuit is needed between CT and comparator circuit. An X-OR gate is used in the proposed system for detecting the lagging time pulse and phase delay between the voltage and current. A precision full wave rectifier is used to convert AC to DC voltage and current. The output from rectifier is passed through the Arduino and measures the equivalent voltage and current. Arduino sends the data to the MATLAB through serial communication, and then the Matlab calculates active, reactive, and apparent power and plots a power triangle. The MATLAB also calculates the required capacitance value to improve the poor power factor to unity.

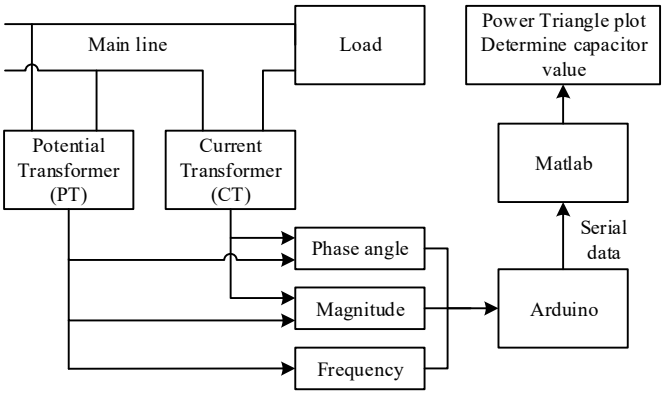


Fig. 1: Basic block diagram of the proposed system

In the previous literature, experts research on the field of power factor measurement. In [3] a system is demonstrated which is used for measuring and correcting the power factor (PF). The required capacitance is switched in or out by solid state relay and continuously monitors the PF. This system can not be used as a portable power factor meter. A PF measuring system based on Arduino is shown in [4] where the active power is calculated by taking the instantaneous value of voltage and current. It needs a high sampling rate. The higher sampling rate may cause sampling error, which affects the accuracy of measurement.

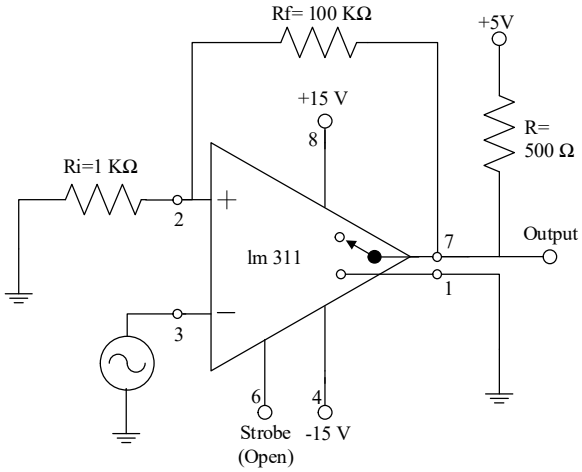


Fig. 2: A high speed differential comparator circuit using Lm 311

An ARM® microcontroller based approach for monitoring the power factor is introduced in [5]. This approach provides much accuracy. The drawback of this approach is its high implementation cost. In [6] power factor control unit based on Arduino is presented where a fixed line frequency is assumed. Moreover, the power line frequency may vary within a small range. The variation of frequency may affect the calculation of the power factor.

Motivated by above difficulties, this paper presents a smart power factor metering system for measuring the PF and frequency of a main-line. The proposed system is able to eliminate the error due to rounding the frequency like 50Hz. The advantage of proposed system includes the use of high-speed differential comparator lm311 which limits the use of the preamplifier circuit and reduces the cost of metering system. The aim of this work is to ensure the accurate PF by ploting a power triangle in the MATLAB environment. The power triangle is used to find the accurate PF by comparing the amount of power used in the load and consumed power from the transmission line.

The rest of this paper is organized as follows. Section II explains the theoretical background of this system. In section III the procedure of measuring power factor is described. The method of calculating voltage is expressed in section IV. The current measurement procedure is shown in section V. Section VI manifests the way of frequency measurement. The evaluation of the experimental results is illustrated in section VII. The conclusion of this paper is given in section VIII.

II. Theoritical Background

The power factor of the circuit $\cos(\phi)$ is measured by taking the phase difference ϕ between the voltage and current in an AC circuit. When the circuit is inductive, the current lags behind the voltage and the power factor is mentioned to as lagging. Most of the residential loads are inductive in nature and therefore have low lagging power factor. The power factor can also be calculated in terms of power drawn by the AC circuit. The ratio of active power to the apparent power is referred to as power factor. The relation between the active, reactive and apparent power is shown in Fig. 3 and discussed in (1), (2) & (3) as follows:

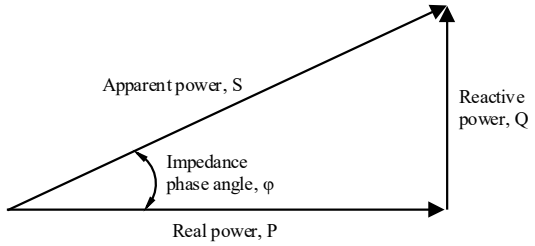


Fig. 3: Relation between the active, reactive and apparent power

$$P = V * I * \cos(\phi) \quad (1)$$

$$Q = \sqrt{S^2 - P^2} \quad (2)$$

$$S = V * I \quad (3)$$

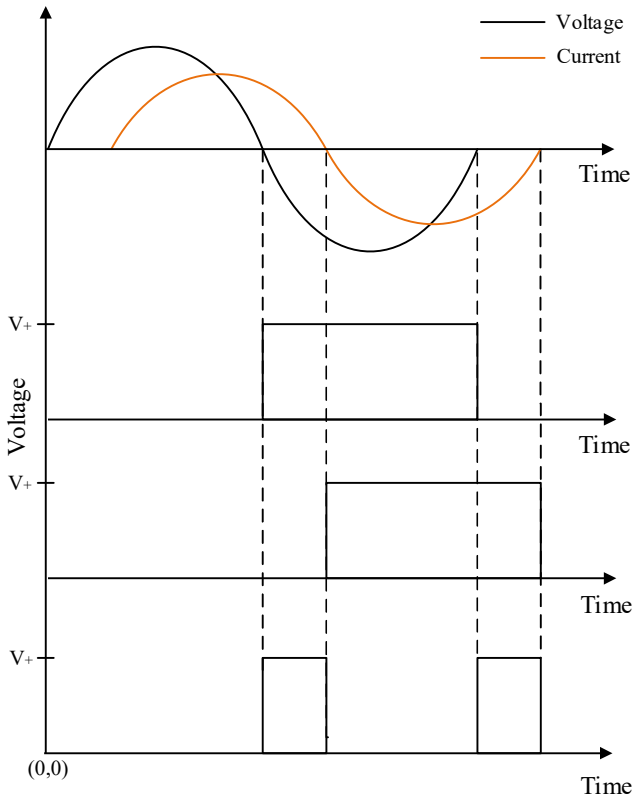


Fig. 4: A phase angle generation from the time delay between the voltage and current

For measuring the phase angle, the sinusoidal output from CT and PT are fed into two different comparator circuit. Lm 311 is chosen to make a comparator circuit shown in Fig. 2 because of its strobe terminal operation and faster response time of 165 nS [7]. The result from strobe terminal acts like a switching output and connecting between a voltage source V_+ and ground. When the strobe switch is open, the lm311 operates normally and gives the positive output voltage V_+ for negative values of the input signal and 0 V for the positive values as reported in Fig. 4.

A Proteus simulation result with the input and output wave shape of comparator circuit is presented in Fig. 5. Here the sinusoidal output of channel A shows the input wave shape and channel B shows the output of the comparator circuit. From Fig. 5, it is observed that the comparator generates an

output of +5 volt for every negative value and returns 0 volts for the positive value of input wave. The switching outputs are passed through the X-OR gate to detect the time delay between the voltage and current. This logic gate detects the inequality between two pulses and generates another pulse. The results obtained from the X-OR gate is reported in Fig. 6. The on time of this pulse denotes the time delay between the voltage and current.

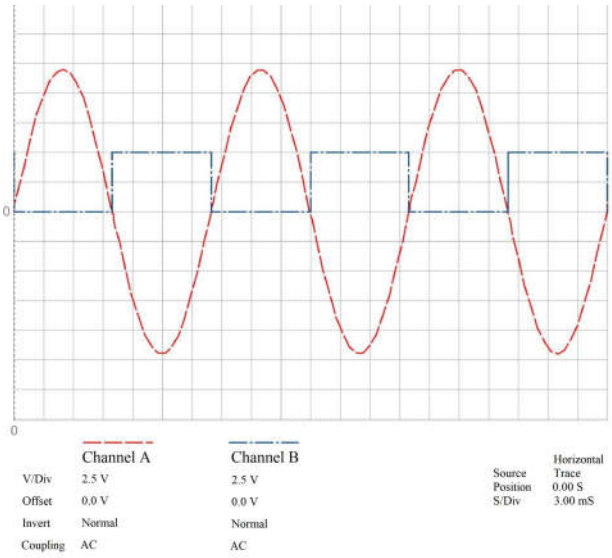


Fig. 5: A simulation result of comparator output

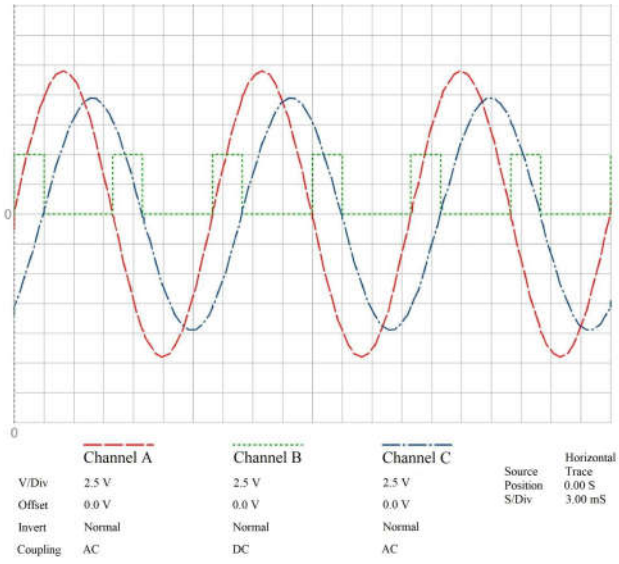


Fig. 6: A simulation result of X-OR gate output

III. Power Factor Measurement

This section presents the brief discussion of the power factor measurement by using phase difference between the voltage and current of an inductive load. The phase difference is obtained by measuring the time delay with the help of Arduino board. After measuring the time delay, the phase difference between the voltage and the current is measured. The relation between the time delay and the phase difference is described by the following equations:

$$\phi_D = 2 * 180 * f * T_D * 10^{-6} \quad (4)$$

$$\phi_R = \frac{\pi}{180}\phi_D \quad (5)$$

Here,
ϕ_D = Phase difference in degrees
T_D = Time delay in μS
f = Frequency in Hz
ϕ_R = Phase difference in radians

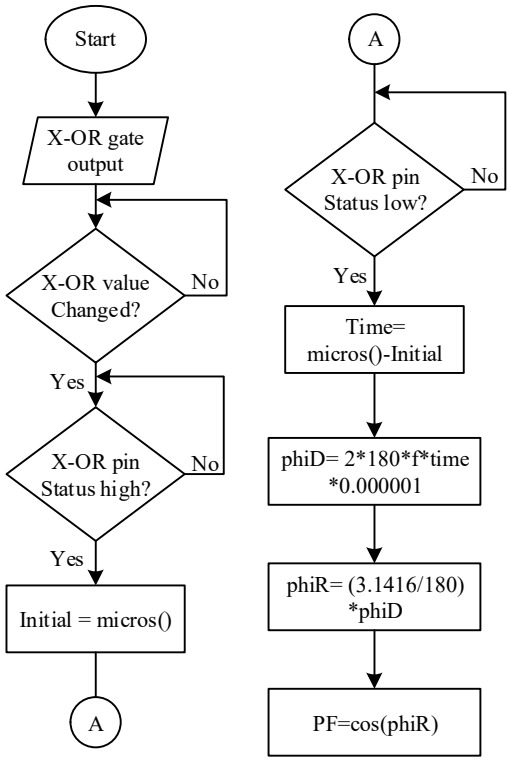


Fig. 7: A flow chart of power factor measurement

A flow chart of the programme for measuring the power factor is shown in Fig. 7. The time delay between the voltage and the current is measured using the pulseIn() function of Arduino. The process of the pulseIn() function is that it waits until the rising edge of X-OR output is detected. While detecting the signal from low to high, it stocks the time in its memory using the function micros(). After that, a recursive function is called which call itself again and again until detecting the falling edge. If the signal goes low, it deducts the current time from the starting time. This process returns the on-time duration of the pulse. Based on this on-time duration, the power factor is then evaluated using (4) and (5). After that, taking the cosine of the phase angle and returns the power factor. The obtained power factor is then put in (1) and calculates the required active power for the load. From (1), it is seen that the load voltage and current also needs to be measured in order to calculate the active power. The proposed system is able to measure the load voltage and current. The measurement process of load voltage and current is discussed in the next section.

IV. Voltage Measurement

For measuring the voltage across the load, the output from the potential transformer is given to a rectifier circuit. A precision full wave rectifier circuit as shown in Fig. 8 is chosen for the rectification purpose.

The rectifier circuit excludes the additional voltage drop across the diode and rectifies low value of AC signal by adding the large gain of OP-Amp. A voltage divider is employed to

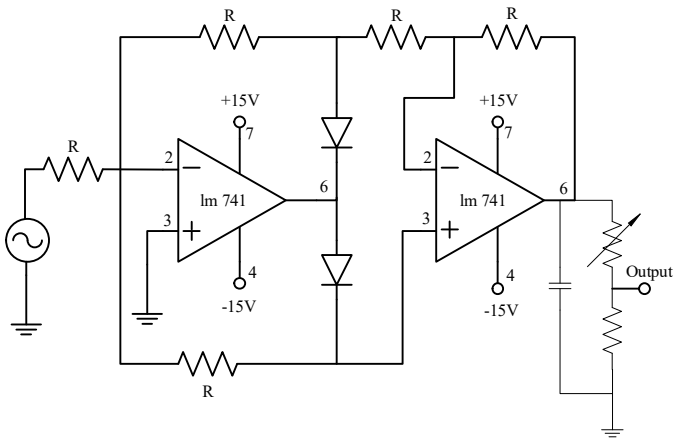


Fig. 8: A precision circuit diagram of full wave rectifier

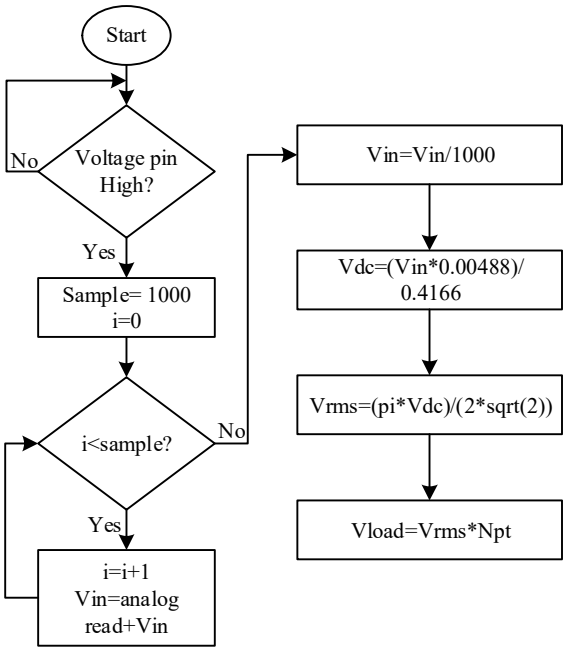


Fig. 9: A flow chart of load voltage measurement

convert the high value of DC signal into Arduino compatible form. The ratio of Arduino input to the total DC voltage is 0.4166 used for the proposed system. As the Arduino uses 10-bit ADC, it gives 0-1023 (1024 steps) for a 0-5V input. After using this step, the Arduino calculates the load voltage by using the following equations:

$$V_{DC} = \frac{ADC\,reading \times 0.00488}{0.4166} \tag{6}$$

$$V_{RMS} = \frac{\pi \times V_{DC}}{2\sqrt{2}} \tag{7}$$

$$V_{load} = V_{RMS} \times N_{PT} \tag{8}$$

Here,
V_{DC} = Rectified DC voltage
V_{RMS} = RMS value of AC voltage at PT secondary
V_{load} = Load voltage
N_{PT} = Transformer ratio of PT

A flow chart of the programme for AC voltage measurement is reported in Fig. 9. For measuring the AC Voltage, Arduino waits for the voltage pin to go high. When the voltage pin goes high, the programme starts sampling for 1000 times. This is done by using a "for" loop. Inside the "for" loop, the programme samples the analog input from the rectifier circuit and stocks it into the memory. When the " for" loop is executed, the average value is computed by summing all the sampled data and dividing it by 1000. After completing this step, the load voltage is calculated by using (6)-(8).

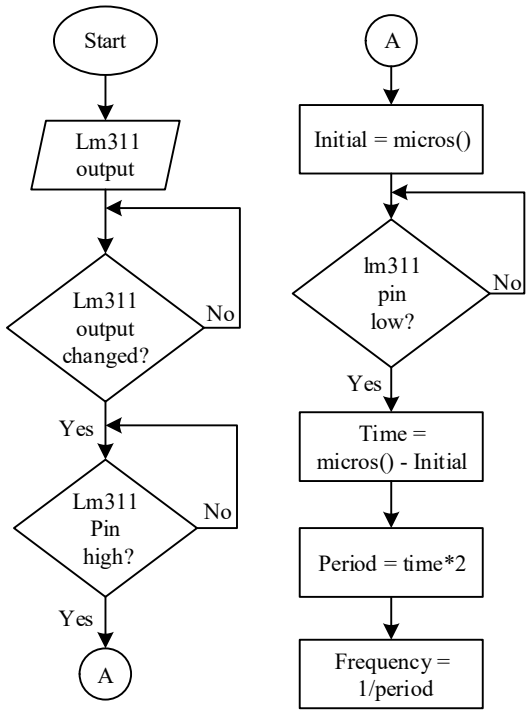


Fig. 10: A flow chart of frequency measurement

V. Current Measurement

A current transformer (CT) of 1:24 turns ratio is used to obtain the high value of load current. As the output of CT is a current signal, it converts the current signal into a voltage signal. The required conversion is done by adding a 100-ohm resistance at the secondary windings of the CT. The converted signal is applied to the full wave rectifier circuit. Arduino reads the voltage in the similar way used in voltage measurement and calculates load current by using the following equations:

$$V_{DC} = \frac{ADC\,reading \times 0.00488}{0.4166} \tag{9}$$

$$V_{RMS} = \frac{\pi \times V_{DC}}{2\sqrt{2}} \tag{10}$$

$$I_{RMS} = \frac{V_{RMS}}{100} \tag{11}$$

$$I_{load} = I_{RMS} \times N_{CT} \tag{12}$$

Here,
I_{RMS} = RMS value of AC current at CT secondary
I_{load} = Load current
N_{CT} = Transformer ratio of CT

VI. Frequency Measurement

The transmission line frequency is not exactly 50Hz in an ideal case. It may vary between 49Hz to 51Hz. Thus, for better accuracy, the measurement of frequency is needed rather than assume it. The output of the comparator circuit is directly fed to the Arduino. Arduino then measures the on-time duration

of the pulse output from the comparator circuit using pulseIn() function. The time period of the original sinusoidal signal is twice the on-time duration of that pulse. This duration is then converted into frequency using the below equations:

$$T = t_{on} \times 2 \tag{13}$$

$$f = \frac{1}{T} \tag{14}$$

Here,
t_{on} = On time duration of comparator pulse output
T = Time period
f = frequency



Fig. 11: Frequency measure using frequency counter

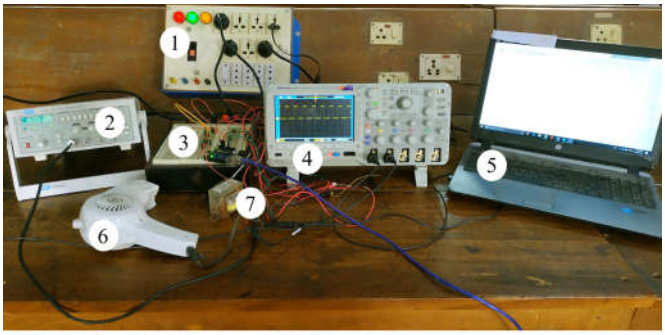
A flow chart of the frequency measurement is shown in Fig. 10. The flow chart is similar as the power factor measurement except that the input of the programme is given from the output of the comparator. After measuring the on time of the comparator output as discussed in the power factor measurement section the frequency is then calculated from the equation (13) and (14). The actual measured frequency of power line is presented in Fig. 11.

VII. Experimental Setup and Performance Analysis

This section presents the experimental setup for measuring the power factor as shown in Fig. 12. The setup consists of a power board, frequency counter, computer, oscilloscope, potential transformer and load. A 200W hair drier is used as an inductive load which contains a motor that shows dynamic behavior along with the time. The function of frequency counter is to measure the actual frequency of main line. A computer is used for plotting the power triangle curve with the help of MATLAB that provides the support of analyzing of power factor. The wave shape of voltage, current, and time difference are observed in an oscilloscope. The performance of proposed system is evaluated by comparing its output and the measuring instrumental output.

A. AC power measured by the measuring device

The power factor, frequency, and the AC power are first measured from the instrumental data. The performance of the proposed system with an inductive load is verified in Fig. 13 and Fig. 14. Here, the channel-1 (yellow wave) presents the voltage wave and channel-2 (blue wave) presents the current wave. From Fig. 14, it is seen that the time delay between voltage and current is 2.96 ms. The frequency counter gives the accurate line frequency value of 50.41 Hz as shown in Fig. 11 and the phase difference between voltage and current is 53.716° obtained from (4). After collecting the instrumental data and finishing some calculations, the system reaches the following parameters presented on table I.



1) Power board
2) Frequency counter
3) Project board
4) Oscilloscope
5) Computer
6) Hair drier
7) Potential transformer

Fig. 12: Experimental set up of the system

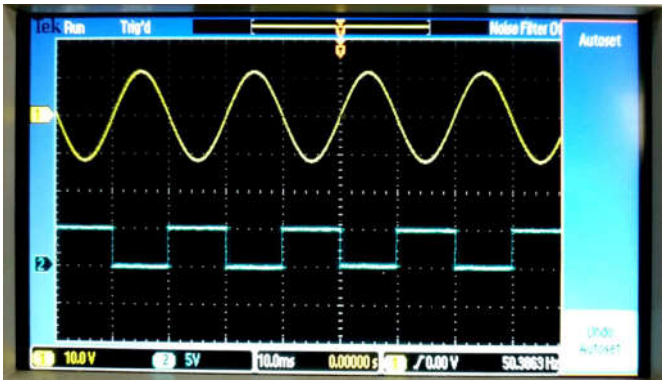


Fig. 13: A generation of pulse wave using high differential comparator



Fig. 14: An experimental output of the proposed system

TABLE I: Parameters measured from the instrumental data

Parameter	Measured value
Power factor	0.598
V_{load}	200 V
I_{load}	1.08 A
Active power	129.17 W
Apparent power	216 VA
Reactive power	173.12 Var

B. AC power measured by the system

The results of the system are displayed on the computer screen. After calculating the AC power, MATLAB plots a power triangle which is shown in Fig. 15. Here, The green, blue and the red line presents the active, apparent and reactive power respectively. The other calculated parameter is displayed

on the MATLAB command window. In table II the measured value of the parameters are shown.

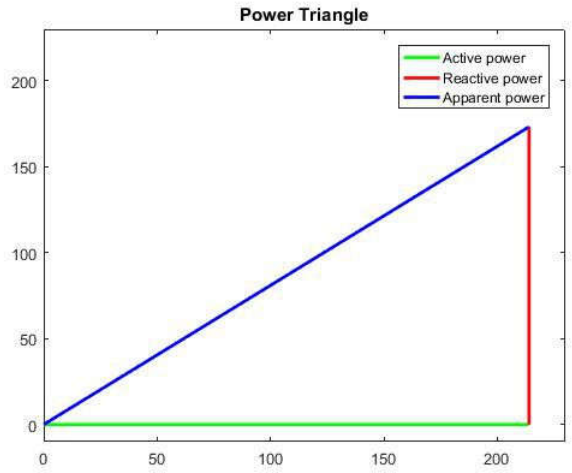


Fig. 15: A triangle power curve using MATLAB

TABLE II: Parameters measured by the system

Parameter	Measured value
Power factor	0.590
Active power	126.25 W
Apparent power	213.90 VA
Reactive power	173.23 Var

After comparing the results, it is observed that the proposed system is able to ensure the required power needed to the load. The proposed system also meets the accurate utility charge within the reasonable range by calculating the appropriate power factor.

VIII. CONCLUSION

A design of high differential strobe terminal based smart power factor metering system is presented in this paper for measuring the power factor of an inductive load. This is done by measuring the load voltage, current and time delay between the voltage and current. In many cases, the variation of line frequency limits the calculation of accurate power factor that may result in higher utility charges. The proposed metering system is able to measure the accurate line frequency and find out the appropriate power factor corresponding to the load. The experimental results are used to proof the validation of proposed system. The obtained result confirms the high effectiveness of proposed system.

REFERENCES

[1] W. M. Grady and S. Santoso, "Understanding power system hannonics," *IEEE Power Engineering Review*, vol. 21, no. 11, pp. 8–11, 2001.

[2] M. Oommen and J. L. Kohler, "Power factor and power factor control alternatives for mines," in *Industry Applications Society Annual Meeting, 1988., Conference Record of the 1988 IEEE*. IEEE, 1988, pp. 1209–1215.

[3] M. B. Khan and M. Owais, "Automatic power factor correction unit," in *Computing, Electronic and Electrical Engineering (ICE Cube), 2016 International Conference on*. IEEE, 2016, pp. 283–288.

[4] P. Machado, T. Abud, M. Fortes, and B. Borba, "Power factor metering system using arduino," in *Power Electronics and Power Quality Applications (PEPQA), 2017 IEEE Workshop on*. IEEE, 2017, pp. 1–6.

[5] S. Saha, T. Tyagi, and D. V. Gadre, "Arm (r) microcontroller based automatic power factor monitoring and control system," in *India Educators' Conference (TIIEC), 2013 Texas Instruments*. IEEE, 2013, pp. 165–170.

[6] R. Dhameliya, K. Domadiya, P. Miyani, H. Savaliya, and P. Jariwala, "Automatic power factor control using arduino uno," *International Journal of Advance Engineering and Research Development (IJAERD)*, vol. 4, no. 4, pp. 2348–4470, 2017.

[7] L. Voltage Comparator, "National semiconductors, texas instruments, literature number: Snosbj1c," 2012.

Converter based Frequency Adjustment and Protection of Grid-tied Wind Farm

Kanij Ahmad
Department of Electrical & Electronic Engineering
Daffodil International University
Dhaka, Bangladesh
Chittagong University of Engineering & Technology
Chittagong, Bangladesh
kanji.eee@diu.edu.bd

Nur Mohammad
Department of Electrical & Electronic Engineering
Chittagong University of Engineering & Technology
Chittagong, Bangladesh
nur.mohammad@cuet.ac.bd

Muhammad Quamruzzaman
Department of Electrical & Electronic Engineering
Chittagong University of Engineering & Technology
Chittagong, Bangladesh
qzaman359@cuet.ac.bd

***Abstract*— Frequency adjustment and protection of a wind turbine due to wind gust is a salient aspect of the grid-tied wind farm. Both of them require conserving system stability. This paper presents an investigation of the converter based frequency adjusting method and wind turbine protection. A grid-tied wind farm model using the Doubly Fed Induction Generator in MATLAB is used. To regulate the active power output, synchronization between wind speed and wind turbine speed is adjusted. Thus, wind farm operated at grid frequency and maximize turbine output. The control strategy of the converter based frequency synchronization of a grid-tied wind farm also includes protection subsystem. The protection system executed by receiving information from the logical block implemented in the wind farm model. It provides wind turbine protection by terminating wind turbine from a grid in abnormal wind conditions. Simulation result validates the analysis and control methods.**

Keywords—Grid-tied wind farm, DFIG (doubly fed induction generator), frequency regulation, wind turbine protection.

I. INTRODUCTION

A huge amount of electricity is required to meet up the growth of electricity demand due to the industrial revolution and modern civilization. The conventional power generation sources such as thermal, nuclear power, hydroelectric, geothermal are used throughout the world. Most of these power plants create harmful effects on the environment. All though we can't avoid these power plants totally, nonetheless it is possible to minimize those energy usages. Optimal generation mix from different sources is considered to be one of the key elements for economic growth of a country [1]-[3]. The worldwide energy policy concerned with the utilization of renewable energy resource [4], [5]. By combining the traditional energy and renewable resources, an optimized, low-cost and environment-friendly power systems can be planned. Among renewable, the wind energy's share sharply rises, and it's growing penetration demands for flexible supply resources and various options [7].

At the present time, wind energy harvesting technology by using wind turbine improves significantly than in the past. Yet the integration of wind power in the exciting power system creates a few technical challenges, power quality issues are prime among them. The term "power quality" means the stability of the voltage, the stability of frequency and the non-existence of various forms of electrical commotion (e.g. fluctuation or harmonic distortion) in the power grid [8], [16]. Combining wind generators with the existing power system could lead to many disturbances like voltage fluctuations, harmonics, flickers and frequency instability [9]. The interconnection of a large number of wind turbines to the grid creates huge trouble to the stability of power system [10], and the uncertainty associated with the uncontrolled wind system is disparaging to the power system operation as well as a control mechanism.

The remainder of the paper is prepared as follows: Section II presents the problem associated with grid-integrated with the wind power system. Frequency adjustment technique is explained in Section III. Section IV provides the intelligent converter control mechanism followed by a protection mechanism in Section V. The simulation results for several plausible scenarios are shown in Section VI. Finally, concluding notes are delivered in Section VII.

II. PROBLEM WITH GRID CONNECTED WIND SYSTEM

Among various renewable sources, wind energy is the speedily expanding energy source, which is regarded as the promotable future of renewable energy [16]. By the end of 2018, the comprehensive capacity of all wind farms installed across the world reached 600 GW, according to a report of the world wind energy association. The rate of extension in the installation of wind-power generation capacity has increased fairly, steadily over the years. So addressing the problem associated with the wind generation system is the real challenge for modeling of a grid-tied wind farm.

When the wind turbine assists a large proportion of the grid load, an irregular change in wind speed can cause voltage inequality, frequency fluctuation and flicker effects on the healthy operation. When a squall of wind hits a wind turbine blade, the blades struggle to speed up because the synchronous generator is locked to the speed of the power grid. In this situation, a huge amount of forces is developed in the gearbox, hub and generator to synchronous turbine shaft with power grid frequency. But it causes damage to the whole coupling mechanism. On the other hand, uncontrolled wind speed causes frequency fluctuation and damage of the structure for a grid-tied wind turbine system. This causes instability in the power system.

III. FREQUENCY ADJUSTMENT TECHNIQUES FOR WIND ENERGY CONVERSION SYSTEM

A grid linked wind farm system always tries to keep the system frequency close to 60/50 Hz. When generation and demand unbalance because of variable wind speed, the

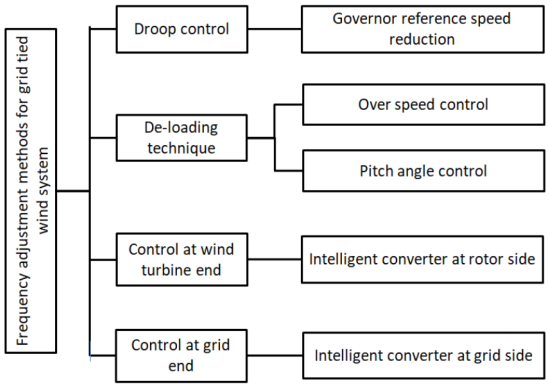


Fig. 1 Frequency tuning methods for a grid-integrated wind power.

system operating frequency starts to decrease based on system inertia and unbalanced power, according to the swing equation of power system stability given in (1) [13].

$$\frac{df}{dt} = \frac{f_0}{2H_{sys}S_G}(P_m - P_e) \quad (1)$$

Where,$\frac{df}{dt}$ denotes the rate of frequency change, $f_0$is the system frequency, H_{sys}stands for the overall system inertia parameter,S_G indicates the rated power of the machine. The terms P_m and P_e are refers respectively to the mechanical power and electrical power [13]. [14].

In order to minimize the negative impact of wind power penetration on system frequency, various inertia and frequency control techniques imply for wind energy conversion system [15], [16]. Fig. 1 elaborates the techniques which apply to adjust the frequency of a grid tied wind power system.

IV. INTELLIGENT CONVERTER FOR FREQUENCY ADJUSTMENT

With the speed change of wind, if the turbine, which converting wind power into electricity allowed to speed up immediately, the stress can be lower on the turbine. Themechanical power from wind gust can be transformed to usable electricity. Most of the wind turbine intended to operate at its standard speed to maintain synchronization of frequency between wind farms and the grid [20]. The frequency of a system linearly increases with the decrease of active power and vivaversa. Any deviation between grid frequency and wind turbine frequency can minimize by active power control of both sides.

A doubly fed induction generator (DFIG) with insulated gate bipolar transistors (IGBTs) based four-quadrant ac-to-ac converter is the best solution for a grid-tied wind turbine system. Fig. 2 presents a model of grid-tied wind farm.Doubly fed electrical generators operate similarly like an AC electrical generator, but the added feature of a DFIG is it can run at speeds to some extentoverhead or lower than its synchronous speed. The converter at rotor side and grid side of a doubly fed induction generator (DFIG) based wind turbine basically regulate the active power which eventually adjusts frequency variation of a grid-tied wind farm [21] – [24]. The four-quadrant converter in the rotor circuit of a DFIG enables decoupled control of active and reactive power of the generator.Fig. 3 illustrates the view of the converter

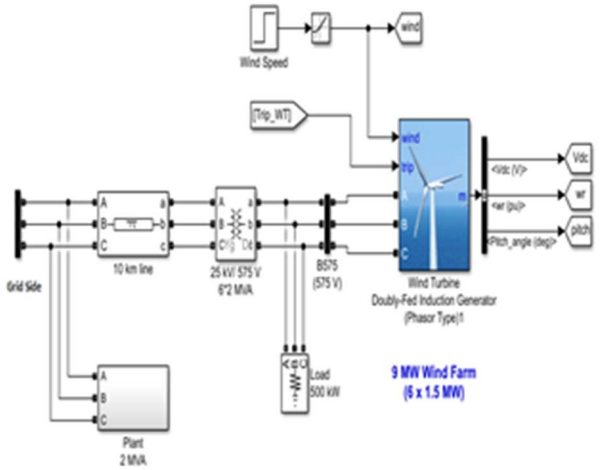


Fig. 2 Grid connected wind farm Simulink model.

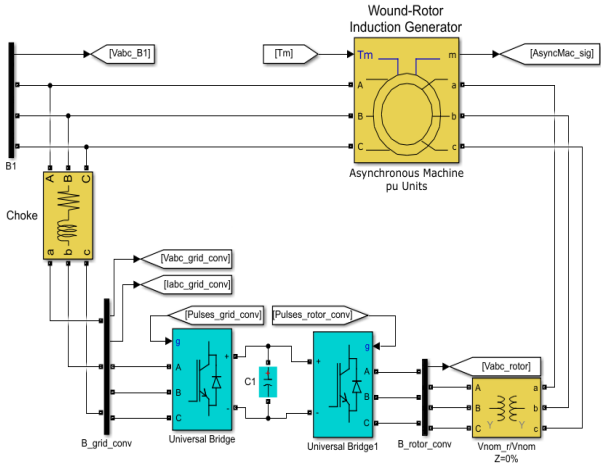


Fig. 3 Doubly Fed Induction generator (DFIG).

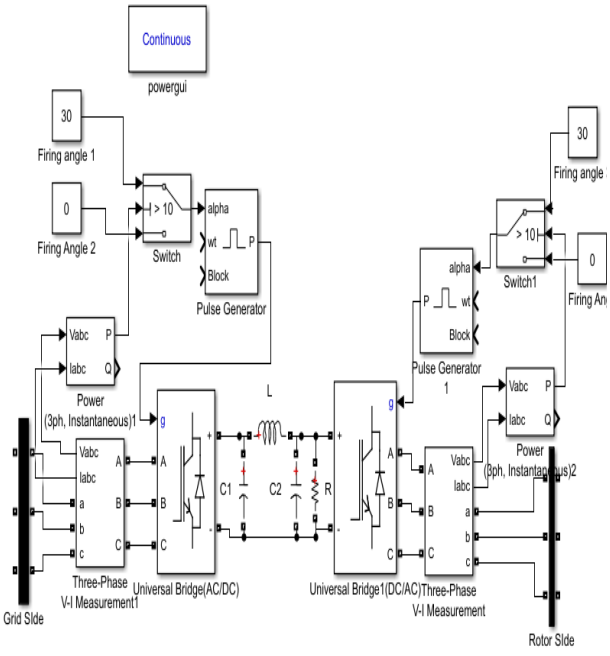


Fig. 4 Back to Back converter (AC/DC/AC) circuit

controlled Doubly Fed Induction Generator (DFIG) [25].The converter associated with DFIG is divided into two components: the grid-side converter (C_{grid}) and the rotor-side converter (C_{rotor}). Fig. 4 shows a converter circuit for controlling both the grid side and rotor side active power.

C_{grid} and C_{rotor} are force commutated Voltage-Sourced Converters that employ power electronic devices (Power MOSFET, IGBTs) to synthesize an AC voltage from a DC voltage [26], [27]. The DC side capacitor acts as the DC voltage source. The coupling inductor, L employs to connect $C_{grid}$to the grid. The brush helps to connect the three phase stator winding to the grid directly and slip ring applies to associate $C_{rotor}$with rotor winding [28].

The induction generator converts the power captured by the wind turbine into electrical power and transmitted to the grid.In Doubly fedInduction generators instead of the field winding is fed by a DC voltage and a stator winding where generates electricity; there are two individual three-phase windings, one stationary and one rotating [29]. One winding is directly connected to the grid and generates 3-phase AC power at the synchronous grid frequency. The other winding generally is known as field winding or rotor winding is connected to 3-phase AC power via a converter to create variable frequency [30]. This input power is adjusted in frequency and phase to compensate for changes in the speed of the turbine. The converter of the rotor side creates the desired magnetic field by controlling its current. The rate of rotation of the output magnetic field is the sum or difference of the rotor rate of rotation and the electric rate of rotation. If the desired frequency 50 Hz, but the shaft of the wind turbine is rotating at 49 Hz, the converter can drive the winding on the rotor with 1 Hz and the output frequency of the generated electricity will be 50 Hz.

In this way, the output frequency will always be in sync whatever the wind turbine speed is. By controlling the rotor currents precisely, the induction machine output power can be controlled. The frequency synchronizes efficiently with the grid frequency while the speed of wind varies [19]. The relationship between stator power and voltage with rotor current shows in (2), (3), (4) and (5) given by [20].

$$P_g = P_s + P_r = -\frac{V_s E_{eq}}{X_{eq}} \sin(\theta - \delta) \quad (2)$$

$$Q_g = Q_s = -\frac{V_s E_{eq}}{X_{eq}} \cos(\theta - \delta) - \frac{V_s^2}{X_{eq}} \quad (3)$$

$$V_s = j\omega_s L_s I_s + V_b \quad (4)$$

$$V_b = j\omega_s L_m I_R' \quad (5)$$

Where

P_s = Stator active Power
P_r = Rotor active power
P_g = Active power of grid
Q_g = Reactive power of grid
Q_s = Stator reactive power
δ = Blade Pitch Angle
V_s = The voltage of the stator
V_b = The back-EMF voltage induced in the stator by rotor current, I_R.
X_{eq}, E_{eq}= Impedance and voltage of DFIG equivalent circuit respectively
L_s,X_s&L_m=stator Inductance and Mutual Inductance

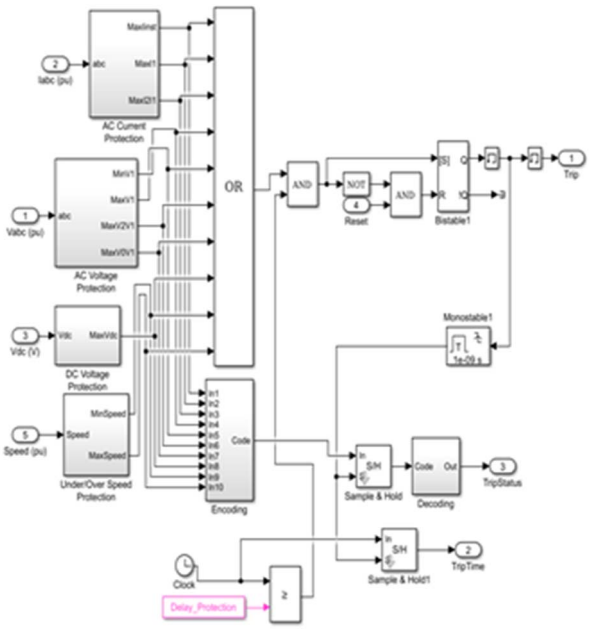


Fig. 5 Logical unit for wind turbine protection.

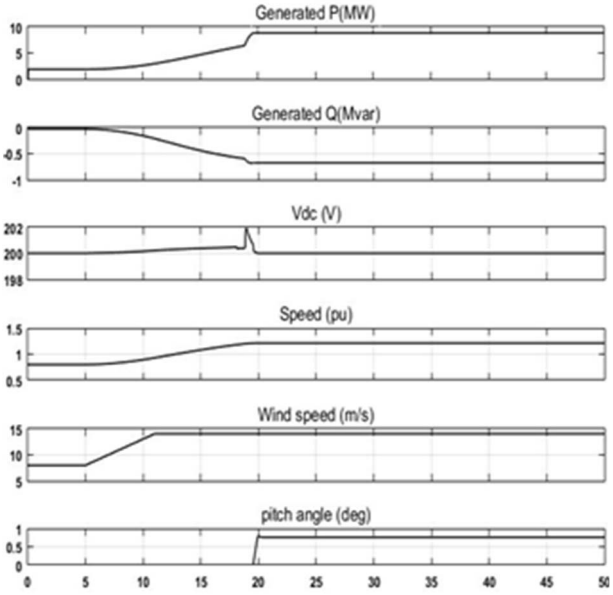


Fig. 6 Turbine response curve for maximum predefined wind speed 15 m/s.

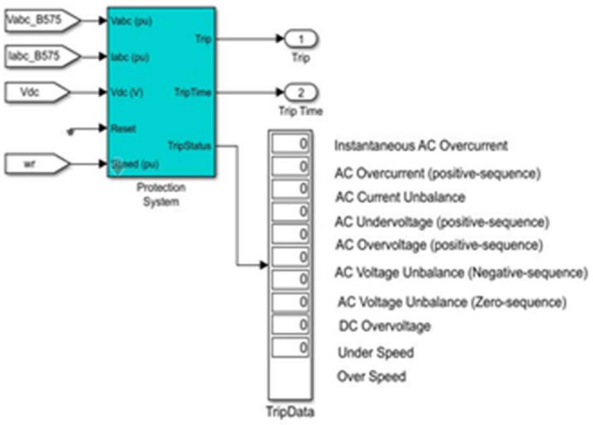


Fig. 7 Wind turbine protection module data (trip status) for maximum predefined wind speed 15 m/s.

The control system induces the voltage control command signals V_{gc} and V_r for C_{grid} and C_{rotor}respectively. By changing the firing angle of power electronics switch, the universal converter can control the power generated from the wind turbine, the reactive power, the DC bus voltage or the voltage at the grid point.

A. Rotor side converter

The rotor-side converter presented in Fig. 4 controls the voltage (or reactive power) and the wind turbine output power. The output power is controlled on the basis of a tracking characteristic which provides a pre-defined power-speed information and maintains the rated output frequency.

B. Grid side control system

The converter of the grid side shown in Fig. 4 regulates the DC bus capacitor voltage. In conclusion, the wind farm empowered a grid converter to produce or dissolve reactive power to adjust the frequency.

V. WIND TURBINE PROTECTION

A trip input is used to provide protection for the wind turbine in an abnormal condition. Fig. 5 indicates a logical unit for wind turbine protection. The block reaches a conclusion about disconnecting wind turbine from the grid on the basis of specific inputs at unexpected conditions.

When any specific input level crosses its reference level, the logical output will be high, both the grid and rotor side are decoupled the wind generator.

VI. SIMULATION RESULT ANALYSIS

The steady state and dynamic performance of a wind farm using The Doubly Fed Induction Generator in MATLAB Simulink R2017a is analysed. The wind farm consists of six 1.5 MW wind turbines. The farm transmits power to a 120 kV grid through a 30 km long 25 kV feeder under 25 kV distribution systems. A 2300V, 2 MVA plant consisting of a motor load (1.68 MW induction motor at 0.93 PF and 500 kW load) is also connected on the 575 V bus of the wind farm.

A. Wind turbine response to a change in wind speed

Initially, wind speed has kept hold at 8 m/s, and then at t=5 Sec., wind speed accelerates to 14 m/s. Waveforms for a squall of Wind (which is in Voltage Regulation Mode) is illustrated in fig 6 where a frequency adjusting converter associate for a grid-connected wind farm. At t=5 Sec., the active power generated from the wind farm elevates smoothly (together with the turbine speed) to reach its rated value of 9 MW in approximately 15 Sec. Over that time period the turbine speed raises from 0.8 pu to 1.21 pu.

At the beginning, the pitch angle of the turbine blades is set to zero degrees. If the wind speed reaches beyond its rated maximum value, the pitch angle is changed from 0 deg to 0.76 deg to limit the mechanical power. Fig. 6 conveys that the pitch angle remains in its initial value zero up to maximum wind speed.

If the wind speed reaches beyond is a predefined maximum value, pitch angle increase with the increase of wind speed to divert the turbine blade out of the wind gust with turbine speed to adjust the system frequency to its

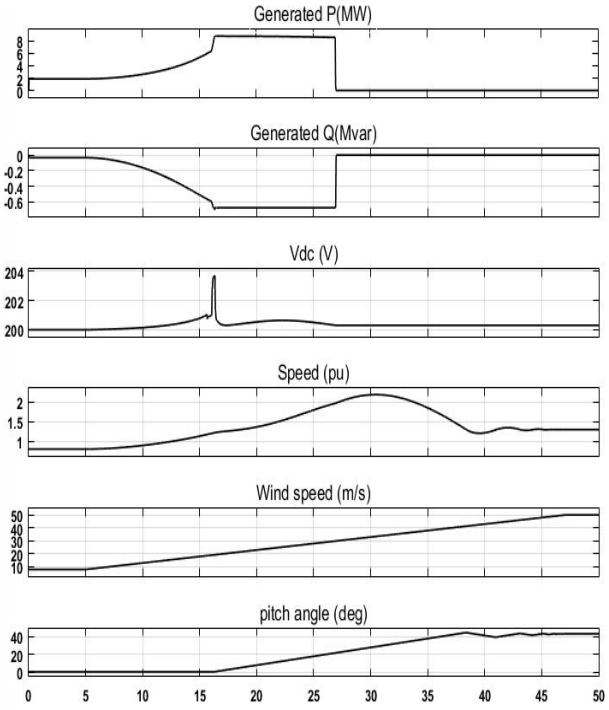


Fig. 8 Turbine response curve for wind speed 50 m/s.

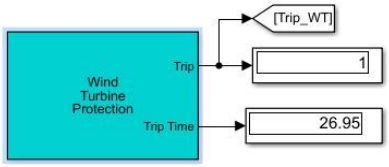


Fig. 9 Wind turbine protection module data (trip time) for over speed wind.

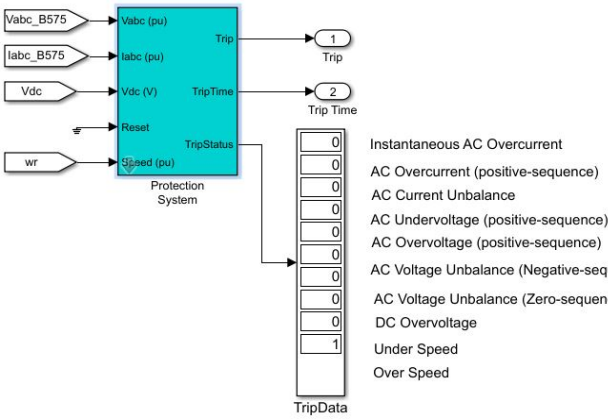


Fig. 10 Wind turbine protection module data (trip input) for over speed wind.

constant value.All the trip status is zero shown in the trip module (Fig. 7) as the wind turbine is operating in between its maximum level.

Fig. 8 implies that the pitch angle of the turbine changes rapidly with the change of wind speed. This is done so to keep the rotor blades at the desirable angle to maximize output in all wind speeds. Blades are angled to protect the wind turbine from wind gust or over speed wind. Fig. 8 also

implies the fact that the power generated from wind turbines is zero when wind speed is over than predefined maximum speed. At that moment trip circuit activated and disconnects the wind farm from the connecting grid.

Fig. 9 and Fig. 10 shows when wind speed (50 m/s) is over than predefined maximum speed (15 m/s), the trip input is high in the wind turbine protection module. The wind turbine trips at 26.95 Sec when the wind speed reached over than maximum speed.

VII. CONCLUSION

In concern to the energy demands and environmental prospect, wind energy is considered as one of the best and clean solutions. But the integration of wind energy in the exciting grid introduces few factors of uncertainty. Interaction of feedback control, intelligent power electronic based controller could improvise frequency adjustment mechanism as well as system instability by ensuring system protection. The analytical study of this paper emphasizes on the fact that a well-designed interfacing of power electronics based converter with grid-tied wind farm can effectively and efficiently synchronize frequency between the power system grid and wind farm by controlling active power and the pitch angle. Also a smart logic control block can protect the coupling structure of grid-tied wind farm in unanticipated wind speed condition. The limitation of this paper is that the other feature of the trip module is not included in the discussion.

REFERENCES

[1] S. Hossain, M. Rahaman, I. Tasnim, and N. Mohammad, “Optimal Energy Mix and Operation Cost in the Presence of Nuclear and Solar PV Generation,” in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1–6.

[2] S. K. Shil, F. P. Polly, M. Z. Islam, and N. Mohammad, “Scenario of Power Generation-Mix in Bangladesh and Australia,” in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1–6.

[3] A. Nath, S. Barua, and N. Mohammad, “Electric Power Generation-Mix for Bangladesh and Its Future,” in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1–4.

[4] M. K. Hasan and N. Mohammad, “An Outlook over Electrical Energy Generation and Mixing Policies of Bangladesh to Achieve Sustainable Energy Targets -Vision 2041,” in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1–5.

[5] M. Barua, S. Mutsuddy, N. Mohammad, and M. A. Razak, “A Survey of Cross Country Generation Mix - Opportunities and Challenges: Bangladesh, Australia, and the U.S.A Perspectives,” in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1–6.

[6] K. Ahmad, N. Mohammad, and M. Quamruzzaman, “Review on Frequency Adjustment for Power Systems with Grid Connected Wind Farm,” 2019 Int. Conf. Robot. Signal Process. Tech., no. 1, pp. 617–621, 2019.

[7] N. Mohammad and Y. Mishra, “Coordination of wind generation and demand response to minimise operation cost in day-ahead electricity markets using bi-level optimisation framework,” IET Gener. Transm. Distrib., vol. 12, no. 16, pp. 3793–3802, 2018.

[8] Sharad W. Mohod and Mohan V. Aware. “Power Quality and Grid Code Issues in Wind Energy Conversion System”.Intech.dx.doi.org/10.5772/54704.

[9] P. Janik, T. Lobos, J. Rezmer, Z. Waclawek, T. Thiringer. "Wind Generator Transients' Computation using Prony Method", 2007 International Conference on Clean Electrical Power, 2007

[10] Tin luu, Adel Nasiri, “output power maximizing of a wind turbine by adjusting rotor speed”, Canadian Conference on Electrical and Computer Engineering, 2008.

[11] S. Mishra, Y. Mishra, Fangxing Li, Z. Y. Dong. "TS-fuzzy controlled DFIG based wind energy conversion systems", 2009 IEEE Power & Energy Society General Meeting, 2009

[12] L. D. Zhang, L. Harnefors, and H. P. Nee, “Power-synchronization control of grid-connected voltage-source converters” IEEE Trans. Power Syst.,vol. 25, no. 2, pp. 809–820, May 2010.

[13] Z.-S. Zhang, Y.-Z. Sun, J. Lin and G.-J. Li. “Coordinated frequency regulation by doubly fed induction generator-based wind power plants”. IET Renew. Power Gener., 2012, Vol. 6, Iss. 1, pp. 38–47.

[14] Ayodele, T.R., and A.S.O. Ogunjuyigbe. "Wind energy resource, wind energy conversion system modeling and integration: a survey", International Journal of Sustainable Energy, 2013.

[15] Sun, HaiShun, Ju Liu, JinYu Wen, ShiJie Cheng, Cheng Luo, and LiangZhong Yao. "Participation of large-scale wind power generation in power system frequency regulation", Chinese Science Bulletin, 2013.

[16] Howlader, Abdul Motin, Naomitsu Urasaki, and Ahmed Yousuf Saber. "Control Strategies for Wind-Farm- Based Smart Grid System", IEEE Transactions on Industry Applications, 2014.

[17] Mridula Sharma, Ajit verma, “A survey report on issues of grid connected wind farm”. International Journal of Science, Engineering and Technology Research (IJSETR), Volume 3, Issue 3, March 2014.

[18] Md. Rabiul Islam, Youguang Guo, Jianguo Zhu. "A review of offshore wind turbine nacelle: Technical challenges, and research and developmental trends", Renewable and Sustainable Energy Reviews, 2014.

[19] D'Rozario, Joan, Sabrina Shams, Shahinur Rahman, Ahmed Sharif, and Enamul Basher. "Cost effective solar-biogas hybrid power generation system", 2015 IEEE International Conference on Industrial Technology (ICIT), 2015.

[20] C. Hamon, K. Elkington and M. Ghandhari, "Doubly-fed induction generator modeling and control in DigSilent PowerFactory," 2010

[21] Mehdipour, Cyrus, Amin Hajizadeh, and Iman Mehdipour. "Dynamic modeling and control of DFIG based wind turbines under balanced network conditions", International Journal of Electrical Power & Energy Systems, 2016.

[22] K. A. Naik, C. P. Gupta. "Improved oscillatory behavior of a grid connected wind farm using IMC PID based pitch angle controller", 2016 IEEE 6th International Conference on Power Systems (ICPS), 2016.

[23] G. Ramtharan, J.B. Ekanayake and N. Jenkins “Frequency support from doubly fed induction generator wind turbines”. IET Renew. Power Gener, 2017, 1, (1), pp. 3–9.

[24] Jie Zhang, Qunjing, Wang,Cungang Hu and Tao Rui. “A New Control Strategy of Seamless Transfer between Grid-Connected and Islanding Operation for Micro-grid”. 12th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2017.

[25] Mohammad Dreidy, H. Mokhlis, Saad Mekhilef. “Inertia response and frequency control techniques for renewable energy sources: A review”. ELSEVIER, Renewable and Sustainable Energy Reviews 69 (2017) 144–155.

[26] Mehdipour, Cyrus, Amin Hajizadeh, and Iman Mehdipour, "Dynamic modeling and control of DFIG based wind turbines under balanced network conditions", International Journal of Electrical Power & Energy Systems, 2016.

[27] Fattah Hassanzadeh, Hossein Sangrody, Amin Hajizadeh and Shahrokh Akhlaghi, “Back-to-back converter control of grid connected wind turbine to mitigate voltage drop caused by faults,” NAPS, November 2017.

[28] Qiu Han, Zhi Qu, Zhi Zhou, Xiaoyang He, Tie Li, Xiaoming Jin, Jinze Li and Zhaowei Ling. “Research on wind power grid-connected operation and dispatching strategies of Liaoning power grid”. IOP Conf. Series: Earth and Environmental Science 121 (2018).

[29] Shao Kang Ma, Hua Geng , Lu Liu, Geng Yang, and Bikash C. Pal “Grid-Synchronization Stability Improvement of Large Scale Wind Farm During Severe Grid Fault”. IEEE transaction on power system, Vol. 33, No. 1, January 2018.

[30] Xiaolong Yue, Dushan Boroyevich, Fred C. Lee, Fang Chen, Rolando Burgos and Fang Zhuo. "Beat Frequency Oscillation Analysis for Power Electronic Converters in DC Nanogrid Based on Crossed Frequency Output Impedance Matrix Model", IEEE Transactions on Power Electronics, 2018.

Design of a Cost Effective Smart Autonomous Power and Switch Controlled Led Street Light

Monir Hossain
Department of Electrical and Electronic Engineering (EEE)
Independent University, Bangladesh (IUB)
Dhaka, Bangladesh.
monir.hossain717191@gmail.com

MD. Tawhid Islam Opu
Department of Electrical and Electronic Engineering (EEE)
Independent University, Bangladesh (IUB)
Dhaka, Bangladesh.
opuasha@gmail.com

Mohaiminul Islam
Department of Electrical and Electronic Engineering (EEE)
Independent University, Bangladesh (IUB))
Dhaka, Bangladesh.
mi.ratul2@gmail.com

Navila Rahman Nadi
Lecturer, EEE Department,
Independent University, Bangladesh (IUB)
Dhaka, Bangladesh
navila@iub.edu.bd

Abstract – This paper implies and demonstrates a unique Smart Autonomous power and switch controlled Led Street Lighting System for Bangladesh. As we live in the age of globalization where streetlights are very much important for a city infrastructure in the crowded regions, the autonomous power and switching of these streetlights can have a huge impact in our electricity usage and minimize the wastage of electricity. Usually street light is needed at the time of sunset until the sunrise, but during peak hours (time 17:00 pm to 20:00 pm), the electricity demand is higher than electricity supply from the National Grid for many under developing countries including Bangladesh as this is the high time of excessive usage of Home Loads like Television, Air-Cooler etc. Proposed herein is a design and practically implementation of a smart autonomous power and switch controlled street light system by using stored solar energy during peak hours and light dependent resistor based energy efficient automatic Streetlight switching rather than the conventional switching operations. Besides temperature sensor and Real time clock (RTC) are being attached to display the updated time and temperature to the nearby passengers of the vehicles and walkers of the street sides. Thus, proper implementation of the proposed system will be able to be a medium for an energy efficient circuit, national grid's power scarcity minimizer and free cost of energy and high reliability.

***Index Terms* - Relay operation, Light Dependent Resistor (LDR), LED Street light, Temperature sensor, Real Time Clock, ARDUINO UNO.**

I. INTRODUCTION

In this modern globalization, Power Scarcity Minimization is the biggest challenge and responsibility whereas natural coals and fossil fuels are at empty line. Lighting is one of the biggest power consumption load and most expensive burden for the under developing countries like our Bangladesh. Recent studies shows that one of the most expensive and vital responsibilities of a city which accounts almost 10% to 38% of the dissipation on electricity [1]. A critical and crucial duty of public officials is Street lighting which has a strategic impact on the stability of social and economy for under developing and developing countries. 8 to 10% of the whole lighting occurs outdoor lighting including traffic lights, street lights and vehicle parking lights [2].

Some of the developing and under developing countries have several outdoor lighting studies,for example, India's central public officials have found the estimation in 2010-2011 of power consumption for public lighting to be about 6.7 TWh annually [3]. Indonesia's officials found on average 15 to 30% consumption occurs for street lighting [4]. Whereas, Libya's street lights causes about 50% power dissipation of the total lighting consumption [5].

In Bangladesh the total road length is mentioned as 4,21,117 km in Local Government Engineering Department (LGED) sites.[6-7] The data shows a diversity of 3,69,999.95 km of rural roads, 30,000 km city roads and 21,118 km of Highways. As a result, approximately 60 to 80 lac street lights are needed to be lighten up the streets of Bangladesh every day. Until now 200-250 watt Pressure Sodium Vapor Lamp (HPSV) are being vastly used in the street of Bangladesh where around 5.475 TWh-7.120 TWh electric energy is needed in every year which is 3.67% to 4.8% of total generation [8] This implies almost 23% to 39% of the total lighting consumption of this street lights considering as per wattage and one light per 100 m road length.

TABLE 1: COMPARISON BETWEEN HPSV AND LED LAMP

Lamp Type	Wattage (W)	Efficiency (lm/W)	Life Span (Hours)
HPSV	**215-250**	**50-150**	**15000**
LED	**15**	**70-160**	**50000**

To get such assumptions the annual Cost of Street lighting is calculated using equation (1) where C is the total annual cost, W is the wattage, H refers the daily working hours, P is the price of electricity in KWh, N is the number of Street lamps.

$$C = (W \times H \times P \times 65)/1000 \quad (1)$$

Power Savings is calculated by equation (2) where S is the Power saving variable in percentage, U1 is the previous lamp's energy dissipation, U2 is the lamp's energy dissipation.

$$S = ((U_1 - U_2)/U_1) \times 100\% \quad (2)$$

Annual Energy Consumption is calculated through equation 3 where E is the annual energy dissipation,

$$E = (W \times H \times 365 \times N)/1000 \quad (3)$$

Battery Requirement (AH) to be calculated by equation (4) where T is battery back up hour, V_b is Battery voltage and Pf is power factor of the source.

$$A_H = (W \times T)/(V_b \times P_f) \quad (4)$$

Minimum charging current required for Battery I is measured by equation (5) where Ta is active Solar Radiation Hours,

$$I = \frac{A_H}{T_a} \quad (5)$$

Required Panel Watt-Peak Wp is measured by equation. (6) as follows:

$$W_p = (W \times T)/T_a \quad (6)$$

Required Solar Panel wattage as Ws is measured by equation (7) where η_s is the efficiency of the System and P_v is the efficiency of the Photovoltaic Solar Panel,

$$W_s = W_p/(\eta_s \times P_v) \quad (7)$$

On the other hand, If LED 60 Watt Street Lights are being used instead of HPSV, 1.314 TWh to 1.71 TWh electric energy will be needed in every year which is 0.88% to 1.15% of total generation power and 13 to 29% of the total lighting dissipation. The two comparisons represents a 4.161 TWh-5.41 TWh energy saving (and 76%) in every year just by using LED instead of HPSV lights. So a crucial conclusion can be provided by the data that for the current scenario of Bangladesh, the usage of LED lights can be the best solution for the street lighting because of the highest energy saving and efficiency.

II. Proposed Design and Hardware Architecture

Improper lighting is dangerous for public transportation where as traditional lighting infrastructures of streets can be mostly improper with huge financial and energy drain. Fortunately, there's a proper, smarter and cost-effectively designed street lighting infrastructure proposed in this paper with a very simple circuit and several feasibilities that will operate automatically on the street lighting on/off with relay coupled. It enhances well enough to find out objects nearby.

TABLE 2. Utilized Components for the Design

Component's Name	Specification
Arduino Uno	Version R3, 5watt, 7-12 volt
LM35	3 pin, 5volt, 10 mv/°C
LDR	1.8 kΩ - 4.5kΩ, 600 nm
RTC	DS1307
5V Relay Module	10 A, Active low
Small LED Strip	60 watt
Battery	12 V
Potentiometer	10KΩ
Resistor	10KΩ
Breadboard	
80 watt Solar Panel	
Charge Controller	
DC Adapter	230 V AC to 12 V
LCD Display	16×2

This circuitry and hardware is very easy to operate and it permits low power dissipation as very few numbers of components have been used.

A. Arduino Uno

Arduino Uno R3 is an ATmega328 AVR based microcontroller board which offers DIP (dual-inline-package) and removable IC with 20 digital i/o (Input/Output) pins. Out of them 6 pins can provide PWM outputs and another 6 pins can have analog inputs. Programs using C++ language provided by IDE platform can easily be loaded into it. The programs can be loaded Computer USB hub. Arduino has an openly extensive supportive community, which is a open source platform for everybody. So it is very easy for the embedded electronics based projects to get started working.

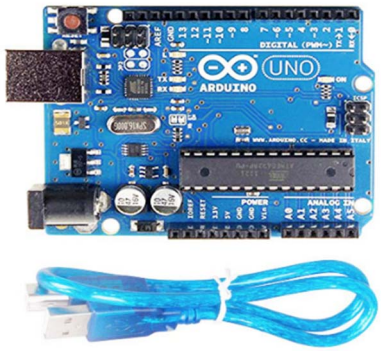


Figure 1: Arduino Uno R3 [9]

It has 16 MHz resonator and 16k flash memory. It consumes too little power and it has the feasibility to use pins for any projects. Sensors can be interfaced with this micro-controller. Data from the sensors can be read easily through ADC (Analog to Digital Conversion) method and based on that data, i/o pins can be controlled as required to any projects. Those pins can be worked as digital signal producer. Whenever the signal is high it means 5V has been provided on the selected pins by the micro-controller or, data has been sent to the output devices. Whenever the signal is low it means the pin is connected with the ground and data can be received. [9]

B. Temperature Sensor-LM35

It is a temperature sensor which is used to measure temperature. It has 3-pin, middle pin works as an output which changes depending on the temperature changes. For every 10MV change in output it indicates 1° C change in temperature.

C. Light Dependent Resistor (LDR)

LDR is a light intensity based sensor whose sensitivity changes depending on the light intensity poured on it. When the increasing strengthen light falls on LDR, the resistance of LDR decreases, and the vice versa.

In darkness when there is not enough light to see anything, the resistance of LDR shows the range of resistance in mega ohms, and when presence of light or in brightness increment occurs then the resistance decreases and shown by few hundred ohms. Based on that resistances variation light intensity measurements are done and decisions are made based on those measurements.

D. Relay Module

Relay is an Electro-Mechanical switch which takes digital or logical signals from various micro-controllers, ICs or chips like Arduino, 8086, AVR, ARM, PIC, DSP, PLC, MSP430, TTL etc. It can be Multichannel and different voltage and current rated. In our case, 5v, 10A 2-Channel active low Relay interfacing was done with the micro-controller board Arduino Uno R3. These relays are used to control any kinds of high current electrical appliances and equipment with safety. These

relays can be controlled through directly with 3.3V or 5V logic output signals from any microcontroller.

E. Real Time Clock (RTC) Module

Real Time Clock (RTC) which is a clock module can provide local time from local time zone. One of the modules of RTC is DS1307 which has 8 pins with an I2C capabilities. The DS1307 consumes very low-power time clock with calendar with SRAM of 56 bytes battery backup. The module provides time as seconds, minutes, hours, and date as day, month, year from local qualified data. Each month's end date is smartly adjusted, for those special months which are less than 31 days. RTC has an integrated circuits (ICs) and it can be easily supervised of timing just like a clock. In addition, it can operate date just like a year calendar. The main feasibility of RTC is that it has a setup of battery backup which helps the module running though there is no power provided externally. As a result, it can keep and continue the clock/calendar. There are many applications of RTC module like integrated and embedded in several systems for example mother boards of computers.

F. Implemented Resistors

Resistor is the most basic passive electrical component with two unidentified terminals that is applied to manage and limit the flow of current in any closed circuit. Current which flows through a resistor is directly proportional to the appearing voltage drops across the resistor. Resistors are of two types –
i) Fixed Resistor – having a fixed resistance value.
ii) Variable Resistor – whose resistance can be varied or changed for example a variable resistor of 8K can be used for the value of resistance from 0 to 8 KΩ
In the proposed design a 10KΩ resistor has been used.

III. FRAMEWORK

In the proposed circuit, Arduino Uno R3 microcontroller has been implemented as the brain of the hardware, which actually reads data from the sensors and takes proper decision based on that data. Arduino reads data from any sensor's voltage which is digitally converted from analog Signal and give an output as numerical format depending which the micro-controller is programmed to take decision independently. Based on those decisions Arduino is programmed to make its output pin high or low.

The very next task of the hardware was done by the Relay which has been used as an active low operational device hereby. This means if any grounding point can be provided then the relay coil will be operational by completing the relay circuit with the combination of magnetic field and mechanical contact. The circuit has two relay operations hereby as there are two power sources which are Battery and national grid.

Additionally, RTC is used to provide time from local time zone and simple interfacing with Arduino Uno. LM 35 temperature sensor is used to read atmosphere temperature and give the projection of daylight. Meanwhile, LCD display has been used to represent the local time and daily temperature to the passerby.

A. Proposed Design

The proposed Smart automatic power and switch control Led Street Light system can be divided in four main parts: The working principal of each part has been described hereby.

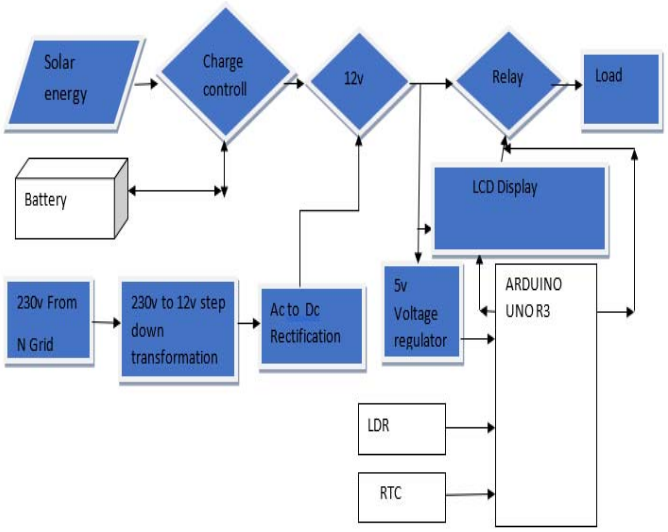


Figure 2: The block diagram of the proposed System

B. Battery and Solar panel setup

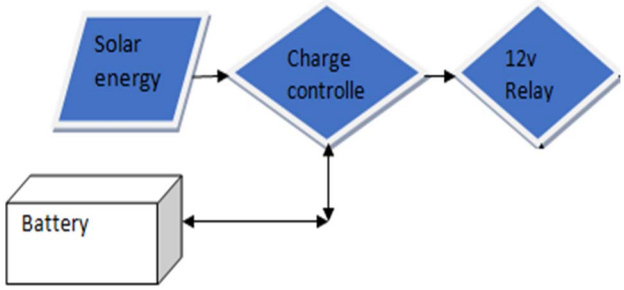


Figure 3: The block diagram of the battery and solar panel setup

During day time the electricity is produced by solar panel then the electricity is stored in battery by the charge controller after that the stored energy active the 12v relay and work as source as well.

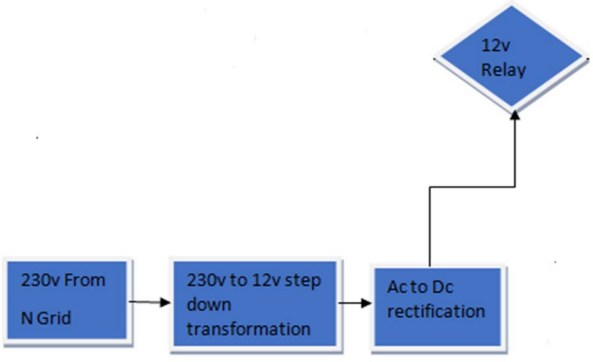


Figure 4: The block diagram of the grid and relay setup

C. *National Grid Connection*

The conventional street light has electricity connectivity which is connected with National grid, then converts 230v AC to 12v AC after that modify 12v AC to 12v DC using rectifier and filter circuit, However the 12v DC always work as backup power, when the battery power is finished then 12v DC will work automatically by the help of relay .Here12v dc relay is used to permit which source will be used, that means output of the relay will act as 12v DC source constantly.

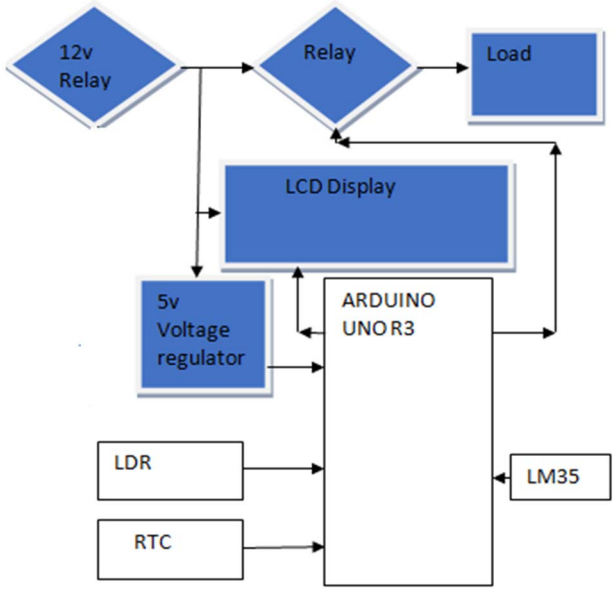


Figure 5: The block diagram of process of voltage regulator

Here, the output of the Relay is considered as the source of the rest of circuit, which is 12V DC. As we know, ARDUINO or Microcontroller needs fixed 5v supply power for long time service that's why we use 7805 voltage regulator to get fixed 5v power for ARDUINO and other appliances.

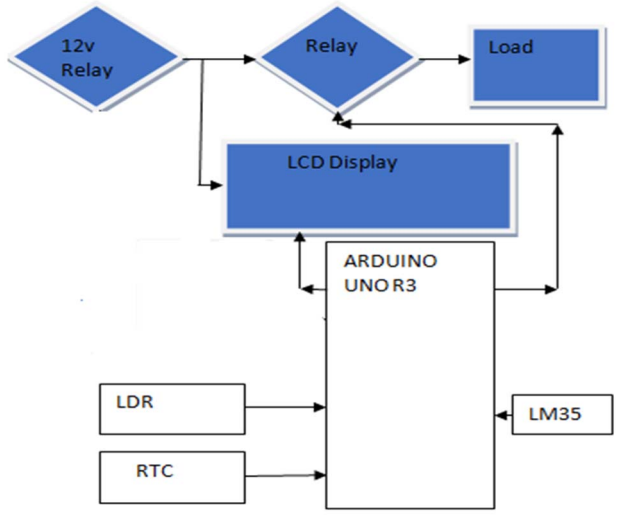


Figure 6: The block diagram of central circuit

ARDUNO Board used as the brain of entire part of street lights. The Light Dependent Resistor (LDR) can sense daylight and dark and send analog signals to the ARDUNO Board, according to ARDUINO processed LDR data the 5v relay will operate switching ON or OFF the load automatically. Temperature sensor senses Current temperature and produce analog also the Real Time Clock (RTC) produce analog signal as well.

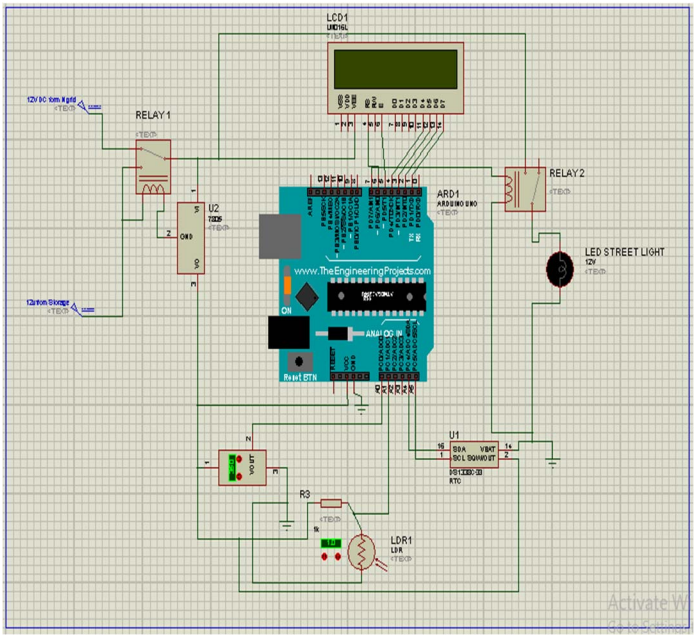


Figure 7: Proteus design of the Smart automatic power and switch control LED Street Light.

According Temperature sensor and RTC data the display will show current time and temperature which helps the people who are traveling or walking on the road.

D. *Experimental Work*

In This project the main focused was on Practical implementation of a cost efficient and highly reliable automatic power and switch control streetlight control system where used two sources of power and display will show local time and temperature. The main difficulty of Current conventional streetlight use only one source either from national grid or storage energy which charge up from solar panel and old switching system is wastage of power.

The project aims are to fabricate and Practical installation a highly reliable Smart automatic power and switch control Led Streetlight system which has two types of sources to reduce energy crisis during peak hours. Usually street lights are needed after sunset to sunrise using either energy directly from national grid or directly stored energy form solar panel (where needed higher rated solar panel and Battery for 6
Hours backup)



Figure 8: Practical installation of the proposed project.

The solar panel converts solar energy to electrical energy presence of sunlight then the electricity is stored in battery by the charge controller after that the stored energy activated the 12v relay and worked as the source as well.

The conventional street light has electricity connectivity which is connected with National grid which converts 230V AC to 12V AC and modifies 12V AC to 12V DC using rectifier and filter circuit. However, the 12V DC always works as a backup power, After peak hours or when the battery power is finished it will shifted automatically by the Relay operation to alternative source which is 12V DC converted energy from national grid. Here12v dc relay is used to permit which source will be used, that means output of the relay will act as 12v DC source constantly.

ARDUNO Uno Board has been used as the brain of entire controlling circuit of streetlights. Here the output of Relay conducted as source of the rest of circuit, which is 12V DC. The control circuit requires fixed 5v DC to operate, that's why we use 7805 voltage regulator to get fixed 5v Dc power for ARDUINO and other appliances.

The present streetlight systems has enormous amount of power wastage due to conventional switching operation. The Light Dependent Resistor (LDR) used to sense daylight, dark and send analog signals to the ARDUNO Board. The Analog signals from the LDR are processed in Analog to digital converter (ADC) of the Microcontroller, A code Written in C language and the program is dumped into the microcontroller, according to LDR data the microcontroller generates signals to activate or deactivate the relay for switching ON or OFF the streetlight automatically. This system is efficient automatic switching control.

Temperature sensor senses Current temperature and produce analog also the Real Time Clock (RTC) produce analog signal as well. The Signals send to the ARDUNO Board. According Temperature sensor and RTC data the display will show current time and temperature which helps the people who are traveling or walking on the road.

IV. Result Analysis

The Smart automatic power and switch control Led Street Lights Are necessary in today's life for safety issues and eliminate unexpected accidents during night. Providing street light is important and more expensive of any city. Usually street lights are needed after sunset to sunrise using either energy directly from national grid or directly stored energy form solar panel (where needed higher rated solar panel and Battery for 12 hours backup)

It is being known that during peak hours (time 5pm to 10pm) the electricity demand is higher than electricity supply of many countries for instance, Bangladesh. This project is designed for two types of sources; these are stored energy from solar panel and energy from national grid. During peak hours the street lights will run by stored energy from solar panel and after peak hours it will shifted automatically by Relay operation to alternative source which is DC converted energy from national grid. As a result this system can reduce energy crisis during peak hours also utilize power of national grid by street lights after peak hours (because after peak hour energy is less). However the solar panel and Battery size will reduce using this idea.

From the proposed design the annual Consumption of Energy is found to be 131.4 KWh through equation03 where W=60 watt LED Bulb, H=6 hours for a single lamp during the pick hour when demand is high. This amount of energy is provided by our system which is from the solar energy. Battery Requirement is found as 30 AH by equ-4 where T=6 hours, Vb=12 volt and Pf =1 as dc power source.

Charging current (minimum) required for 30 AH Battery I is calculated as 4.29 Amps. by equation 5 where considering Ta =7 Hours

Required Solar Panel Peak wattage Wp is measured as 51.5 watt by equ. (06) and Rated Solar Panel wattage, Ws is found 80.5 watt by equ. (07) where considering ηs =80% and ηpv=80%

Power Savings of national grid by our system is calculated through equ-02 where U1 =213.6 KWh and U2 =131.4 KWh. We found 38.5% power savings and minimization of pick demand may occur just because of our proposed hardware And 76% power savings can be done just using the LED light.

Annual Cost minimization for a single Street lighting is calculated through equation (1) and found that worth of 1314 BDT can be saved through our system [P=10 BDT in pick hours in KWh].

In large scale for example 1 lac street light, 13.14 GWh energy demand can be minimized and around worth of 130 million BDT can be retained just for using free solar energy.

A. Cost Analysis

The costing is done for single Street lighting System. In large scale installation the costing would be 70 to 80 percent less.

Single Street Lighting system's initial installation cost can be retrieved in 6 years approximately through the service we are getting from it.

TABLE 3 : COST EVALUATION OF THE PROPOSED SYSTEM

Sl	Equipment name	Price (per unit) BDT	Net Price BDT
1	Solar panel (80 watt)	50	4000
2	Battery 30AH	3000	3000
3	Hardware(Arduino Uno R3, LDR,LM35)	600	600
4	Led street lamp (60 Watt)	300	300
5	RTC Module	100	100
Total costing			8000

Table 2 shows how much we have to invest to produce the system where a power generation small unit also present. On the other hand convention system only consumes power where National Grid is the only power provider and no alternative power generation can be done through it.

B. Findings

The whole analysis and findings are summarized in a table would be helpful to make comparison between our system and the previous system. [10] The table 4 shows a summarized view of the findings of the system where it shows 38% demand minimizing capability of it. In this modern era, where power is a very serious issue this percentage can have a huge impact in the economical and infrastructure of the country. One of the alternative solution of resolving the electricity crisis is by implementing solar plants which gives energy free of cost except plant installation cost which is very less comparing with other power plant.

TABLE 4: COMPARISON BETWEEN CONVENTIONAL AND OUR PROPOSED STREET ELECTRIFICATION

Technology	Annual Energy Consumption		Demand Minimizing of National Grid	Energy Savings
	National Grid	Solar powered Battery		
Previous and Conventional Street lighting system	100%	0%	0%	0%
Our Proposed Street Lighting System	40%	60%	38.5%	76%

Table 4 shows how much Energy burden on National Grid can be reduced and energy savings can be done through our proposal in practically over conventional street lighting system. Above comparison and benefits shows that our system will be very handful for our motherland Bangladesh. Both people and government will be benefited highly which shows that Our Street lighting system is the best option for our Bangladesh. Not only that there are some features which makes our project proposal more useful and those feature can be operated with Zero Cost.

VI. CONCLUSION

As being a part of the resident of developing country, some necessary initiatives should be taken as soon as possible to solve the most challenging fact of power crisis of this country. In this aspect the proposed Smart Autonomous Power and Switch Controlled Led Street Lighting system will be the first initiative to overcome the situation. Our Government and Officials of Electrification Board of Bangladesh should think about the present situation of Power generation and be cautious about using the electricity for any purposes. As it has been represented here that the system is helpful and enormously beneficial for the power crisis, it can be implemented practically with further advancement to make it applicable in the Street Lighting Sector by which power crisis of our motherland can be slightly solved by minimizing the power loss and applying the renewable solar energy in this sector.

VI. REFERENCES

[1] Johnson et al., "Energy Savings Potential for Street Lighting in India," Ernest Orlando Lawrence Berkeley National Laboratory, Rep. LBNL-6576E, 2014

[2] Mary Ashe at all, "2010 U.S. Lighting Market Characterization ," Navigant Consulting, Inc., Final Report, 2012

[3] "All India Electricity Statistics," Central Electricity Authority, India, 2010-2011

[4] Tempo. " A working meeting minister sudirman call the electricity crisis".[online] Available at tempo.co/read/news/2014

[5] Ashraf Khalil, Ali Ashcibc, "The Chances and Challenges for Renewable Energy in Libya". In the Proceedings of the Renewable Energy Conference, Palermo, 2015, pp. 1-6

[6] List of roads in Bangladesh, https://en.wikipedia.org/wiki/List_of_roads_in_Bangladesh

[7] LGED Bangladesh, retrieved from http://www.lged.gov.bd/ViewRoad2.aspx

[8] Imtiaz Mahmud, Economic Feasibilty Study of Solar Powered Street Lights in Bangladesh" National Conference on Renewable Energy, At Khulna University, Khulna, Bangladesh

[9]A. Jain and C. Nagarajan, "Efficient Control Algorithm for a Smart Solar Street Light," 2015 9th International Conference on Next Generation Mobile Applications, Services and Technologies, Sep. 2015.

[10]Bangladesh Power Development Board, Retrieved from http://www.bpdb.gov.bd/bpdb/index.php?option=com_content&view=article&id=360&Itemid=170

Analysis of Air Foils and Design of Blades for a Low-Speed 250W Vertical Axis Wind Turbine Suitable for Coastal Areas of Bangladesh

Mahtab Murshed
Department of Electrical and Electronic Engineering
Independent University, Bangladesh
Plot-16, Block-B, Bashundhara,
Dhaka-1229, Bangladesh
E-mail: mahtab_murshed@hotmail.com

Md. Yeasin Arafat
Department of Electrical and Electronic Engineering
Independent University, Bangladesh
Plot-16, Block-B, Bashundhara,
Dhaka-1229, Bangladesh
E-mail: nealmagik@gmail.com

M. Abdur Razzak
Department of Electrical and Electronic Engineering
Independent University, Bangladesh
Plot-16, Block-B, Bashundhara,
Dhaka-1229, Bangladesh
E-mail: razzak@iub.edu.bd

***Abstract*—The potential of wind energy in Bangladesh is not so good but has reasonable prospect if modern technology and optimum design of air foils and blades can be used. Coastal regions of Bangladesh including Cox's Bazar, Chittagong, Teknaf, Patenga, Mongla, and Kutubdia, are most suitable points of harnessing wind energy. Although the velocity of wind in these regions is not perfect for harnessing energy from a big sized wind turbine rather it is quite capable to generate electricity from small sized turbine in an efficient way. Vertical axis wind turbine (VAWT) is perfectly appropriate for extracting the maximum power from wind in such low wind speed area. NACA foil is mainly prioritized to build this type of small sized VAWT system because of the low maintained mechanism. The small-sized VAWT system needs more analyzed wind data to work in an optimized condition. The positions and efficiency of the turbine system are highly dependent on the blade lift and drag coefficient because of random wind flow directions. Counting all these points, this paper examines the wind data available in these areas, and describes the design of a vertical axis wind turbine (VAWT) suitable for small sized electricity generation system in an efficient way by optimizing turbine blades and its foil mechanism.**

Keywords—Vertical Axis Wind Turbine; Power Generation; Coastal areas; Blade design; Windpower

I. INTRODUCTION

The electricity generation is increasing rapidly due to it's high demand [1]. Currently, the energy is produced mostly from fossil fuel as its main sources. The use of these non-renewable fuel sources has genuine harmful effects on surrounds. An abnormal amount of fossil fuel uses can cause several environmental damages like ozone layer exhaustion, acid rain, climate change and so on [2]. It is obvious that fossil fuel is decreasing over the period due to excessive use to generate power. To slow the decreasing rate of fossil fuel and make environment more reclamation, renewable energy is the only alternative solution except nuclear energy which has great negative impact and risk for the livable environment. Currently, many countries adopted renewable energy as their primary electricity generation sources over fossil or nuclear because of less or negligible effect on the environment. Besides the renewable sources are quite available and affordable. There is no cost in raw renewable energy sources.

Also, harnessing wind energy across all over the globe is increasing due to raw energy availability and use of a higher grade of optimized technology. Most of the European countries are adopted this energy expansion over other regions of the globe because of huge availability of wind in these areas.

The situation in Bangladesh is quite different where most of the power generation system is fully depended on conventional fuel sources and alternate energy development is not making that much impact over this. But the development in the solar system is really appreciable in renewable sector although the expansion is not fully satisfactory yet [3]. Currently, different approaches are taking place to improve this solar energy and make it more handy and mainstream. To increase renewable energy generation in this country the only focus on the solar system is not quite enough. Finding the availability of other sources like wind, geothermal, tidal, hydro, biomass, etc. and harnessing energy from these sources are relatively challenging [4]. Few small-sized winds or biomass projects are taking place although most of them are only showcasing the structure and lies on idea of the projects, and many of them are not in working conditions. The main reason behind this kind of wasting technology and resources is the lack of adequate analysis and modification of the technology suitable for Bangladeshi weather and environment. In the context of Bangladesh, the wind has a great prospect. Maybe the available wind energy is not capable to harness a huge amount of electricity but it is quite capable to generate in a micro amount in an efficient way [1, 5, 6]. Currently, the annual average wind velocity in Bangladesh is 2 m/s which is quite low to develop the huge size wind turbines all over the country [5, 6]. But the coastal regions and some hilly areas of Bangladesh has a reasonably good average wind speed to demonstrate the wind energy harnessing in an effective way [1]. The availability of electricity is not regular in these hilly or coastal regions. Hence proper utilization of wind energy can make an off grid or on grid network suitable for home, small business model, several types of commercial enterprises and so on [7]. Proper installation of a bunch of wind turbine can also use for harvesting crops, pumping drinking water which can significantly reduce the use of diesel generator [2]. Still different NGO and government do not show proper interest for harvesting this energy due to the absence of appropriate information about wind velocity, mechanism, weather adjustment, and its future opportunities to expand with latest available technology [1, 5]. Hence the available wind velocity is not adequate for normal or high-speed turbine system, only low-speed turbine system is more efficient over the existing available model.

Counting all these obstacles, the focus of this research work is to examine the actual wind prospect in the coastal region of Bangladesh to determine the available energy and analyze of airfoils to develop a vertical axis wind turbine (VAWT) perfectly suited for low-speed wind to extract the available energy with the maximum efficiency. In low-speed small size power generation, VAWT is much smaller and work efficiently over horizontal axis wind turbine (HAWT). The analysis of air foils of the turbine blade is done using QBlade software which is specially design for existing and new type blade foils analysis made by Hermann Föttinger Institute of TU Berlin.

II. Wind Prospects in Coastal Region of Bangladesh

Currently the electricity generation in Bangladeshi grid is quite high but most of them are directly or indirectly depend on fossil fuel [2], although most of the coastal areas are not grid connected. The geographical position of Bangladesh is one of the biggest advantages over other small countries because of comparatively large coastal regions. Figure 1 shows the coastal area geographically where the scope between 20°34'- 26°38'N and the longitude between 88°01-92°4E [2]. Strong wind flows are fairly available in the southern part of Bangladesh where 724 km longest coastline is located with 200 km long hilly areas [6]. The average wind velocity in these areas are recorded 5 m/s or greater at a height of 20-30 meter which has decent potential to feed small-sized wind turbine [6, 8, 9]. Although some projects like Kutubdia island and Muhuri dam has been established long ago but all of them are out of order because of lack of proper knowledge about wind prospect and maintenance. Bangladesh government along with several NGOs do some research earlier where it shows clearly that to generate electricity now in a small amount, the desired wind velocity is available from late March till the beginning of October and the maximum peak will be in June-July [1, 6]. Besides coastal areas, many waterfront areas, for instance, Teknaf, Patenga, Cox's Bazar, Chittagong, Hatiya, Mongla, and Kutubdia can also potentially stable to generate wind energy because of availability good wind velocity [10]. The average wind velocity in these areas are recorded to be precise around 3.5 m/s [10, 11].

III. Analysis of Wind Velocity at Coastal Regions

To find out the optimum energy generation in coastal regions of Bangladesh, at first it is needed to analyze the wind velocity, air density on that areas and maximum power calculation for the desired system carefully.

A. Wind Data Analysis

It is quite important to measure the minimum, maximum and the average wind velocity on that particular area where the wind turbine system is supposed to be installed. Table I shows the monthly measured average wind velocity of some coastal regions targeted for wind turbine installation. The wind velocity is measured at a height of 20 meters from the ground [11]. The average yearly maximum wind velocity has been recorded 5.52 m/s and the yearly average minimum wind speed is 2.44m/s where May and July can be considered as the most prospective months and November to January can be considered as the least prospective month for wind power extraction. This indicates a good amount of wind energy can be harvested and utilized during the monsoon period in the coastal regions of Bangladesh [7].

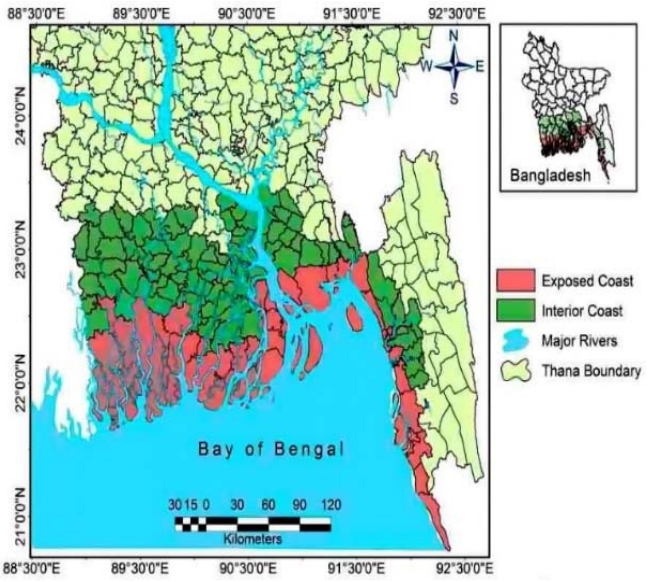


Fig. 1. Wind Prospective Coastal Regions in Bangladesh

B. Wind Velocity Calculation

The wind shear exponent can be used to determine the velocity of wind, v at a reference height, h_{ref} and wind velocity at practice height, h [11–14] as equation (1) and plotted in Fig.2.

$$v = v_0 \left(\frac{h}{h_{ref}}\right)^{\alpha} \quad (1)$$

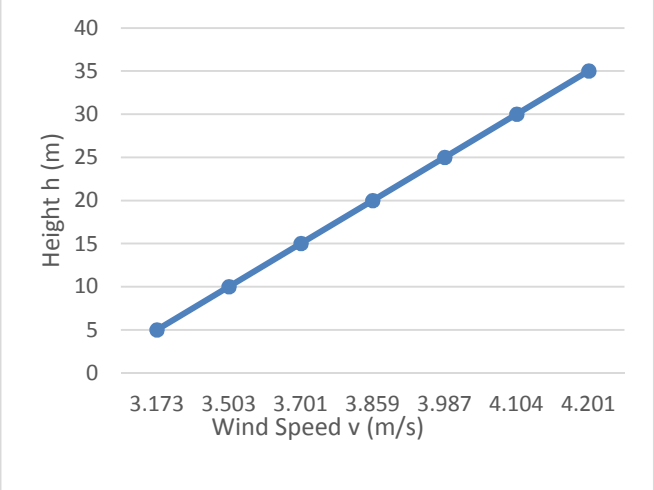


Fig. 2. Height from ground vs. wind speed

C. The Density of the Air Factor

The density of the air, ρ is the mass per unit volume of Earth's atmosphere. Power extraction from a wind's kinetic energy significantly depends on various factors such as air moisture and air density. Here the density of the air has been considered since it has a great impacts during power extraction. Within a settled atmospheric pressure, a list of the density of the air at various temperature is given in Table II [15]. In this research work ρ =1.16 kg/m^3 is considered as the average coastal temperature is 30℃ in Bangladesh.

Table I. Wind Speed (m/s) at a height of 20m in Various Coastal Areas of Bangladesh

Locations	Monthly wind velocity (m/s)												Yearly Maximum	Yearly Minimum	Yearly Average
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec			
Chittagong	3.65	2.87	4.94	5.02	5.52	6.88	7.08	6.85	4.63	2.81	3.38	2.21	7.08	2.21	4.66
Cox's Bazar	3.76	3.86	4.53	5.56	3.84	4.13	3.84	3.96	3.21	3.27	2.58	3.27	5.61	2.58	3.82
Teknaf	3.71	4.01	4.39	4.02	3.33	3.88	3.85	2.71	2.43	2.21	1.58	1.77	4.38	1.58	3.16
Hatiya	3.05	2.64	4.18	3.98	4.83	6.48	5.76	2.63	2.97	2.78	3.07	2.58	5.76	2.58	3.73
Patenga	6.23	6.34	7.37	7.91	8.48	8.68	9.19	8.55	7.49	6.94	6.72	5.92	9.21	5.92	7.47
Kutubdia	1.78	1.82	2.34	2.74	2.78	3.66	3.62	3.15	2.12	1.46	1.18	1.28	3.66	1.18	2.32
Mongla	1.06	1.25	1.76	2.52	2.93	2.64	2.47	2.36	1.84	1.26	1.03	1.02	2.93	1.02	1.85

D. Calculation of Maximum Available Power

The kinetic energy available in the wind cannot be converted more than 59% into the mechanical energy as stated in Betz law where practically the power extraction is usually calculated less than 45% [11]. By considering VAWT's rotor blades swept area, $A = rl$, the velocity of air = v, length of the VAWT blade = l of the density of the air = ρ and the theoretical Betz limit C_p=0.59, the output power, P (maximum) availability can be calculated by,

$$P_{available\ (max)} = \frac{1}{2}\rho A v^3 C_P \qquad (2)$$

By considering the average wind velocity from the height 5m to 15m, v= 3.437 ms^{-1} as calculated in section III.B, length of the blade, l= 10m, the radius of a swept area by blades, r = 3m, the total swept area can be calculated, $A = 30\text{m}^2$, air density, ρ = 1.16 kg/m^3 as measured in Table II, Betz coefficient, C_p = 0.59 (theoretical) and considering Betz coefficient, C_p = 0.35 (practical), the maximum output power availability for the proposed low-speed vertical axis wind turbine (VAWT) can be found using the equation (2) as

$$P_{available\ (Max)} = \frac{1}{2}\rho A v^3 C_P$$
$$= \frac{1}{2} \times 1.165 \times 30 \times 3.437^3 \times 0.59$$
$$\approx 480W\ (theoretical)$$
$$P_{available\ (Max)} = \frac{1}{2}\rho A v^3 C_P$$
$$= \frac{1}{2} \times 1.165 \times 30 \times 3.437^3 \times 0.35$$
$$\approx 270W\ (practical)$$

Table II. Air density vs. Temperature

Temperature, T (℃)	Air Density, ρ (Kg/m^3)
35	1.143
30	1.165
25	1.184
20	1.204
15	1.226
10	1.247
5	1.269
0	1.292

IV. Design of Blades for a 250W VAWT

A. Calculation of Blade Radius

In consideration of the average wind velocity as calculated in section III.B (from height 5m to 15m), v = 3.437 ms^{-1}, desired output power, $P_{desired\ (Max)}$= 250W, density of air at 30℃ temperature, ρ = 1.165 kg/m^3 (as measured in Table II), and considering the Betz coefficient, C_p ≈0.35 (practical), the blade radius, r can be determined using eq.(1) as

$$r = \sqrt{\frac{2P_{available\ (max)}}{\rho v^3 C_P \pi}}$$
$$= \sqrt{\frac{2 \times 250}{1.165 \times (3.437)^3 \times 0.35 \times 3.1416}}$$
$$\approx 3.0\text{m}$$

B. Selection of Air Foil

There are lots of NACA airfoil models available right now. For low-speed small size VAWT system, numerous studies suggested that NACA 0021 is one of the lightest weights, robust, efficient and shapeable air foil. As the blade length covers 2/3rd of the entire pole which is 15m in length and the velocity of wind is under 5 ms^{-1}, weight is the primary reason to select NACA 0021 airfoil as for the research. Another important goal is to reduce the cost as the NACA 0021 air foil is one of the low costs and commercially available airfoils which is suitable for small size low-speed blade. With a wide operational range, this airfoil performs at its best with an optimum efficiency [15, 16]. Selection of the air foil in the Qblade simulation software with 21% thickness and 100 measurement points are shown in Fig. 3. The operational points and air foil pressure point analysis of the blade (NACA 0021) are shown in Fig. 4 where the angle of incidence α = 18 (as simulated).

Although blades of VAWT generally operates in high inflow angle range, due to the limitation of the simulation software used in this study, the airfoil's inflow angle range is limited to take from -25° to +25° as shown in Figs. 5-8 where the curve shows the coefficient of lift vs. coefficient of drag behavior, lift coefficient vs. pitching moment coefficient analysis and incidence angle vs. pitching moment coefficient

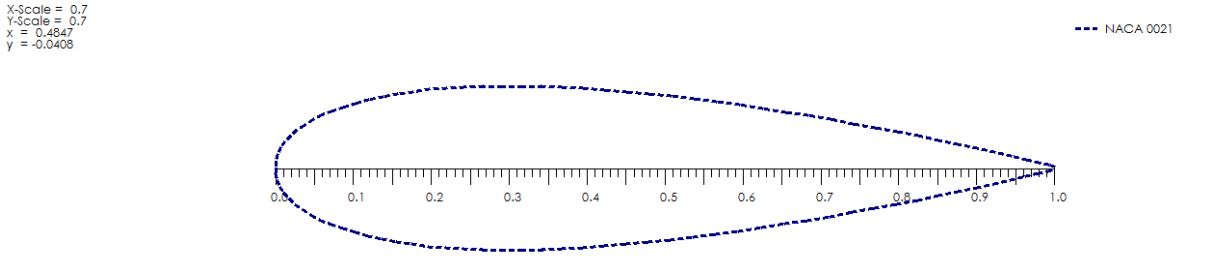


Fig. 3. NACA 0021 Airfoil Selection

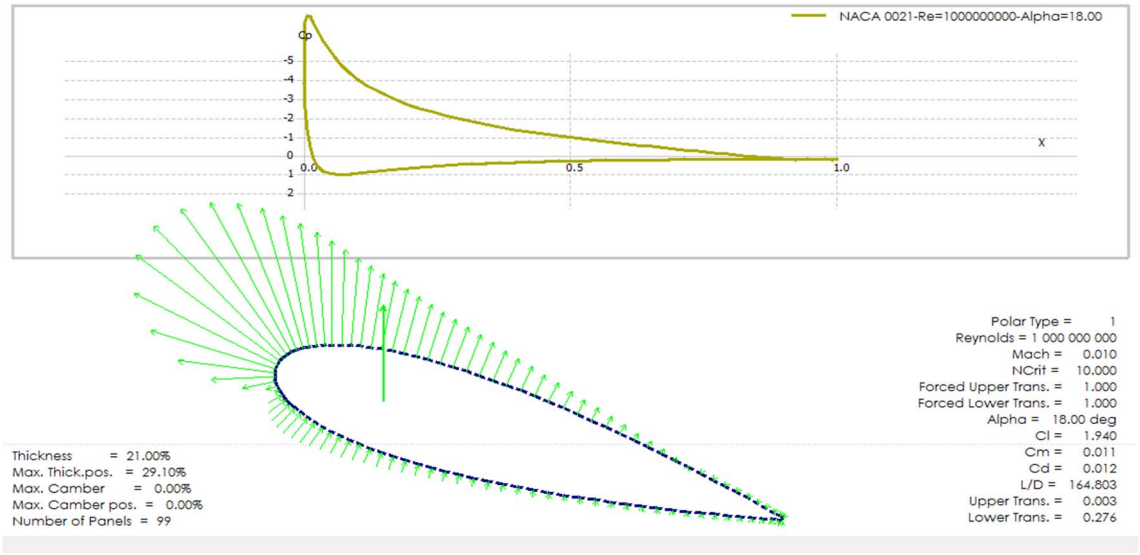


Fig. 4. NACA 0021 operation point and pressure analysis

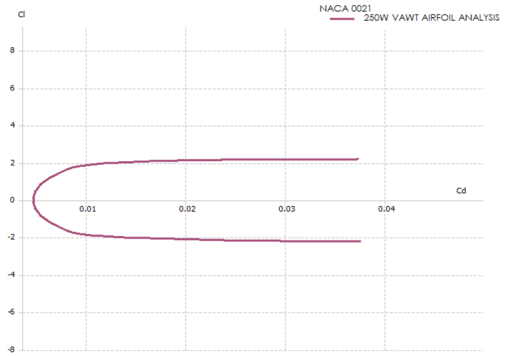


Fig 5. NACA 0021 Lift coefficient, C_l vs. drag coefficient, C_d (inflow angle --25° to +25°)

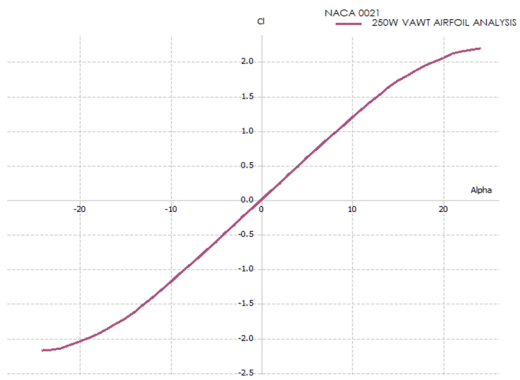


Fig. 6. NACA 0021 Lift coefficient, C_l vs. Pitching moment coefficient, α (inflow angle -25° to +25°)

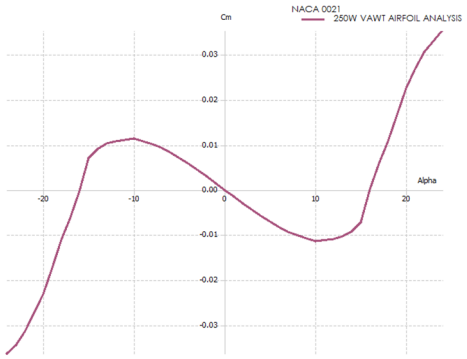


Fig. 7. NACA 0021 Incidence angle, α vs. Pitching moment coefficient, C_m (inflow angle -25° to +25°)

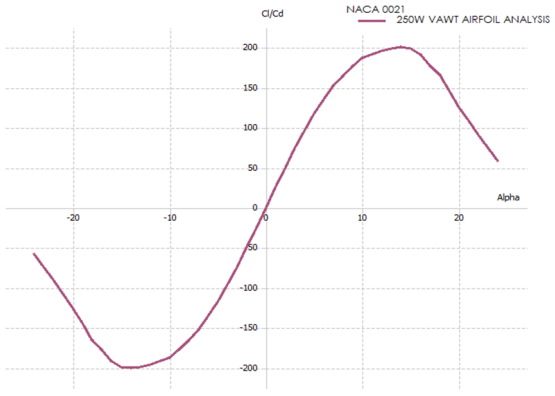


Fig. 8. NACA 0021 Glide ratio C_l/C_d vs. Incidence angle, α (inflow angle -25° to +25°)

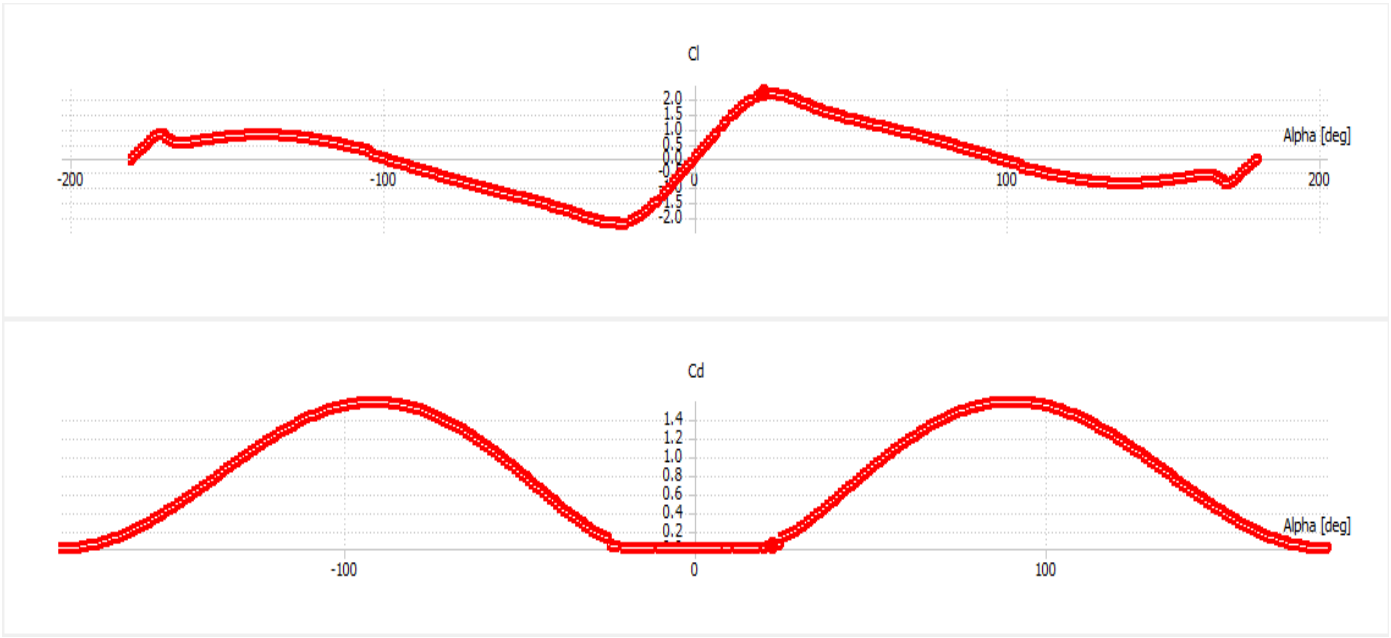


Fig. 9. 360 Polar extrapolations of NACA 0021 airfoil responses

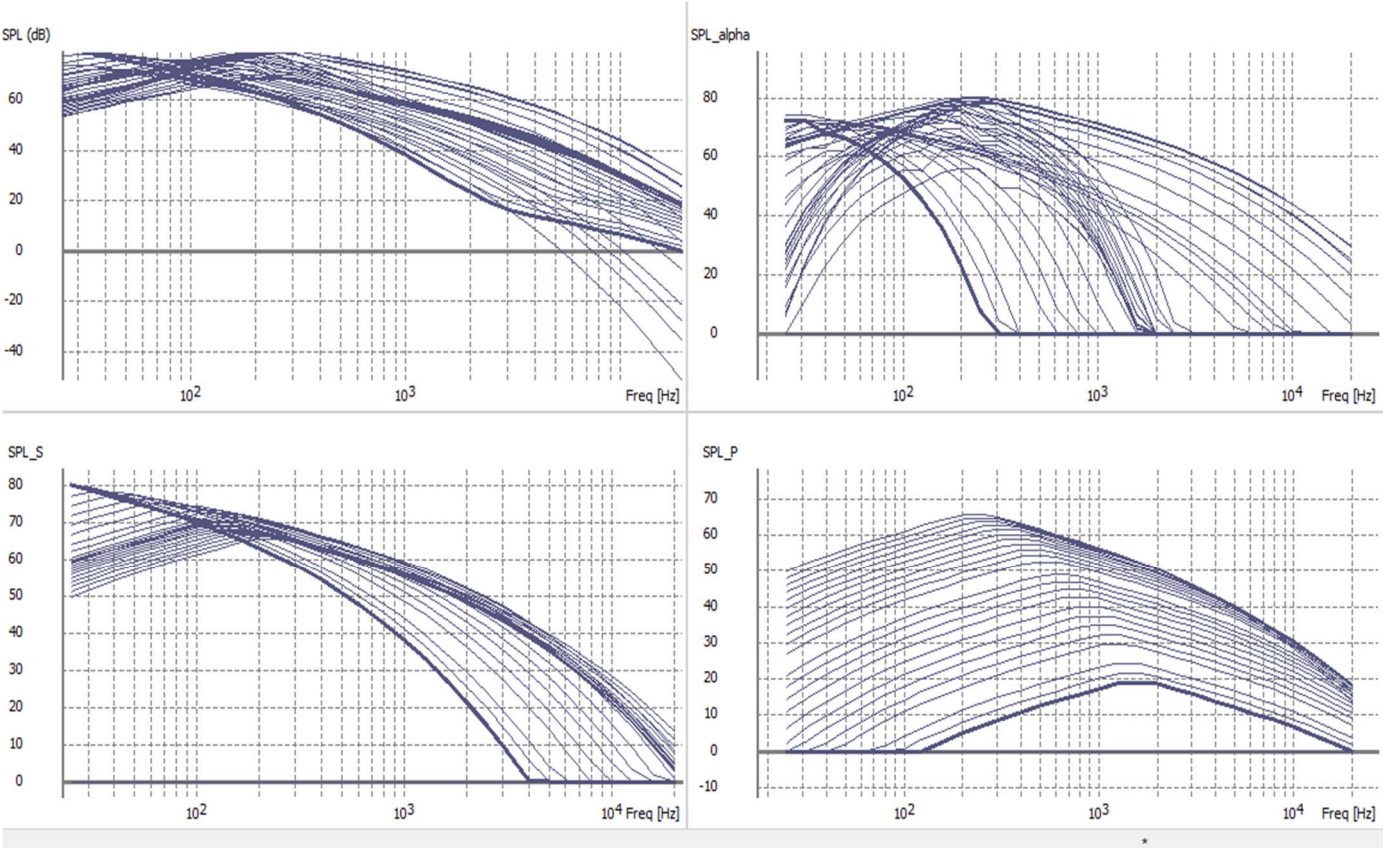


Fig. 10. Noise simulation of NACA 0021 at all operational points

with Glide ratio vs. Incidence angle, respectively. To minimize the limitation, an extrapolation of 360° of the inflow angle has been executed within the Qblade in order to converge in practice and to achieve an error-free analysis as shown in Fig.9, where angle of incidence = α , coefficient of lift = C_l, lift = L, coefficient of drag = C_d , drag = D and pitching moment coefficient = C_m.

A simulation of possible noise and interferences at different operational points are made within the Qblade as shown in Fig.10. For the purpose of this research, NACA 0021 Airfoil analysis was made to understand the blade element momentum theory. To find out the response of the selected airfoil drag, lift, the angle of incidence, glide ratio and other coefficients were analyzed. The 250W low-speed VAWT's designed blades with single blade view and rotor view using NACA 0021 air foils at different points of the blades are shown in Figs. 11 and 12, respectively, which will be used to install the wind power system in the proposed location.

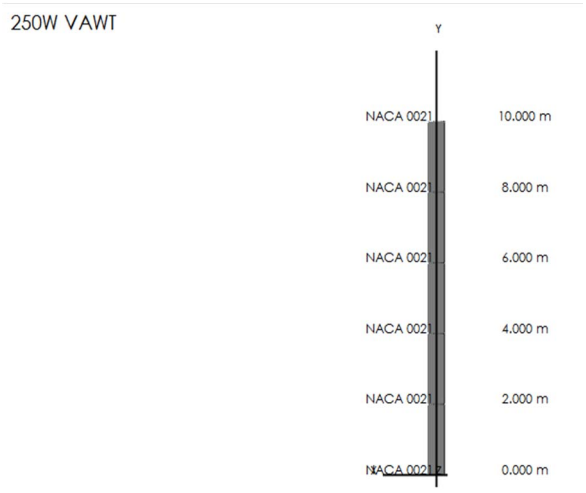


Fig. 11. Design of the Blade

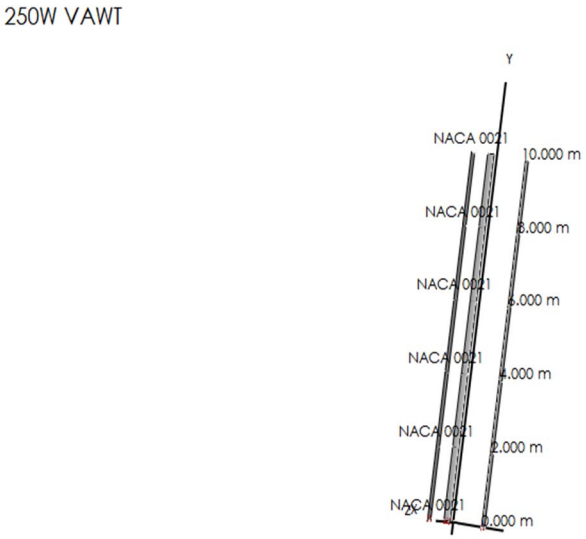


Fig. 12. 3-Blades rotor view

V. CONCLUSION

In this research work, the theoretical and statistical analysis have been studied for a small-sized low-speed vertical axis wind turbine perfectly suitable for the areas where wind velocity is quite low like coastal region of Bangladesh using NACA 0021 airfoil. The pole heights of the VAWT is keeping optimum to build at a low cost and can extract the efficient energy from available wind. In a small scale, the vertical axis turbine makes smaller in diameter and protect all the system quite well than horizontal axis turbine. That's why the VAWT is used here for low wind flow rate. Single airfoil can make this blade more robust and maintenance free. This type of small system is good for power solution for thousands of people living in these areas which can significantly reduce the pressure on fossil fuel and can have some contribution to growing electricity demand.

References

[1] A. N. M. Mominul Islam Mukut, Md. Quamrul Islam, Muhammad Mahbubul Alam, "Analysis of Wind Characteristics in Coastal Areas of Bangladesh", Journal of Mechanical Engineering, The Institution of Engineers, Bangladesh, Vol. ME39, No. 1, June 2008.

[2] M. Murshed, Y. Arafat, M. A. Razzak, "Analysis of Air Foils and Design of Blades for a Low-Speed 250W Horizontal Axis Wind Turbine Suitable for Coastal Areas of Bangladesh," 4th International Conference on the Developments in Renewable Energy Technology (ICDRET), 07-09 January, 2016, Dhaka, Bangladesh, pp.1-6.

[3] Y. Arafat, M. Murshed, M. M. Hasan, M. A. Razzak, "Design Aspects and Performance Analysis of Inner and Outer Rotor Permanent Magnet Alternator for Direct Driven Low-Speed Wind Turbine," 2nd IEEE International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), 27-28 February, 2016, Chennai, Tamil Nadu, India, pp.1-6.

[4] "Installed capacity report- BPDB", Bangladesh power development board, Available at http://www.bpdb.gov.bd/bpdb/index.php?option=com_content&view=article&id=5&Itemid=6 [May 21, 2015]

[5] Anwar Hossain, "Wind Energy in Bangladesh", Sustainable Development Networking Programme. Available at: http://www.sdnbd.org/wind.htm [June 19, 2015]

[6] M. J. Khan, M. T. Iqbal, S. Mahboob, "A wind map of Bangladesh", Renewable Energy (2004), Volume: 29, Issue: 5, Publisher: Elsevier, pp. 643-660.

[7] M. Murshed, Y. Arafat, M. A. Razzak, "Design of blades for a low-speed 400W wind turbine suitable for coastal area of Bangladesh" 3rd International Conference on Green Energy and Technology (ICGET), 11-12 September, 2015, Dhaka, Bangladesh, pp.1-6.

[8] T. V. Ramachandra, B. V. Shruti, "Wind energy potential mappings in Karnataka, India, using GIS", Energy Conversion and Management 46 (2005) pp.1561-1578.

[9] "Record Year for Wind Energy: Global Wind Power Market increased by 43% in 2005", Global Wind Energy Council Press Release, Brussels, 17th February 2006.

[10] "Final Report of Solar and Wind Energy Resource Assessment (SWERA)- Bangladesh", Renewable Energy Research Centre (RERC), University of Dhaka, Bangladesh, Available at http://xa.yimg.com/kq/groups/14503985/2014718169/name/ SWERA_Bangladesh_FullReport.pdf[July 18, 2015]

[11] Sultan Ahmmed and Quamrul Islam, "Wind power for rural areas of Bangladesh", 3rd international Conference on Electrical and Computer Engineering (ICECE), 28-30 December, 2004, Dhaka, Bangladesh.

[12] M. L. Kubik, P. J. Coker, C. Hunt, "Using meteorological wind data to estimate turbine generation output: a sensitivity analysis", World Renewable energy congress 2011, 8-13 May 2011, Linkoping, Sweden.

[13] H. Panofsky and J. Dutton, "Atmospheric Turbulence: models and methods for engineering applications", Pennsylvania State University: John Wiley and Sons, 1984.

[14] Wikipedia: web-based encyclopedia project supported by the Wikimedia Foundation, "Density of air" available at http://en.wikipedia.org/wiki/Density_of_air [June 11, 2015]

[15] John McCusker, "Design and Optimization of a Small Wind Turbine", An Engineering Project Submitted to the Graduate Faculty of Rensselaer Polytechnic Institute Available at: http://www.ewp.rpi.edu/hartford/ ~ernesto/SPR/McCosker-FinalReport.pdf [June 23, 2015]

[16] M. H. Mohamed, A. M. Ali, A. A. Hafiz, "CFD analysis for H-rotor Darrieus turbine as a low-speed wind energy converter", International Journal of Engineering Science and Technology, 18 (2015) pp.1-13.

Genetic Expression Analysis To Detect Type Of Leukemia Using Machine Learning

Umid Kumar Dey
Chittagong, Bangladesh
umiddey@gmail.com

Md. Sajjatul Islam
Assistant Professor
Chittagong Independent University
sajjatulislam@ciu.edu.bd

Abstract—**One of the worst epidemics in the history of mankind is the deadly disease known as cancer. There are several types of cancer and the one that is more commonly heard of these days is leukemia. There are two types of leukemia – acute myeloid leukemia (AML) and acute lymphocytic leukemia (ALL) – and the purpose of this study is to take into account the gene expression data of several people and predict what type of leukemia they have by using three machine learning algorithms, XGBoost, Random Forest Classification and Artificial Neural Networks. The dataset's dimensionality was reduced using principal component analysis (PCA) before using the algorithms on them.**

Keywords— **genetics, leukemia, cancer, machine learning, xgboost, neural networks, pca, random forest**

I. INTRODUCTION

Death and destruction have been in the history of humankind ever since its inception. Wars and conflicts have taken many lives while via atrocious means and it is perhaps no different now. However, war and conflicts are not the only aspects sweeping humankind as diseases have taken a prominent place in the society. One of the deadliest among them is cancer. This is a disease that has taken a lot of lives and is continuing its surge by clasping people in its claws. According to the World Cancer Research Fund [8], approximately 17 million people have been afflicted with cancer in 2018, out of which approximately 437000 cases were of leukemia, which constitutes of 2.6% of all cancer. Researchers are desperately trying to eradicate this abomination and this is where machine learning could come in handy.

The first step to solve a problem is to be aware that the problem exists. In other words, knowing the type of cancer a person has as soon as possible can help the rejuvenation process take place faster. With the advancement of technology, humankind has been able to store gargantuan amount of data. With this data, researchers can use Machine Learning algorithms to find patterns and then predict something before anything else can. The process used in this study is also similar as the main idea is to take the genetic expression data and use it to find out the type of leukemia a patient has via supervised learning. At first, the dataset that consists of 7129 independent attributes was narrowed down using principal component analysis and then XGBoost, Artificial Neural Networks and Random Forest Classification was used in order to correctly classify the type of leukemia that a specific person has.

This study can be confined and indexed down to five sections. While the next section – II – constitutes of the literature review in this field, Section III defines the system used for the purpose of this study. The experimental design is expounded in Section IV and the results are offered in Section V before finishing off by giving a conclusion in Section VI.

II. RELATED WORKS

Cancer forecast is one of the most investigated subjects around the earth because of its unmatched prominence. In [1], the researches reveal that ALL have a distinguishing gene expression outline that is stable with an early hematopoietic antecedent articulating select individual HOX genes and multilineage markers. They also added that clustering algorithms reveal that lymphoblastic leukemia with MLL translocations can evidently be alienated from conventional acute lymphoblastic and acute myelogenous leukemia. The researchers in [2], meanwhile, predicted malignancy with 98.14% on Wisconsin's Breast cancer dataset with the use of numerous if-then rules and J48 classifier. In the following study [3], Bellaachia et al. proved that C4.5 was the best at accurately predicting cancer whilst having variables like race, marital status, age, tumour size and so on. The 86.7% accuracy of C4.5 beat the other two techniques, Naïve Bayes and Neural Networks.

The researchers in [4] executed a detailed study of numerous exploration papers and associated a surfeit of data mining techniques to settle that data mining is very beneficial for forecasting the probabilities of breast cancer as it delivers respectable accuracy rates. For [5], the authors discovered that Naïve Bayes was the best technique to predict lung cancer as it could thoroughly identify the variables. The other techniques they used were if-then rules, Neural Networks and Decision Trees. In [6], Decision Tree came out victorious when it came to predicting the malignancy of a tumour. They used various mining techniques but Decision Tree gave the highest accuracy of 93.62%. For the agricultural research in [7], Dey et al. used four mining techniques and a equation of their own to forecast the harvest of rice in Bangladesh. At the end of the study, they found out that their non-linear modified regression equation produced the best results overall.

III. LEUKEMIA TYPE RECOGNITION SYSTEM (LTRS)

A. Data preprocessing

The system with which the analysis was done is encapsulated in the Figure 3.1 below.

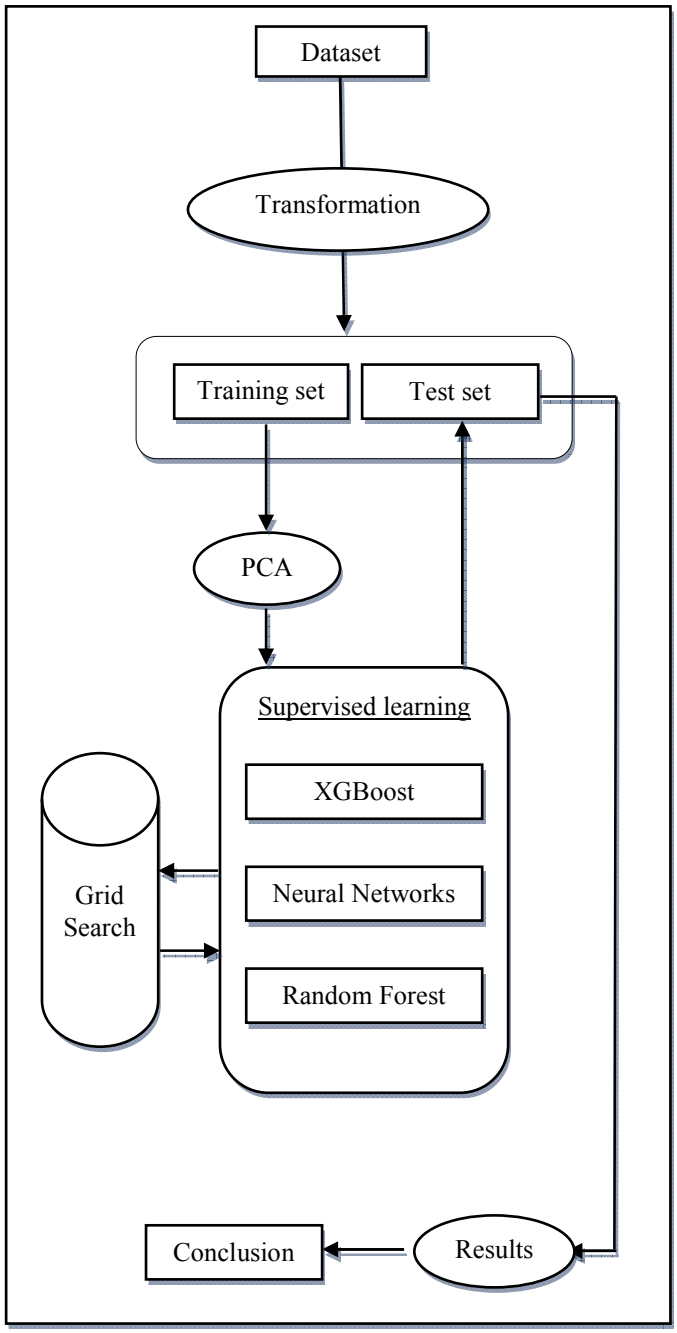


Figure 3.1: Leukemia Type Recognition System (LTRS)

To further elaborate, the dataset containing the genetic expression had to be transformed at first for the purpose of the study. After that, they were split into two: training set and test set. With the use of Principal Component Analysis (PCA), the dimension of the independent variables in both the training and test set was reduced to just 6 components from a whopping 7129. PCA is an unsupervised dimensionality reduction technique. The goal of PCA is to identify the patterns in the independent variables of the dataset and then detect the correlation between them. The idea behind PCA is to find the direction of maximum variance in a high dimensional data – such as this dataset – and then project it into a smaller dimensional subspace whilst also keeping most of the information provided by the dataset. PCA can be defined by the following steps:

1. Standardising the dataset
2. Obtaining the eigenvalues and eigenvectors from the correlation matrix or just perform Singular Vector Decomposition
3. Sorting the eigenvalues in descending order while selecting the k eigenvectors that relate to the biggest eigenvalues, where k is the number of dimensions of the new subspace
4. From the chosen k eigenvectors, construct a projection matrix, W
5. The final step concludes the transformation of the original dataset, X, using the projection matrix W as a means to obtain the k-dimensional subspace Y.

After reducing the dimension of the large dataset, the training set was passed on for supervised learning with the use of three Machine Learning algorithms: XGBoost, Artificial Neural Networks and Random Forest Classification. They are classification algorithms and were chosen because of the fact that the output is binary (which suits classification more) and that the computational requirement to run these aren't very high. Most importantly, for a dataset like this, classification algorithms give better accuracy.

B. Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting, or simply known as XGBoost, is among the most powerful gradient boosting algorithms created for the purpose of not only speeding up the execution but also improve the performance of the model. It has become renowned in Kaggle for its competition-winning ability. The principle idea behind XGBoost is to combine many 'weak learners' trees to form a strong learner that performs better and much faster than the others. The number of boosted trees to fit in this model was set at 40 while the booster used was gblinear.

C. Artificial Neural Networks

Scientists have been trying to make systems that can work like the human brain. Artificial Neural Network (ANN) is basically an attempt simulate the working methodology of the brain. The brain consists of many neurons that send and receive signals, allowing humans to make a decision accordingly. The idea behind ANN is similar as there are layers of neurons connected to each other to make a network. Some networks are feed-forward, which means that it only moves in the forward direction but the one used is back-propagated, wherein the error between the test result and the

actual result is sent back to the network as means to adjust the weights and reduce the error rate in the following cycle. This is the oversimplified idea of ANN.

Delving deeper, Neural Networks can consist of many layers between the input layer and the output layer. When the input layers are fed with the data, it passes it to "hidden" layer in the middle, whose neurons then add a weight to the information received from the input layer. Every neuron adds up the input it receives and only "fires up" if the sum of it is more than the threshold value. The threshold here is, simply, decided by the activation function and for this study, the activation function used for the hidden layers is Rectified Linear Unit. It is one the most popular functions in Neural Networks and can be represented as follows:

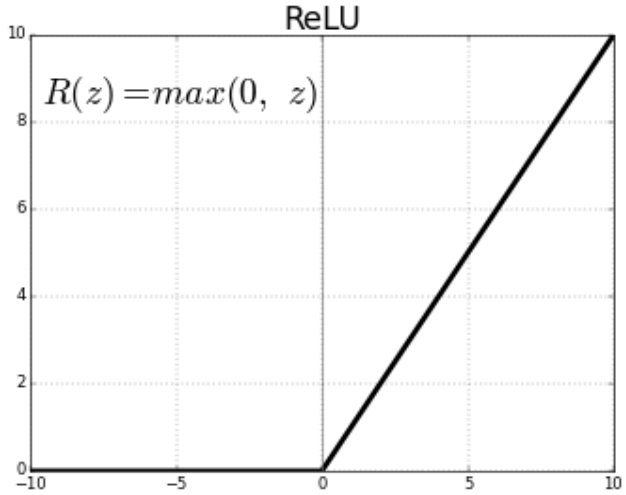


Figure 3.2: Rectified Linear Unit activation function [9]

For the output layer, another activation function was used. Since the output has only two types – acute myeloid leukemia (AML) and acute lymphocytic leukemia (ALL) – the sigmoid activation function was used as the output for this function is constrained to either 0 or 1.

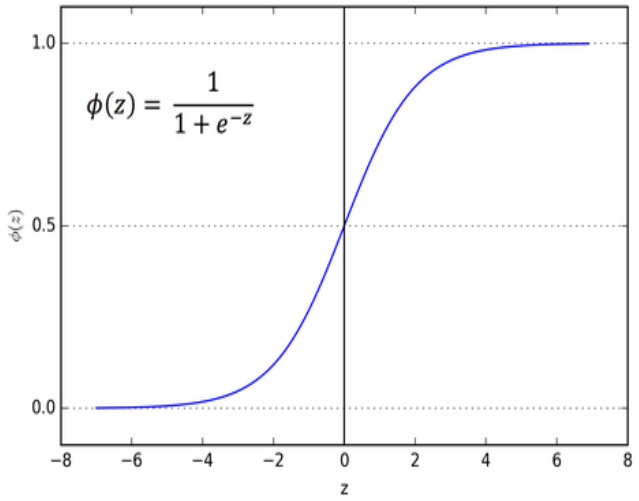


Figure 3.3: Sigmoid activation function [10]

After that, the output is cross-checked with the real result and if there are errors, the feedback is given to the network. This is basically back-propagation since after being made aware of the feedback, the hidden layers adjust the weights as a means to reduce the error rate in the next epoch. An epoch is simply the cycle from the moment the data enters to the input layer(s) to the moment an output comes about. For this study, an epoch of 100 was chosen. Since PCA reduced the dimensionality to 6 components, the input layer for this ANN had 6 neurons. Furthermore, three hidden layers with 6 neurons each was chosen as it gave the best results.

D. Random Forest Classification

Random Forest Regression basically chooses some data points in the dataset and builds a decision tree associated to those points. After that, multiple decision trees are built to cover the entire dataset. Once it is done, a new data point's category is chosen by a voting system, wherein the trees will predict where the new point belongs and the category that gets the most votes will have the new point. This procedure can be narrowed down to four steps:

1. Select random data points, K, from the training set
2. Build a decision tree based on those selected K data points
3. Choose the number of trees, N, to be built and repeat steps 1 and 2
4. To predict which category a new data point belongs to, assign the new data point to the category that gets the most votes from the decision trees

For this study, the number of trees used for Random Forest Classification was set at 10 with minimum sample leaves per node kept at two and minimum sample split at 0.1.

E. Grid Search

After training the algorithms with the data, the best hyper-parameters were selected using Grid Search. Hyper-parameters are parameters whose values can be chosen by the user. For instance, the number of trees in Random Forest Classification is a hyper-parameter. Parameter tuning optimizes the model to such an extent that the accuracy of the prediction increased. To find out the best hyper-parameter values, Grid Search was done on the dataset corresponding to the classifier (or learner) to find the best parameter values for that specific classifier. For instance, to find out the optimal hyper-parameter values for Random Forest Classification, once the Random Forest classifier was trained with the training set, the classifier along with the training set was passed onto Grid Search, which then scans the data to find the optimal hyper-parameter values corresponding to the Random Forest classifier. This is how the hyper-parameter values for Random Forest Classification, ANN and XGBoost was found out. After that, those values were implemented on the classifiers and they were once again trained on the training set.

F. Evaluation

The results were then obtained by running the trained learners on the test set and evaluated using Accuracy, Precision, Recall and F1 score. For the context of this study, where the objective is to predict the type of cancer, Accuracy and modified F1 score are given the highest priority. They are calculated as follows:

1. Accuracy = (True AML + True ALL)/Total data points
2. AML Precision = True AML/(True AML+False AML)
3. AML Recall = True AML/(True AML+False ALL)
4. AML F1 score = 2 * (AML Precision * AML Recall)/(AML Precision + AML Recall)
5. ALL Recall = True ALL/(True ALL+ False AML)
6. ALL Precision = True ALL/(True ALL + False ALL)
7. ALL F1 score = 2 * (ALL Precision * ALL Recall)/(ALL Precision + ALL Recall)

Here, Accuracy is mainly the ratio of the correct predictions over the total data points in the dataset. AML Precision, meanwhile, is the measure of how accurate a model is when it makes AML predictions and ALL Precision is the measure of how accurate the model is when it makes ALL predictions. AML Recall is how good a model is at correctly predicting the AML instances and ALL Recall is how good the model is at correctly predicting ALL instances. Finally, AML F1 score is the accuracy based on AML Precision and AML Recall while ALL F1 score is the accuracy based on the ALL Precision and ALL Recall

IV. Experimental Design

A. The dataset and modifications made

The dataset was taken from Kaggle [11] and it has the genetic expression of 72 people, where each person consisted of 7129 genes. As mentioned in section III, the dataset had to be modified. The reason being, the dataset had gene expression for 72 people but with the independent variables in the rows instead of the columns. There was also a categorical variable named Call after each independent variable and it was removed because it was deemed as unnecessary. The dataset was transposed with the independent variables taking the columns. The dataset consists of the measurements relating to ALL and AML samples from the bone marrow and peripheral blood.

Table 4.1: Sample of dataset before preprocessing

Gene Description	Gene Accession Number	1	call	…

As mentioned earlier, the dataset, in its raw format, had the independent variables as rows. The 'call' variable was removed entirely since it was redundant. The header Gene Description and Gene Accession Number were dropped but their values were kept. These values are the types of genes and there are 7129 of them.

Table 4.2: Sample of dataset after preprocessing

1	-214	-153	…(other 7127 gene values)
2	-139	-73	…(other 7127 gene values)
3	-76	-49	…(other 7127 gene values)

Here, the values correspond to the patient as in 1 is patient number 1 and the values in its row are the genetic values relating to it, of which there are 7129 values.

There are two target variables – ALL and AML – where AML stands for *acute myeloid leukemia* and ALL means *acute lymphocytic leukemia.* Since having both of these variables would lead to a trap, a single dummy variable was taken where 0 meant that the said patient had ALL while patients with AML were given the value of 1. The dataset was split into training and test set with a split ratio of 0.65 for training set and 0.35 for test set, which means 26 data points were assigned for the test set and 46 for the training set.

B. Tools used and calculations made

Python, along with libraries such as scikit, numpy, pandas, tensorflow, keras and so on, was the only tool used to conduct this research. The calculations made were corresponding to the test set and the results were evaluated against the actual values. If the prediction was 0, then the person is thought to be afflicted by ALL and if 1, then AML was the type of leukemia in that person. These results were represented on a confusion matrix in in order to calculate the False AML, False ALL, True AML and True ALL, which, in turn, was used to calculate the Accuracy, AML Precision, AML Recall, ALL Precision, ALL Recall and F1 score.

V. Results and Discussion

After the trained learners were run on the test set, the resulting output, as expected, assisted in evaluating the methods used in this study. Keeping in mind the fact that a result of 1 means that the specific person is predicted to have AML type leukemia and a result of 0 means that a person is predicted to have ALL type of leukemia, the confusion matrix for the three algorithms are as follows:\

Table 5.1: Confusion matrix for XGBoost

	Predicted ALL	Predicted AML
Actual ALL	14	0
Actual AML	2	10

Table 5.2: Confusion matrix for ANN

	Predicted ALL	Predicted AML
Actual ALL	14	0
Actual AML	2	10

Table 5.3: Confusion matrix for Random Forest

	Predicted ALL	Predicted AML
Actual ALL	14	0
Actual AML	5	7

From the confusion matrices in Table 5.1, 5.2 and 5.3, the Accuracy, Precision, Recall and F1`Score were calculated and the resulting graphs obtained were as follows.

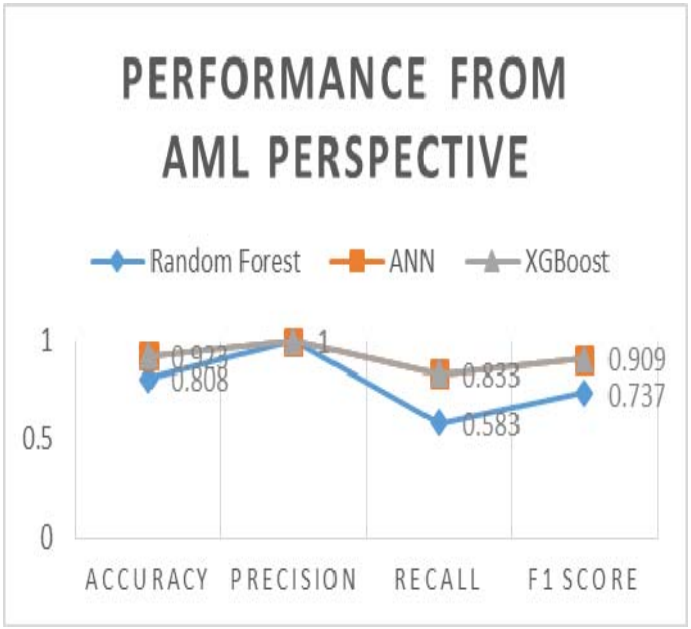


Figure 5.1: Performance from AML perspective

From Figure 5.1, it can be seen that ANN and XGBoost have the exact same results and are far superior to Random Forest when it comes to predicting the AML type leukemia. The Precision of 1 for all three learners simply means that when the model predicts that the type of leukemia is AML, it is correct every time. However, the Recall value for ANN and XGBoost mean that the model can predict whether a person has AML with 83% correctness.

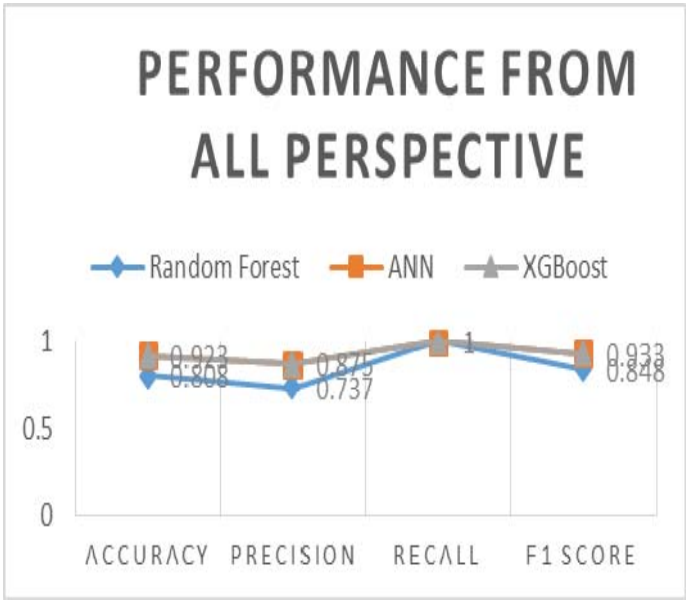


Figure 5.2: Performance from ALL perspective

In Figure 5.2, as expected, ANN and XGBoost fare better than Random Forest but it should be noted that the latter is actually performing better while predicting ALL than it was in Figure 5.1. Random Forest in AML prediction had an F1 Score of 0.737 while in Figure 5.2 is at 0.848. Indeed, even XGBoost and ANN have a higher F1 score while predicting ALL when compared to predicting AML. The Recall value of 1 for all three learners mean that the model can predict whether a person has ALL with 100% correctness. The Precision value, meanwhile, suggests that if the model predicts that a person is afflicted by ALL, then it is right 87.5% of the time for ANN and XGBoost and 73.7% with Random Forest. The accuracy, for obvious reasons, stays the same from both perspectives.

VI. Conclusion

From section V, Results and Discussions, it could be safely concluded that the model used here does a good job at predicting the type of leukemia a patient has, especially when it comes to forecasting ALL type leukemia. With just two incorrect outcomes, out of a possible 26 (0.923 * 26 = 24), in XGBoost and ANN it is very difficult to better the model presented here. In fact, comparing this with other studies, A.C Tan et al [12] got a predictive accuracy of 91.18% while T. Li et al [13] did slightly better at 93%. Furthermore, A.B. Olshen et al secured similar results as their study [14] also had only two incorrect outcomes out of a possible 34.

The use of PCA to decrease the dimensionality of the dataset not only made the computation faster but also gave better results since when the dimensionality wasn't reduced at all, the results were far poorer, which is why they weren't included here. Similarly, the use of Grid Search allowed the discovery of the best possible hyper-parameters associated to the learners used, which, in turn, helped the model produce better results. Without tuning the parameters, the results were some way off to what was found when the parameters were

tuned. Finally, the power of Neural Networks and XGBoost became apparent here as even a great ensemble algorithm like Random Forest fell pale in comparison to the former two. However, considering everything from computational time and predictive power, XGBoost edges out even Neural Networks as it was much faster to simulate than Artificial Neural Networks.

REFERENCES"

[1] S. A. Armstrong et al, "MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia", Nature Genetics, Volume 30, January 2002

[2] S. S. Shrivastava et al., "An Overview on Data Mining Approach on Breast Cancer data," International Journal of Advanced Computer Research, Volume-3 Number-4 Issue-13 December-2013.

[3] A. Bellaachia and E. Guven, "Predicting Breast Cancer Survivability Using Data Mining Techniques", [Online]. Available: http://www.siam.org/meetings/sdm06/workproceed/Scientific%20Datasets/bellaachia.pdf

[4] S. Gupta et al., "Data Mining Classification Techniques Applied for Breast Cancer Diagnosis and Prognosis", Indian Journal of Computer Science and Engineering.

[5] V. Krishnaiah et al., "Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques", International Journal of Computer Science and Information Technologies, Vol. 4 (1), 2013, 39-45.

[6] S. Kharya, "Using Data Mining Techniques for Diagnosis and Prognosis of Cancer Disease", International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.2, No.2, 2012.

[7] U. K. Dey et al., "Rice Yield Prediction Model Using Data Mining," International Conference on Electrical, Computer and Communication Engineering (ECCE), February 16-18, 2017, Cox's Bazar, Bangladesh

[8] World Cancer Research Fund, "Worldwide cancer data", last accessed on November 4, 2018, https://www.wcrf.org/dietandcancer/cancer-trends/worldwide-cancer-data

[9] Medium, "ReLU : Not a Differentiable Function: Why used in Gradient Based Optimization? and Other Generalizations of ReLU", last accessed on November 8, 2018, https://medium.com/@kanchansarkar/relu-not-a-differentiable-function-why-used-in-gradient-based-optimization-7fef3a4cecec

[10] Towards Data Science, "Activation Functions: Neural Networks", last accessed on November 8, 2018, https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

[11] Kaggle, "Gene expression dataset (Golub et al.)", last accessed on April 12, 2019 https://www.kaggle.com/crawford/gene-expression

[12] A.C. Tan et al., "Ensemble machine learning on gene expression data for cancer classification," Applied Bioinformatics, Open Mind Journals, 2003

[13] T. Li et al., "A comparative study of feature selection and multiclass classification methods for tissue classification based on gene expression," Bioinformatics, Volume 20, Issue 15, 2004, 2429–2437

[14] A.B. Olshen et al., "Deriving quantitative conclusions from microarray expression data," Bioinformatics, Volume 18, Issue 7, 2002, 961-970

Sensor-Less Solar Irradiance Estimation and Maximum Power Point Tracking Using Artificial Neural Network

Arif Rubayet
Department of Electrical and Electronic Engineering
Ahsanullah University Of Science And Technology
Email: a.arifrubayet@gmail.com

M.M. Atiqur Rahman
Assistant Professor
Department of Electrical and Electronic Engineering
Ahsanullah University of Science And Technology
Email: atiq_aust@yahoo.com

***Abstract*—Maximum power point tracking using artificial neural network model has been implemented. Maximum Power point tracking is used to track the maximum power from a PV panel under different conditions. This model is tested for a variety of irradiance from the light source. The proposed ANN-based algorithm does not have any sensor for prediction purpose. In order to cope with real life maximum disturbance, it is tested under non-linear time domain and non-linear condition state as well. That shows noteworthy performance in estimating maximum power point (MPP). It has the competency to predict irradiance and maximum power point(MPP) with a very tiny amount of error using two series neural networks. Simulation results and experimental studies manifest that proposed sensorless ANN-based model has a minor amount of undesired oscillations and much more effective in tacking maximum power under swiftly changing condition.**

Index Terms*—Maximum power point trackers, photovoltaic systems, ANN algorithm, irradiance, sensor, I–V curve, DC/DC converter.

I. Introduction

Global warming is being commented over the past years. Day by day we are forwarding in a path that needs electricity by anyhow. Many methods have developed for energy generation and carry them out. Most out of them use fossil fuel for generation purposes, which has the ultimate result of decreasing the amount of fossil fuel from limited sources and increasing pollution and adverse effect. However, we need energy sources that do not affect the environment. Renewable energy sources can be a good option to meet the demand for todays electricity famished world. The renewable energy sources (RES) have a little impact comparing to other energy sources like fossil fuel. Solar power is one of the widely used energy sources with a negligible amount of adverse effect on the environment [1]. All other criteria for a clean planet need to be considered. Solar power is available for most of the time when energy demand is high. The solar power generation technique is described as a photovoltaic (PV) system, which has a remarkable performance among other renewable energy sources. Over past years, many MPPT techniques [2], [3] have been proposed and implemented by researchers; such as fuzzy logic method [4], incremental conductance method, and artificial neural network method. Among them, fuzzy and artificial neural networks are getting much more response because of their robustness and accuracy [5]. Increase in output energy is 45.2% for using MPPT in the sunny environment [6]. Our main focus would be given to the implementation of Artificial neural network and maximum power point tracking under different unfavorable conditions. Selecting the efficiency depends on the method and accuracy of tracking MPP. The PV system shows a non-linear Current and Voltage characteristic under different conditions like irradiance and temperature, which depends on selecting the maximum power point. Therefore, tracking and determination of MPP have a tremendous effect on the efficiency [7]. MPPT ensures the maximum possible power to deliver from the solar cell, and even it may happen in low light. Modern MPPT system has approximately 93% to 97% efficiency in the conversion. Comparing with traditional MPPT technique, it is concluded that ANN based technique has a swift response to MPP tracking and more accuracy under different environmental conditions [8]. Perturbation and observation technique is a widely used method for MPPT, but the MPP tracking performance is not much responsive, and it results in undesired oscillation, which has consequences of disturbance in tracking operating point [9]. Artificial Neural networks can be used to overcome such MPPT problem [10]. ANN can handle non-linear equation and few data loss effectively [11], [12]. Very reasonable and low-cost ANN based MPPT techniques take voltage, and current as input [13]which has multistage output because of two series of the neuron network. Our proposed technique composed of two cascaded ANN, first one for irradiance prediction and the second one is for MPP, thus required training sets have been minimized. The complexity of the PV grid integration is increasing day by day and also been reported by industries frequently [14]. As PV arrays, Voltage and current are used for the input; no other sensor is used, which decreases the possibility of being complex. This presented technique has the specialty of calculating irradiance and MPP with a nominal amount of error using Neural network (NN) based algorithm. For any changes in irradiation, it can estimate the new MPP, which ensure better

efficiency and better performance. The simulation results and experimental studies indicate the positive outcome of this algorithm. This paper has five major parts. Introduction part is discussed in the first section. PV model and different behavior are presented in the second part. In the third section maximum power point tracking strategy is discussed, proposed model simulation results are shown in the fourth part. Finally, in the fifth section, the conclusion part is discussed.

II. Solar PV Model

The fundamental part of a PV model is the photovoltaic cell. Current is proportional to irradiation. Thus, depending on different condition solar irradiation varies with time, which has consequences of output power variation. The mathematical expression for I(V) curve of a PV model is given by the following expression.

$$I_{pv} = i_{pv} - I_0 \left[e^{\frac{(V_{pv}+I_{pv}R_s)}{\eta N_s V_t}} - 1 \right] - \frac{V_{pv} + I_{pv}R_s}{R_{sh}} \quad (1)$$

$$V_t = \frac{K_b T}{e} \quad (2)$$

Where I_{pv} denotes output current of the Solar cell, V_{pv} is Output voltage, V_t is the thermal Voltage, I_0 is reverse saturation current and K_b is Boltzmann's constant.

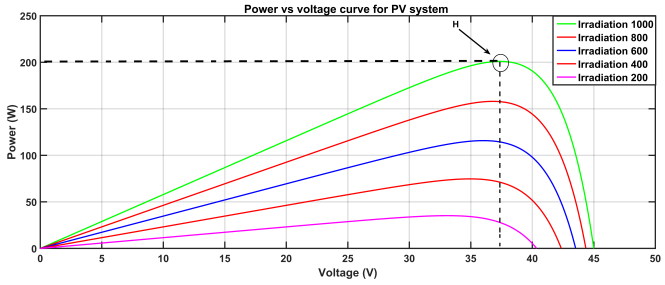


Fig. 1. Power vs voltage curve for PV system

Effect of different irradiation for the PV system can be easily observed in Fig. 2. For this experiment we have taken (1000-800-600-400-200) W/m2 as our primary irradiation levels, that is shown in both Fig. 1 and Fig. 2. Under different irradiation level, the variation of current and voltage is shown in Fig. 2.

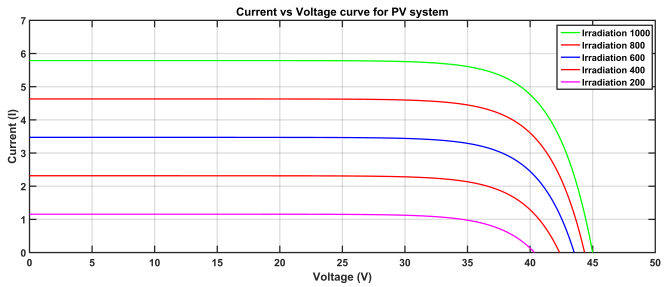


Fig. 2. Current vs Voltage curve for PV system

III. ANN Based MPPT Algorithm

As discussed before, MPP depends on climatic conditions. Because it is directly dependent on irradiation, and different climatic conditions give rise to different irradiation. This basic concept is demonstrated clearly in Fig. 3. It shows how MPP is obtained from our proposed model. After data extraction from MATLAB, those are used for neural network training. Thus different layer will be created. ANN1 predicts the irradiation and ANN2 will provide MPP. Any significant change in irradiation results in MPP variation. For distribution, DC-DC buck-boost converter is used.

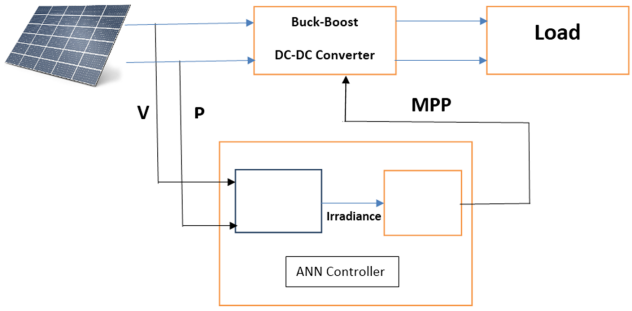


Fig. 3. ANN based MPPT for PV system

In Fig.1, (H) represents the maximum power for its corresponding voltage level V* and current level I*. The artificial neural network is used for the internal measurement of current, voltage and MPP. Different types of algorithms are used to train the ANN. Among them, feed forward back propagation is very popular. Location selection, stock market prediction, distinct pattern recognition can be carried out efficiently with this propagation method. Feed-forward backpropagation algorithm is chosen because of its high-performance rate. This type of network has a massive number of virtual way to carry out information as per networks need, and this information can move forward or backward direction. This propagation network uses a gradient-descent search technique to minimize the error for the network. In the following Fig. 4, a basic feed forward back propagation steps are shown.

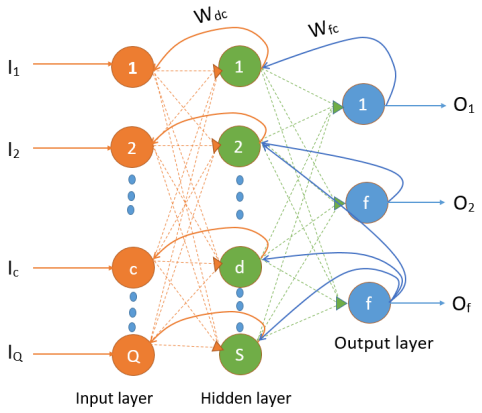


Fig. 4. Feed forward back propagation layers

$$u_c = \sum_{d=-1}^{n} I_d W_{cd} \tag{3}$$

$$V_f = \frac{1}{1 + e^{(-u_c)}} \tag{4}$$

Expression (5) stands for the actual output.

$$O_c = \frac{1}{1 + e^{(-s_c)}} \tag{5}$$

Difference between our targeted output square and actual output square is denoted as the sum of squared error, which would be expressed as (6).

$$E = \frac{1}{2} \sum_{m-1}^{M} E_m = \frac{1}{2} \sum_{m=1}^{M} \sum_{c=1}^{1} (t_c^m - O_c^m)^2 \tag{6}$$

Weights are updated from the minimization of error function using gradient descent method. Continuously those weights are updated. For c^{th} neuron, d^{th} weight update equation would be expressed as (7).

$$W_{cd}(t+1) = W_{cd}(t) + \eta \left(\frac{\partial E_m}{\partial W_{cd}(t)} \right) \tag{7}$$

Where η denotes the learning rate, $W_{cd}(t)$ denotes the old weight, $W_{cd}(t+1)$ for the new weight.

IV. Results and Discussion

For this proposed ANN-based model, two series networks are used, and those are tested under different conditions. Using MATLAB software this model was trained. Validation performance is shown in Fig. 5.

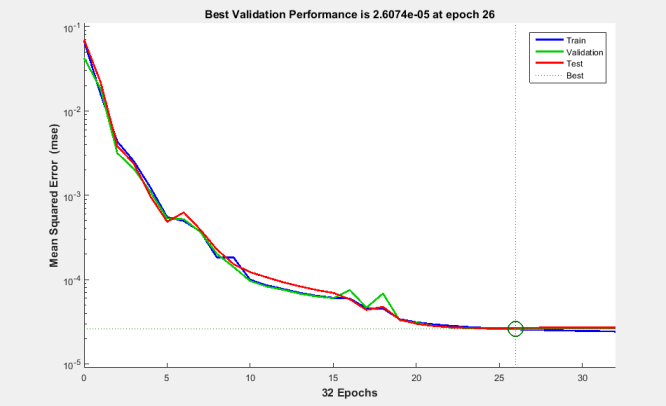


Fig. 5. Validation performance for the proposed system

Where the solid blue line represents training set, green line for validation, a red line for testing the trained model. Best validation performance is indicated, which is at epoch 26. Epoch is iteration that indicates one forward path and one backward path of all our training data (represented by a blue line). To adjust the weight of the neural network, a training set is used. The validation set is used to minimize over-fitting. It's not adjusting the weight of the network and data sets. Testing set is used for testing the final solution of the neural network which ensures the actual predictive power of the network. Here two layers and ten neurons are taken. Tan-sigmoid function is used. During training state, a regression curve is found, which is shown in Fig. 6.

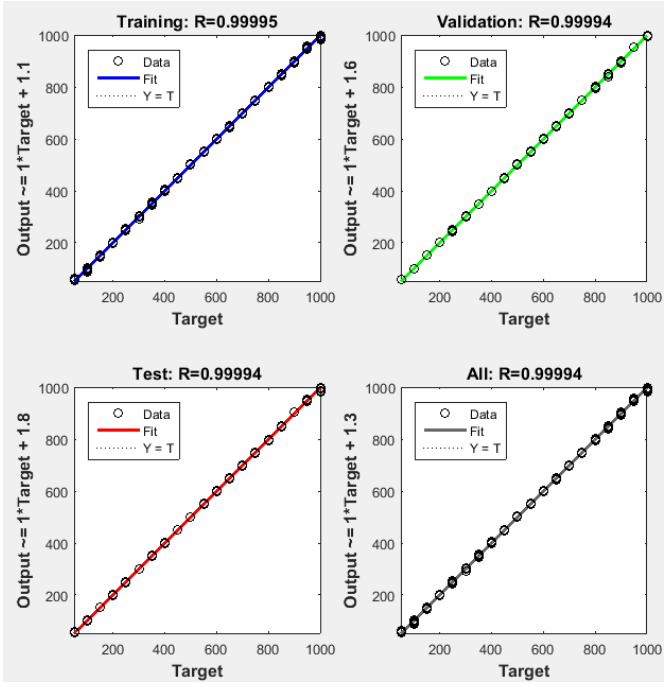


Fig. 6. Regression analysis graph for the proposed model

In Fig. 6, it is simply demonstrated about training performance. This regression performance is generated from MATLAB internal logic, and it gives an idea of whether the training was good enough or not. From Fig. 6, it can be mentioned that our proposed model has performed better in this training stage as values are close to the unit. For training R= 0.99995, for validation R= 0.99994 and value for overall performance is R=0.99994. Thus solid lines are following dotted data very accurately. To avoid any kinds of misrepresentation, this analysis was tested for several times and confirmed.

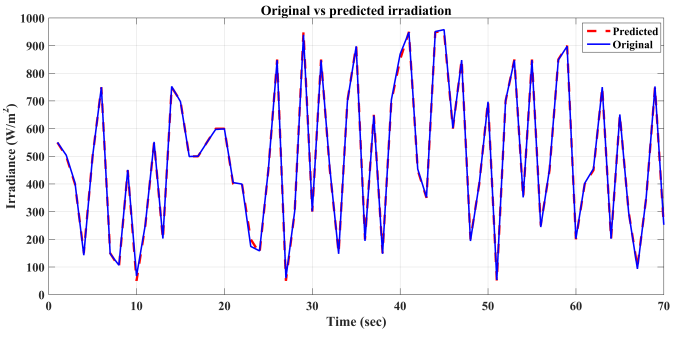


Fig. 7. Original vs predicted irradiation

Among two series networks, the first network estimates the

irradiance which is shown in Fig 7. Blue solid line represents the given irradiation, and red dotted line stands for predicted irradiation. Predicted irradiance is following the original one very closely which results in a tiny amount of error. This error is the difference between original irradiation and predicted irradiation which is shown in Fig. 8.

Table 1: Comparision between MATLAB estimation and given data

Serial No	Given data	Estimated data	Error
01	550	548.5773	1.4227
02	750	750.7156	-0.7156
03	150	150.5157	-0.5157
04	250	253.8099	-3.8099
05	100	102.9519	-2.9519

For ANN1 network, up to 1000 W/m2 irradiation was taken into consideration. Which can be seen from Fig. 7. From table 1, we can get an idea of error volume. All of those data are taken from the above mentioned Fig. 7.

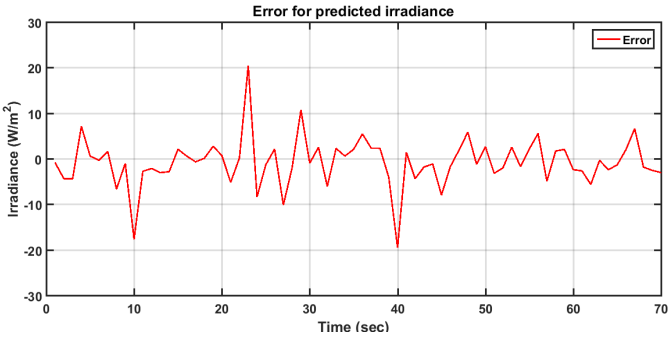


Fig. 8. Error for predicted irradiance

In Fig. 8, simulation result shows the maximum error is about 20. Which is 2% of the maximum irradiation taken into account. Thus the proposed project shows how accurately it can estimate a suitable irradiance for the next (ANN) network. For significant changes in any condition, it has the ability to get better estimation with low error portion. Next ANN2 network works for maximum power point estimation. Which will be in series with the previous network ANN1. In Fig. 3, it is clearly shown that the output of ANN1 is used as an input of ANN2. Therefore, this proposed model can estimate MPP more accurately.

Fig. 9 demonstrates how the estimated MPP is following the given (original) MPP with a low error portion. Hence, it suggests the proficiency of the model. It can be mentioned from the previously stated graphs that error rate depends on the range of original value, where the range is a difference between the maximum and the minimum of the value. Fig. 9 exhibits that estimated MPP almost overlaps the original MPP that leads to a small variance from the ideal value. Difference between those two arrays is simulated which represents an error graph, showed for MPP estimation. In Fig. 10 that is clearly illustrated.

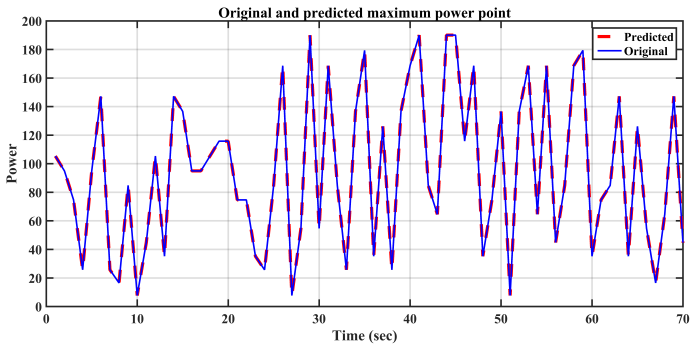


Fig. 9. Original vs predicted maximum power point

Table 2: Comparision between MATLAB estimation and given data

Serial No	Given data	Estimated data	Error
01	54.7055	54.6699	0.0356
02	35.2267	35.2245	0.0022
03	84.8565	84.8326	0.0239
04	95.004	95.0429	-0.0389
05	147.2778	147.2814	-0.0036

From table 2, we can clearly see how precise those estimations are. Here all of those data are chosen from before mentioned Fig. 9.

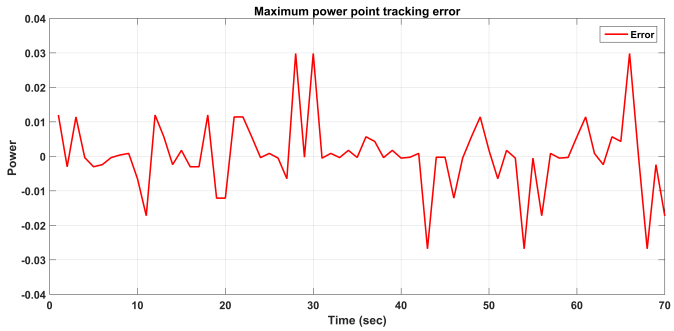


Fig. 10. Maximum power point tracking error

It can be mentioned from Fig. 10 that maximum error for MPP estimation is about 0.03, which is almost negligible. Hence, our proposed ANN-based maximum power point tracking system can function effectively.

V. Conclusion

In this paper, Sensor-Less solar irradiance estimation and maximum power point tracking algorithm is proposed to enhance the efficiency of Maximum power point extraction. Our proposed ANN-based algorithm does not need any sensor to track MPP. The simulation and experimental results of the presented algorithm are compared with traditional results, and that demonstrates the proposed algorithm has better performance

than conventional MPPT method. It exhibits swift response and accurate estimation of irradiance and MPP while tested under unfavorable and non-linear condition. The proposed model has little amount of error which implies the prosperity of this research.

ACKNOWLEDGMENT

This research work is supported by Ahsanullah University of Science and Technology(AUST). The authors acknowledge the facilities provided by AUST.

REFERENCES

[1] Y. Sukamongkol, S. Chungpaibulpatana, and W. Ongsakul, "A simulation model for predicting the performance of a solar photovoltaic system with alternating current loads," *Renewable energy*, vol. 27, no. 2, pp. 237–258, 2002.

[2] T. Esram and P. L. Chapman, "Comparison of photovoltaic array maximum power point tracking techniques," *IEEE Transactions on energy conversion*, vol. 22, no. 2, pp. 439–449, 2007.

[3] S. Jain and V. Agarwal, "Comparison of the performance of maximum power point tracking schemes applied to single-stage grid-connected photovoltaic systems," *IET Electric Power Applications*, vol. 1, no. 5, pp. 753–762, 2007.

[4] B. N. Alajmi, K. H. Ahmed, S. J. Finney, and B. W. Williams, "Fuzzy-logic-control approach of a modified hill-climbing method for maximum power point in microgrid standalone photovoltaic system," *IEEE Transactions on Power Electronics*, vol. 26, no. 4, pp. 1022–1030, 2011.

[5] L. Elobaid, A. K. Abdelsalam, and E. E. Zakzouk, "Artificial neural network based maximum power point tracking technique for pv systems," in *IECON 2012-38th Annual Conference on IEEE Industrial Electronics Society*. IEEE, 2012, pp. 937–942 .

[6] A. Bahgat, N. Helwa, G. Ahmad, and E. El Shenawy, "Maximum power point traking controller for pv systems using neural networks," *Renewable Energy*, vol. 30, no. 8, pp. 1257–1268, 2005.

[7] Y. Liu, M. Li, X. Ji, X. Luo, M. Wang, and Y. Zhang, "A comparative study of the maximum power point tracking methods for pv systems," *Energy Conversion and Management*, vol. 85, pp. 809–816, 2014.

[8] A. A. Kulaksız and R. Akkaya, "A genetic algorithm optimized ann-based mppt algorithm for a stand-alone pv system with induction motor drive," *Solar Energy*, vol. 86, no. 9, pp. 2366–2375, 2012.

[9] N. Yamamura, M. Ishida, and T. Hori, "A simple wind power generating system with permanent magnet type synchronous generator," in *Power Electronics and Drive Systems, 1999. PEDS'99. Proceedings of the IEEE 1999 International Conference on*, vol. 2. IEEE, 1999, pp. 849–854.

[10] K. Al Khuffash, L. A. Lamont, and Y. Abdel-Magid, "Using artificial neural networks to enhance the accuracy of the photovoltaic simulation model," in *Environment and Electrical Engineering and 2017 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–5.

[11] K. Kumar, S. Sharma, L. Jain, and R. Al Khaimah, "Standalone photovoltaic (pv) module outdoor testing facility for uae climate," *Submitted to CSEM-UAE Innovation Center LLC*, 2007.

[12] S. A. Kalogirou, "Artificial neural networks in renewable energy systems applications: a review," *Renewable and sustainable energy reviews*, vol. 5, no. 4, pp. 373–401, 2001.

[13] P. Petchjatuporn, W. Ngamkham, N. Khaehintung, P. Sirisuk, W. Kiranon, and A. Kunakorn, "A solar-powered battery charger with neural network maximum power point tracking implemented on a low-cost pic-microcontroller," in *TENCON 2005 2005 IEEE Region 10*. IEEE, 2005, pp. 1–4.

[14] Y. Sun, S. Li, B. Lin, X. Fu, M. Ramezani, and I. Jaithwa, "Artificial neural network for control and grid integration of residential solar photovoltaic systems," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1484–1495, 2017.

Electricity Theft Detection to Reduce Non-Technical Loss using Support Vector Machine in Smart Grid

Rafia Nishat Toma[1], Md Nazmul Hasan[1], Abdullah-Al Nahid[1], and Bo Li[2]

[1]Electronics and Communication Engineering Discipline, Khulna University, Khulna-9208, Bangladesh
[2]Beijing University of Posts and Telecommunications, Automation Dept., Beijing, 100876, China
{rafiatoma.eceku@gmail.com, nazmul@ece.ku.ac.bd, nahid.ece.ku@gmail.com, deepblue.lb@gmail.com}

***Abstract*—Among the various reason behind Non-technical losses in smart grid, losses due to electricity theft have become major apprehension in power system industries. A significant amount of consumption of electricity in a fraudulent manner decrease the supply quality, increase generation load and the real consumers have to pay excessive electricity bill that finally affects in overall economic condition. The advance metering infrastructure (AMI) has the ability to monitor the consumption detail of every consumer, record the consumption pattern, billing them as well as find any types of abnormalities. The communication capabilities of smart grid have facilitated the utilities to store their consumers' consumption patterns. With this database it is possible to formulate a theft detection model by machine learning algorithm by analyzing the recorded data of smart meter. In this paper, Support vector machine (SVM), one of the prominent machine learning classifiers applied with principle component analysis to train the data collected from smart meter and calculate the prediction accuracy with the test data. Here the principle component analysis reduces the dimension of the data to make the further processing less complex. After that, by using grid-search method the best-suited meta-parameters for SVM has been selected where the almost 90% accuracy rate is achieved. The classification results with respect to different meta-parameters of SVM (regularization parameter C and γ) have presented. Moreover, the obtained result specifies that the applied techniques possesses higher accuracy and less false positive rate for real time consequences.**

***Index Terms*—Classification techniques, electricity theft, grid search method, principle component analysis (PCA), smart grid, support vector machine (SVM).**

I. Introduction

In the last few years, the new idea of a smart grid (SG) became a critical issue in the contemporary world and has attracted cumulative attention of energy users', policy makers and engineers. The combined electricity network of computers and power system infrastructure which monitor and manage energy use as well as can incorporate the nature and activities of all consumers connected to it intelligently is known as Smart Grid. The grid is a collection of digital systems and rejuvenation of current electrical system which extends customers' and electrical companies' ability to observe, control, and foretell energy usage [1]. The overall process is designed to efficiently deliver sustainable, economical, and secure electricity supplies and reduce loss due to energy wastage [2].

Among various technologies, the key technical areas used in smart grid operation are integrated communications, advanced components, sensing and measurement, innovative control methods and finally, improved interfaces and decision support. To execute the operation of Smart grid through these technologies, some basic components like electric power generators, electric power substations, transmission and distribution lines, controllers, collector nodes, smart meters, and distribution and transmission control centers have been implemented [2].

Among the devices use in smart grid, smart meter is one of the most important devices. It provides two way communication between subscribers and utilities to exchange informations such as voltage, current, phase angle, real time energy consumption data and control information [3]. The combination of the unique meter identifier of smart meter, the electricity consumption value, the data timestamp provides the smart meters' data which helps the utilities to optimize end users' bill and also help to control the power transmission and consumption rate.

Though many countries have already started the implementation of smart grid system, the power utilities suffered a dramatic economic loss due to some technical and non-technical issues. The main reason of technical loss (TL) due to the internal electrical resistance and affected components like transformers, generators, and transmission lines [4].

On the other hand Non-technical losses (NTL), also known as Transmission and Distribution losses can be estimated with various types of techniques. Fraud due to the electricity theft, errors and delay in meter reading and billing, irregular bill submission, power stealing by any means, faulty energy meters or un-metered supply and ignored unpaid bills are the main reasons of NTL. The energy supplied to the consumers by stealing of power does not measured by the energy meter which makes a variation between the energy supplied through the meters and the registered amount of energy by the meters [1].

Undoubtedly, electricity is a key input for the development for many underdeveloped and developed countries as many production processes solely depend on it. Because of the NTLs significant damage in economies which results loss of profit and revenue of utility companies as well as reduces the consistency and steadiness of electric power grid. Among the causes of NTL, electricity theft poses a serious concern for the power system utilities as they have to accept excessive loses due to electricity theft. The real time fraud detection

and current action is the only way to get rid from this. By analyzing of all customers data on grid, the detection of illegal consumption is possible. As the SG has been implemented many countries and the ICT field becomes more advance, various methods have been already proposed to identify and restrict the electricity theft. Most of the current methods focus on detecting fraudulent consumers at distribution phase with limited set of parameters. Keeping these issues in mind, this paper focuses on presenting a framework to measure the accuracy of machine learning model to detect electricity theft. The rest of the paper is organized as follows. Section 2 summarizes a brief literature review. Section 3 presents the methodology that contains brief description of Support Vector Machine (SVM), Principle Component analysis and the overall working procedure. While section 4 provides the experiment results generated using SVM classification. In the end, Section 5 presents the conclusions of this paper.

II. Previous work

The NTL, especially electrical theft poses a serious concern in the power system domain because of its huge amount of economic losses. In the recent times, many researcher s got involved to develop and propose several techniques through an extensive and fecund research to detect as well as minimize electricity theft. Energy theft detection system (ETD) can be categorized into three major groups, as state, game theory and classification based [5]. The state based detection was proposed in [6]–[8] by using improved sensors and the devices to improve the accuracy of ETD and reduce the false positive rate. Due to the higher cost of equipment and software, this model is quite difficult to implement.

On the other hand, because of the expansion of internet and emergent concern of cyber-attacks, the ICT can play a vital role to avoid electrical theft [9]. To make data computation, transmission, and storage protected, Wei et al. [10], [11] proposed a security frame work by using Markle Hass tree concept. The security threats can be linked with the metering infrastructure which is discussed in [12], and various countermeasures as suggested tackling them to ensure data confidentiality, integrity, and availability in smart grid.

Monitoring of consumers' load profile by data from electric meter was one of the effective suggested criteria to find fraudulent customers. Nagi et al. [13] applied data mining method by calculating the average daily consumptions of end users over a two-year period by finding the abrupt changes in load profile. Angelos et al. [14] computed average consumption, maximum consumption, standers deviation, sum of the inspection remarks and the average consumption of neighborhood were selected from six months usages profile and detect the abnormalities in data pattern.

In the field of theft detection, the support vector machine has been widely used technique in recent times. The SVM has been applied to detect anomalies on manually collected data by Joker et al. [15]. Based on the irregularities in data frame, [16] presented a SVM based approach to detect fraud consumers. To make the detection more effective, a model of ANN incorporate with SVM has been proposed by Depru et al. [17]. Not only this, Nagi et al. [18] implemented fuzzy logic with SVM to improve the overall performance. An improved coding technique was proposed and both SVM and a rule-engine based algorithm were applied on encoded data to improve efficiency in [19]. Moreover, different techniques such as one-class SVM, optimum path forest and C4.5 decision tree has been combined by Martino et al. [20] to achieve higher accuracy.

III. Methodology

This work classified normal and theft electricity data based on supervised learning technique. The overall working flow is described in Fig. 1.

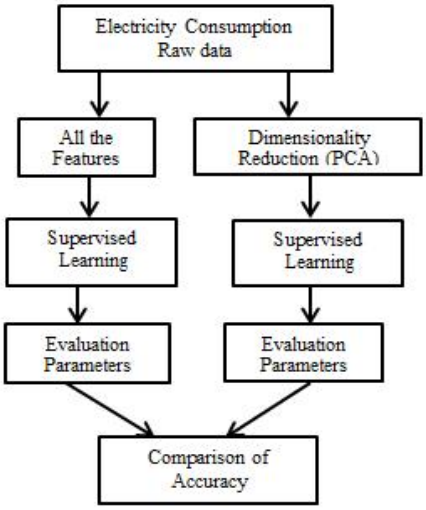


Fig. 1. Proposed framework for comparing accuracy

A. Support Vector Machine

Support Vector Machine (SVM) is powerful and widely used learning algorithm used for both supervised classification or regression tasks. In this work SVM is used for theft classification. SVM classifies labeled data by using a separating hyperplane as depicted in Fig.1 [21].

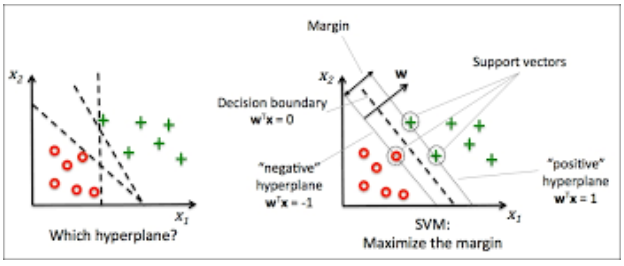


Fig. 2. Concept of Support Vector Machine [21]

The figure depicts a visualization of simple classification involving only two class which are linearly separable. Here

the positive and the negative hyperplanes can be formulated as follows:

$$w_0 + w^T x_{pos} = +1 \quad (1)$$

$$w_0 + w^T x_{neg} = -1 \quad (2)$$

Subtracting these two equations would results the following equation:

$$\frac{w^T(x_{pos} - x_{neg})}{\|w\|} = \frac{2}{\|w\|} \quad (3)$$

The left side of equation 3 represents the margin between positive and negative hyperplanes. The goal of SVM is to maximize this margin. For more complex data points which are not linearly separable as shown in Figure 2, SVM uses kernel method to create hyperplanes. The kernel method utilizes a mapping function , to transform the original features into a higher dimensional space where it is possible to separate them by linear hyperplanes which is also shown in Figure 2.

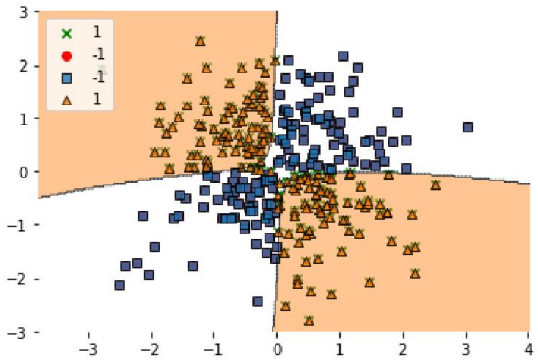


Fig. 3. Support vector Machine for linearly inseparable

Performance of SVM classifier can be optimized by selecting some tuning parameters one is the regularization parameter or C parameter. In simple terms this parameter defines the amount of misclassification in training data. A higher value of C will create small margin hyperplanes which will try to classify all training data correctly. Whereas, smaller value of C will produce hyperplanes with larger margin but doing so it will allow some outliers or misclassified points in training data. This is depicted in Figure 3. Higher value of C can lead to overfitting the training data.

Gamma is another tuning parameter of svm that is used for rbf kernel. Higher the value of gamma the tighter the hyperplane boundary. Large gamma leads to high bias and low variance models, and vice-versa. A classifier with high gamma value might have a high generalization error on test data. Figure 4 displays how change in gamma value changes the decision regions.

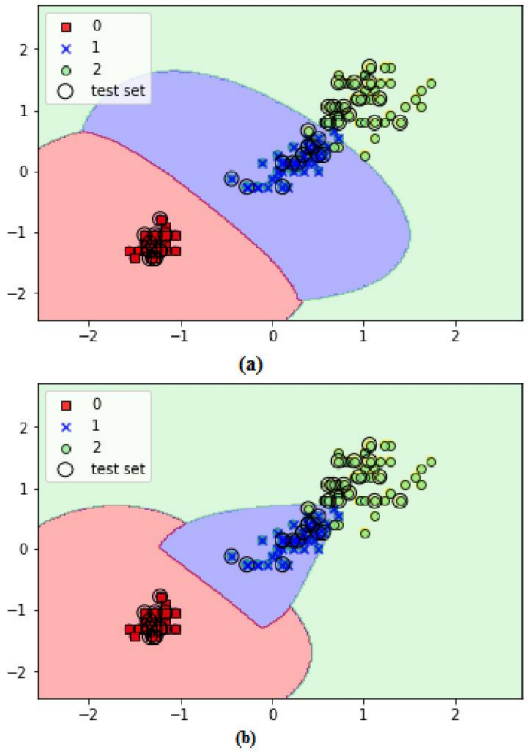


Fig. 4. Decision regions for (a) C = 1 (b) C = 1000

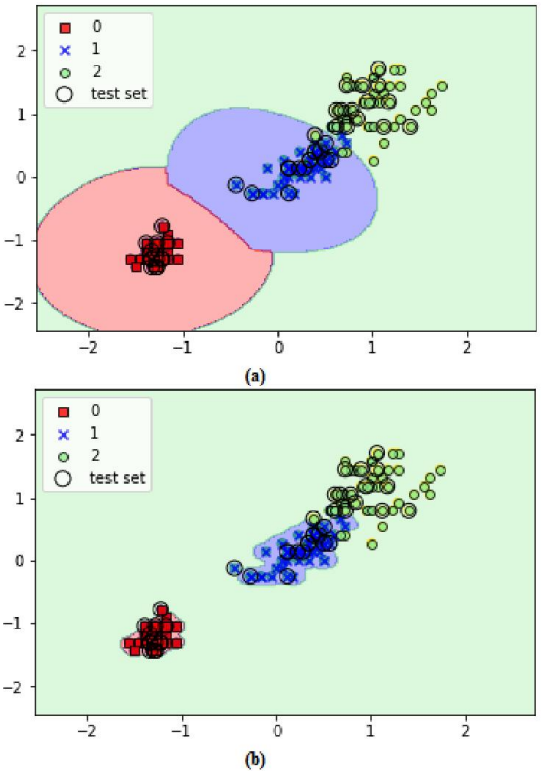


Fig. 5. Decision regions for (a) gamma = 1 (b) gamma = 1000

B. Principle Component Analysis

A huge number of features for each training instance are involved in any machine learning problem, which makes the training awfully slow and also creates over fitting problem. It is mandatory to reduce the number of features/dimensions to enhance the model's performance and facilitates to find out an optimal solution for any respective machine learning problem. A well-known method for this task is Principal Component Analysis (PCA), that aims is to reduce the dimensionality of data set contains many variable interrelated with each other, by keeping the variation present in the data set. It is statistical method that uses an orthogonal transformation to transform a set of interpretations of possibly correlated variables into another set of linearly uncorrelated variable values named principle components. The data set must be scaled on which the PCA techniques will be applied. To implement PCA on a data set, first the data has to be normalized, after that the value of covariance matrix, eigenvalues and eigenvectors needs to calculate. At the next step from the data we have to choose components and form a feature vector for next step. By forming principal components with all the calculative values this process comes to an end.

C. Data Acquisition

The data we used in this paper includes approximately 10,000 users daily real electricity consumption data over the year of 2015 which represents in time series. The dataset is released by State Grid Corporation of China. The dataset is available at (http://www.sgcc.com.cn/). The information in this data set includes customers consumption signature in daily basis.

D. Data Preprocessing

The data contains almost 10,000 users' data over one year. The real world data contains very inconsistent data, there was some missing data as well as the value of some data was too high or immeasurable. To reduce this problem, data preprocessing is applied, which is one of the important parts in data mining techniques. In this process, to reduce noise we have to apply statistical method and converted the data to normal consumption values.

E. Data Normalization

For further processing, the load data needs to be represented using a normalized scale for the SVM classifier [22], [23]. Though we have the data for 365 days, the daily average kWh consumption feature data is normalized as follows:

$$\mathrm{N} = \frac{\mathrm{L} - \min(\mathrm{L})}{\max(\mathrm{L}) - \min(\mathrm{L})} \tag{4}$$

where L represents the current kWh consumption of the customer, and min(L) and max(L) represent the minimum and maximum values over 1 year consumption feature set. The load profiles of customers is found with each load profile being represented by the 365 normalized daily average kWh consumption features.

F. Choice of Appropriate kernels

Karnel can be defined as the set of mathematical functions which is used to set SVM algorithm. The main aim of kernel function is to take data as input and transform it in any required form. Based on the mathematical interpretation, the functions can be various types , such as linear, nonlinear, polynomial, radial basis function (RBF), and sigmoid. To find the best result, one have to choose the best result that fits the respective problem [24].

G. Choice of Modeling Parameters

The selection of two SVM model parameter (C, γ) are plays a vital role to calculate accuracy of classification problem. Thus, selecting the perfect values for for these parameters is an important task. A lot of practical approaches, asymptotical optimization [25], cross-validation (CV), and grid search [26] are available for selecting and optimizing parameters.

H. Train and Test the classifier

Once the selection of kernel and modeling parameters has been defined, the data needs to be divide into training and test sets. The SVM classifier is trained first with the training data and then new test data is classifying based on the trained classification model. Depending on these factors, the data is evaluated and the related results of this classification are described in the next section.

I. Data Compression

Principal component analysis (PCA) is used to reduce problems related to high dimensional and co linearity of any high volume data set, while retaining the variability of the data set. PCA is a classical statistical method for transforming attributes of a data set into a new set of uncorrelated attributes but keep the data type unchanged. Generally, with this initial PCA step, a significant progress in the performance of machine learning methods can be achieved. [27].

J. Grid Search method

Grid search is a search mechanism based on deined subset of hyper-parameter space. By defining minimal value, maximal value, and number of steps, the hyper-parameters are specified [28]. Grid search optimizes the SVM parameters (C, γ, degree) as a performance matrix. To identify good hyper-parameter combination is the main objective, so the classifier can predict test data accuracy more appropriately.

IV. Results and Discussions

To apply SVM classifier, we need to preprocess the data to convert it into SVM compatible format. The next step is to data normalization which is used to standardized the data using sample mean and standard deviation for each feature for optimal performance. After that, we split the data set into training and test data. First, we extract the accuracy rate with respect to number of samples without applying PCA. The training sample accuracy is quite high for the number of sample 0 to 2500.

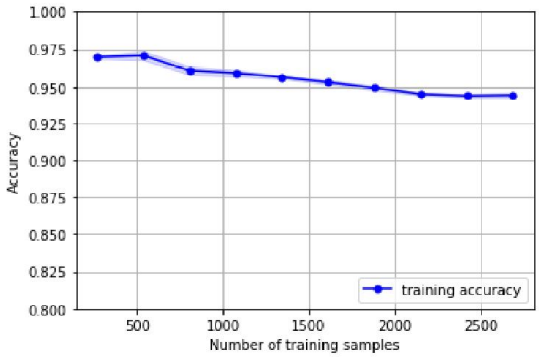


Fig. 6. Learning Curve for the data set

PCA acts as a key component for increasing accuracy for high dimension data. As we have to work with almost 9700×366 dimension data, apply PCA will reduce computational complexity. After that, we have to choose an appropriate kernel with the modeling parameter. In this work, RBF kernel with modeling parameters C and gamma (γ) are taken. For a range of C and gamma (γ),the training data and test data accuracy has been illustrated.

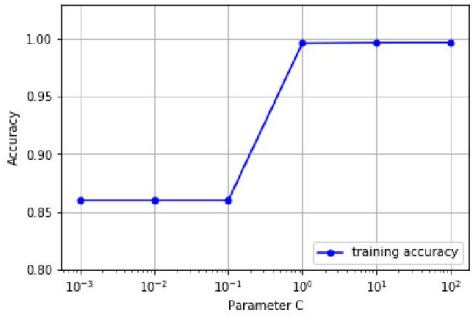


Fig. 7. Validation Curve for Parameter C

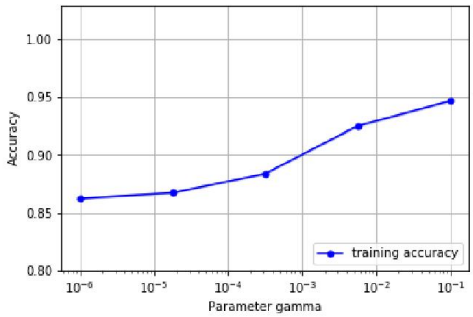


Fig. 8. Validation Curve for Parameter gamma

The confusion matrix of this data set for applying SVM is given in Fig. 9. From this matrix we get true positive(TP), false positive (FP), true negative (TN) and false negative (FN) values, where the ratio of misclassyfied data is quite low for both cases of SVM, with or without PCA [29], [30].

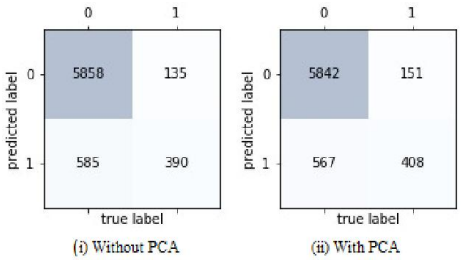


Fig. 9. Confusion matrix

The SVM is trained on the basis of parameters C and gamma (γ). Three different ratio of training and test data as 60-40, 70-30 and 80-20 has been taken in order to test the accuracy, that is given in Table I. The highest accuracy we achieved is 90%.

TABLE I
RESULT FOR DIFFERENT RATION OF TRAINING AND TEST DATA

Training date/Test data	Precision	Recall	f1-score	Accuracy
60/40	0.89	0.90	0.89	89%
70/30	0.89	0.90	0.89	90%
80/20	0.89	0.90	0.89	90%

Principal component analysis along with SVM improves the accuracy of machine learning model in the classification for any high dimensional data set [27]. In this work, we have in total 365 components to analysis. Because of applying PCA, we achieve the maximum accuracy (90%) with 50 components only, which finally reduce the complexity of overall model.

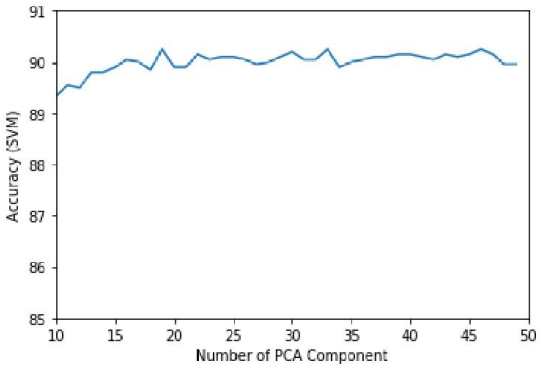


Fig. 10. Accuracy calculation for PCA components

Finally, the grid search method has applied on SVC to estimate the optimum classification by optimizing the kernel and other meta-parameters (C and γ). For this study, a range of values for both parameters has taken into account to detect the optimum value. Finally, it is found that the optimal parameters C = 100 and γ = 0.001 for RBF kernel where the accuracy is approximately 90%.

The dataset used in this work is relatively new and until now only two research works have been performed on this dataset. Multi-scale DenseNet is investigated in [5] and other used convolutional neural network [31]. In this work shallow machine learning algorithm is used. We have used different performance matric (accuracy) whereas the authors of the previous works used AUC curve. As the performance metric is different, so, it is not possible to compare our result with the other ones. In future, other algorithms and similar performance metrics can be investigated for comparison.

V. Conclusion

Electrical theft leads to huge losses to the utilities in the power sector. The illegal use of electricity through tapping by consumers is the main reason behind this. To identify the malevolent consumers and find the accuracy of smart meter data with Support Vector Machine (SVM) has been presented in this paper. The consumption data provides as the input of the SVM classifier to train the collected data set first. Then classifier applied to classify the customers as normal or fraud. The experimental results obtained show that the by choosing the optimal meta-parameters of SVM effectively by grid search method can increase accuracy of the system. The optimal value of C is 100 and gamma is 0.001. The SVM classifier identifies fraudulent consumers with an accuracy of 90% and false positive rate as low as almost 9%.

References

[1] X. Yu, C. Cecati, T. Dillon, and M. G. Simoes, "The new frontier of smart grids," *IEEE Industrial Electronics Magazine*, vol. 5, no. 3, pp. 49–63, 2011.

[2] A. Mavridou and M. Papa, "A situational awareness architecture for the smart grid," in *Global Security, Safety and Sustainability & e-Democracy*. Springer, 2012, pp. 229–236.

[3] J. Zheng, D. W. Gao, and L. Lin, "Smart meters in smart grid: An overview," in *Green Technologies Conference, 2013 IEEE*. IEEE, 2013, pp. 57–64.

[4] P. Glauner, J. A. Meira, P. Valtchev, R. State, and F. Bettinger, "The challenge of non-technical loss detection using artificial intelligence: A survey," *arXiv preprint arXiv:1606.00626*, 2016.

[5] B. Li, K. Xu, X. Cui, Y. Wang, X. Ai, and Y. Wang, "Multi-scale densenet-based electricity theft detection," *arXiv preprint arXiv:1805.09591*, 2018.

[6] C.-H. Lo and N. Ansari, "Consumer: A novel hybrid intrusion detection system for distribution networks in smart grid," *IEEE Transactions on Emerging Topics in Computing*, vol. 1, no. 1, pp. 33–44, 2013.

[7] Z. Xiao, Y. Xiao, and D.-C. Du, "Non-repudiation in neighborhood area networks for smart grid," *IEEE Communications Magazine*, vol. 51, no. 1, pp. 18–26, 2013.

[8] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, and S. Zonouz, "A multi-sensor energy theft detection framework for advanced metering infrastructures," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1319–1330, 2013.

[9] R. Gholami, X. Guo, S.-Y. T. Lee *et al.*, "Information and communications technology (ict) international spillovers," *IEEE transactions on engineering management*, vol. 56, no. 2, pp. 329–340, 2009.

[10] L. Wei, H. Zhu, Z. Cao, W. Jia, and A. V. Vasilakos, "Seccloud: Bridging secure storage and computation in cloud," in *Distributed Computing Systems Workshops (ICDCSW), 2010 IEEE 30th International Conference on*. IEEE, 2010, pp. 52–61.

[11] L. Wei, H. Zhu, Z. Cao, X. Dong, W. Jia, Y. Chen, and A. V. Vasilakos, "Security and privacy for storage and computation in cloud computing," *Information Sciences*, vol. 258, pp. 371–386, 2014.

[12] S. N. Lighari, B. B. Jensen, A. A. Shaikh *et al.*, "Attacks and their defenses for advanced metering infrastructure," in *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2014 6th International Congress on*. IEEE, 2014, pp. 148–151.

[13] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and M. Mohamad, "Nontechnical loss detection for metered customers in power utility using support vector machines," *IEEE transactions on Power Delivery*, vol. 25, no. 2, pp. 1162–1171, 2010.

[14] E. W. S. Angelos, O. R. Saavedra, O. A. C. Cortés, and A. N. de Souza, "Detection and identification of abnormalities in customer consumptions in power distribution systems," *IEEE Transactions on Power Delivery*, vol. 26, no. 4, pp. 2436–2442, 2011.

[15] P. Jokar, N. Arianpoo, V. C. Leung *et al.*, "Electricity theft detection in ami using customers' consumption patterns." *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 216–226, 2016.

[16] J. Nagi, A. Mohammad, K. Yap, S. Tiong, and S. Ahmed, "Non-technical loss analysis for detection of electricity theft using support vector machines," in *Power and Energy Conference, 2008. PECon 2008. IEEE 2nd International*. IEEE, 2008, pp. 907–912.

[17] S. S. S. R. Depuru, L. Wang, V. Devabhaktuni, and P. Nelapati, "A hybrid neural network model and encoding technique for enhanced classification of energy consumption data," in *Power and Energy Society General Meeting, 2011 IEEE*. IEEE, 2011, pp. 1–8.

[18] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and F. Nagi, "Improving svm-based nontechnical loss detection in power utility using the fuzzy inference system," *IEEE Transactions on power delivery*, vol. 26, no. 2, pp. 1284–1285, 2011.

[19] S. S. S. R. Depuru, L. Wang, V. Devabhaktuni, and R. C. Green, "High performance computing for detection of electricity theft," *International Journal of Electrical Power & Energy Systems*, vol. 47, pp. 21–30, 2013.

[20] M. Di Martino, F. Decia, J. Molinelli, and A. Fernández, "Improving electric fraud detection using class imbalance strategies." in *ICPRAM (2)*, 2012, pp. 135–141.

[21] S. Raschka and V. Mirjalili, *Python machine learning*. Packt Publishing Ltd, 2017.

[22] N. Sikder, K. Bhakta, A. Al Nahid, and M. M. Islam, "Fault diagnosis of motor bearing using ensemble learning algorithm with fft-based preprocessing," in *2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*. IEEE, 2019, pp. 564–569.

[23] K. Bhakta, N. Sikder, A. Al Nahid, and M. M. Islam, "Fault diagnosis of induction motor bearing using cepstrum-based preprocessing and ensemble learning algorithm," in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. IEEE, 2019, pp. 1–6.

[24] A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar, and S. Mishra, "Decision tree and svm-based data analytics for theft detection in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 1005–1016, 2016.

[25] V. Cherkassky and Y. Ma, "Practical selection of svm parameters and noise estimation for svm regression," *Neural networks*, vol. 17, no. 1, pp. 113–126, 2004.

[26] C.-W. Hsu, C.-C. Chang, C.-J. Lin *et al.*, "A practical guide to support vector classification," 2003.

[27] T. Howley, M. G. Madden, M.-L. OConnell, and A. G. Ryder, "The effect of principal component analysis on machine learning accuracy with high dimensional spectral data," in *Applications and Innovations in Intelligent Systems XIII*. Springer, 2006, pp. 209–222.

[28] I. Syarif, A. Prugel-Bennett, and G. Wills, "Svm parameter optimization using grid search and genetic algorithm to improve classification performance," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 14, no. 4, pp. 1502–1509, 2016.

[29] A.-A. Nahid, M. A. Mehrabi, and Y. Kong, "Frequency-domain information along with lstm and gru methods for histopathological breast-image classification," in *2017 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. IEEE, 2017, pp. 410–415.

[30] A.-A. Nahid and Y. Kong, "Local and global feature utilization for breast image classification by convolutional neural network," in *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. IEEE, 2017, pp. 1–6.

[31] Z. Zheng, Y. Yang, X. Niu, H.-N. Dai, and Y. Zhou, "Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1606–1615, 2018.

A Machine Learning Approach to Protect Electronic Devices from Damage Using the Concept of Outlier

Sunanda Das, Muhammad Sheikh Sadi, Md. Ahsanul Haque, Md. Milon Islam
Department of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
sunanda@cse.kuet.ac.bd, sadi@cse.kuet.ac.bd, ahsanulhaque865@gmail.com, milonislam@cse.kuet.ac.bd

***Abstract*— Most of the appliances that we used in our everyday life are electronic devices, i.e. TV, Air Conditioner, Refrigerator, etc. Excessive voltage, current, temperature, etc. can harm the devices and in extreme cases, the devices can be completely damaged. We proposed a system to monitor the electrical behaviors of the devices in real-time. The system is trained with an unlabeled dataset and capable of identifying outliers. We have used a clustering technique, i.e. Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to learn the label of a dataset and then apply machine learning algorithms, i.e. Support Vector Machine (SVM) and Decision Tree to predict the label of the new data. From the prediction, the system determines whether the device is operating in safe mode or not. In this work, we have achieved the accuracy of 98.61% in detecting outliers using SVM with 'rbf' kernel. Hence, if the device operates beyond safe mode, we can shut down the device.**

Keywords—Outlier; Outlier Detection; Clustering; Density-based Clustering; DBSCAN; Support Vector Machine; Decision Tree.

I. INTRODUCTION

People depend on the use of electronic devices in their everyday life. Most of the cases, the end users cannot identify the poor performance of the device and can only understand the weak performance when the device isn't working any longer. Hence, we need to find a way to protect these devices.

An outlier is an observation point or a set of observation points that have different properties and are inconsistent with other observations [1]. They are generated by any kind of disturbances in the system. Hence, outliers indicate the erroneous condition of the system. Outliers detection techniques try to find unusual patterns in the system. There are various approaches like distance-based approaches, depth-based approaches, density-based approaches, deviation-based approaches, etc. to detect outliers in the system [2].

In this work, a density-based approach, i.e. DBSCAN is used to detect outlier. DBSCAN is a hierarchical clustering method and can determine the number of clusters automatically. Moreover, DBSCAN can successfully identify arbitrary shaped cluster. This work describes an advanced system which can monitor the electrical characteristics of the device and identify disturbances like overvoltage, flicker, etc. This work considers these disturbances as the outlier and proposes an effective way to detect it. Then, supervised learning algorithms, i.e., SVM, Decision Tree are used for training a model to predict a new sample. Hence, we can decide whether the new sample data is an outlier or not.

The remaining part of the paper is organized as follows: Section II demonstrated the related works that have been done in this field. The proposed methodology with some theoretical explanations is investigated in Section III. The result is illustrated in details with some performance measures in Section IV. The conclusion of the paper is drawn in Section V.

II. RELATED WORK

For a system, outlier or anomaly may come at any time and can cause significant harm to the system. In case of protecting the system, we need to detect the outliers. In recent years, outlier detection has drawn considerable attention in the research community. For identifying outliers in feature spaces, Eskin et al. [3] proposed a system using different machine learning techniques. These techniques include one class SVM, K-nearest neighbor, and cluster-based estimation. They used these algorithms to identify the points that are in the sparse areas. Among the three algorithms, they obtained the detection rate of 98% and a false positive rate of 10% for SVM over the KDD Cup 1999 Data. Ester et al. [4] investigated the effectiveness of DBSCAN. They concluded that DBSCAN is good at detecting outlier and discovering a cluster of arbitrary shape. Their experiment on SEQUOIA 2000 benchmark data shows that DBSCAN outperforms the well-known algorithm CLARANS (Clustering Large Applications based on RANdomized Search) by a factor of at least 100 regarding efficiency.

For unsupervised anomaly detection, Kingsly Leung and Christopher Leckie [5] presented fpMAFIA, a new density-based and grid-based clustering algorithm over the 1999 KDD Cup data set. Their results indicate that their new approach fpMAFIA able to achieve a high detection rate while it suffers from a high false positive rate compared to the other methods. Bakar et al. [6] experimented different techniques like linear regression, control chart and Manhattan distance to detect outliers. Their experimental results demonstrate that the control chart technique gives reliable results in detecting outliers than linear regression. Ramaswamy et al. [7] demonstrated a profoundly efficient partition-based algorithm for discovering outliers. Their algorithm first partitions the input data set into disjoint subsets and as soon as it concludes that they do not include outliers then, the algorithm prunes entire partitions. Breunig et al. [8] talked about LOF (Local Outlier Factor). According to them, each object has its own LOF. The factor depends on how isolated the object is regarding its surrounding neighbors. Their outcome shows that the approach is very assuring in recognizing significant local outliers.

Arning et al. [9] presented a linear algorithm to detect deviation in data. Their investigation reveals that the effectiveness of the algorithm mainly depends on the

dissimilarity function. A dissimilarity function catches how different is a new data from the data items seen previously. For distinguishing outliers, Jiang et al. [10] suggested a two-phase clustering technique. In the first phase, they used a modified k-means algorithm. In the second phase, they applied an OFP (Outlier Finding Process) for finding outliers in the clusters that are generated by phase one. Hawkins et al. [11] suggested a Replicator Neural Networks (RNN) based outlier detection approach. An RNN is trained on sampled dataset so that it can predict the new data. This approach identifies cluster labels which are used to interpret the generated outliers.

III. The Proposed Methodology for Outlier Detection

The primary objective of the work is to determine the outliers of the electronic devices. Fig. 1 illustrates the detailed design of the proposed methodology. The main steps of the methodology include dataset preparation, preprocessing, DBSCAN clustering to detect the outliers and learn the label of the outlier and normal data, from that labeled dataset, predict new samples using Support Vector Machine (SVM) and Decision Tree.

A. Dataset Preparation

Ohm's law illustrates the relationship between current, voltage, and resistance. It informs us that the current i, passing through a circuit is directly proportional to the voltage v, and inversely proportional to the resistance r.

$$i = v/r \tag{1}$$

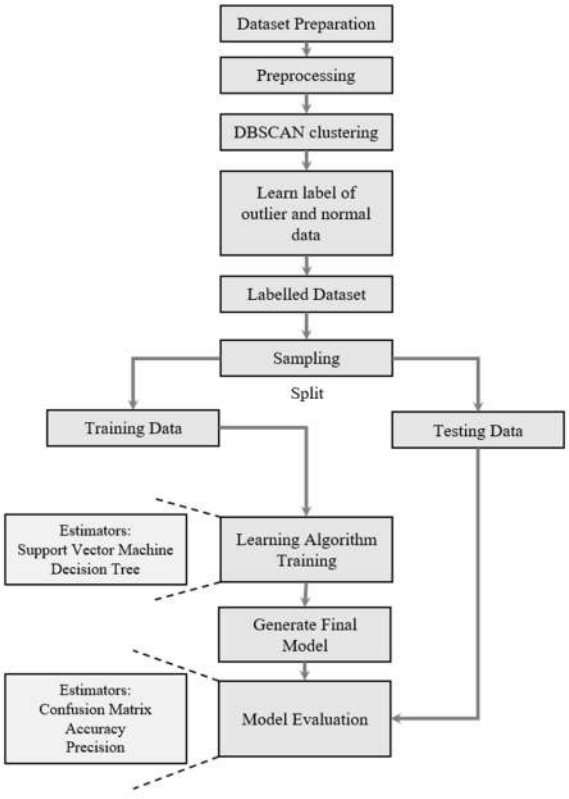


Fig. 1. Proposed methodology of the system.

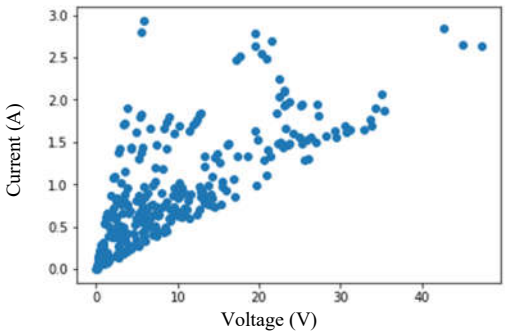


Fig. 2. The generated dataset for the system.

A synthetic dataset is prepared for the work. For preparing the dataset, we have used the (1) and randomly generate v and r and then calculate the corresponding i. The dataset contains two features, i.e., voltage (v) and current (i). Fig. 2 illustrates the generated dataset.

B. Preprocessing

The real-world data can be noisy and sometimes there are missing values in the dataset. The amount of noise and the number of missing values mainly depend on how we collect the data. A large amount of the datasets that are used to solve a machine learning problem need to be handled carefully [12]. For preprocessing, we have used z-score normalization. We get the rescaled features after using the standardization i.e. normalization. Standardizing the features is vital for our dataset as the dataset has measurements on different units. The z-score of a sample x is estimated as:

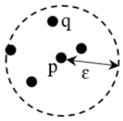
$$z = (x-\mu)/\sigma \tag{2}$$

where μ denotes the mean of the training samples, and σ denotes the standard deviation of the training samples.

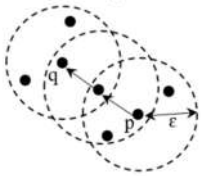
C. Clustering and Learning the label of the dataset

Clustering is mainly used for grouping objects that have the same kind of properties [13]. There are many techniques for clustering objects depending on dataset's perspective. In this work, DBSCAN is applied for clustering purpose. DBSCAN discovers neighbors, i.e. the member of the same cluster by using two parameters, i.e. MinPts and radius ε. The formation of a cluster depends on the notions of directly density-reachability, density-reachability and density-connectivity of points [14]. A cluster can be formed as the maximal set of 'density connected points' in the feature space.

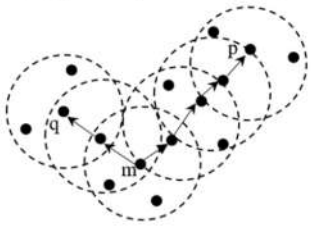
Let, consider two points p, q. From Fig. 3 it can be inferred that, q is directly density-reachable from p regarding ε if and only if q is in p's ε-neighborhood. q is density reachable from p regarding ε if there is a series of points p_1, …, p_n in such a way that $p_1 = p$, $p_n = q$ and for each i= 2, . . ., n it is true that p_i is directly density-reachable from p_{i-1} regarding ε. q is density connected from p regarding ε if and only if there is a point m such that q is density reachable from m and p is also density reachable from m [15].



(a) q is directly density reachable from p



(b) q is density reachable from p



(c) q is density connected from p

Fig. 3. The concepts of directly density reachability, density reachability and density connectedness of points.

DBSCAN defines different classes of points as followings

- Core point
- Border point
- Outlier

Let, u be another new point. u can be a core point if its neighborhood defined by radius ε contains at least or more points than the number of points defined by parameter MinPts. u can be a border point if it a member of a cluster and its ε neighborhood contain less points than MinPts but u is still density reachable by other points in that cluster. u will be an outlier if it is not a core point and at the same time not a border point. Fig. 4 represents the notion of core point, border point, and outlier. Hence, it can be realized that, all the points inside the same cluster are respectively density connected and if a point is density reachable from any member point of that cluster the point is also a member of that cluster as well.

We have used DBSCAN in our dataset. It finds the number of clusters automatically and also finds the outlier. It also returns the labels of the calculated clusters and outlier. Hence, we learn the label and use a predictive learning algorithm, i.e. SVM (Support Vector Machine) and Decision Tree to predict the label of the newly arrived data.

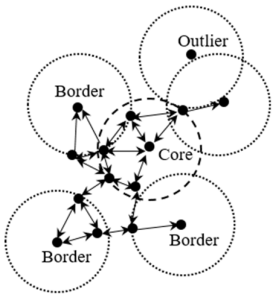


Fig. 4. The concept of core point, border point and outlier.

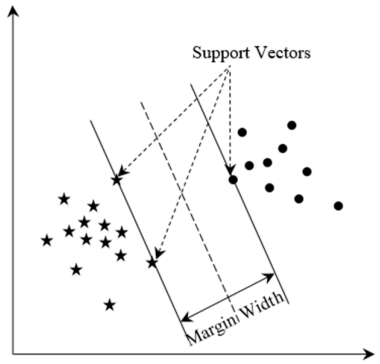


Fig. 5. Support vector points and margin width of hyperplane of SVM.

D. Prediction of outlier

Support Vector Machine (SVM) is a supervised learning algorithm in machine learning which is used for the classification problem. The main ideology behind the concept of SVM is finding the optimal hyperplane that confidently separates different classes [16].

The closest points to the hyperplane consist the support vector points. The distance between the support vector points and the hyperplane is called the margin. Fig. 5 shows support vector points and margin width of hyperplane for SVM. The support vector points play an important role for finding the hyperplane that maximizes margin width. All other points are irrelevant for finding the hyperplane.

Let, n be the number of dimensions. Thus, the equation of the hyperplane can be given as the following

$$\begin{aligned} y &= w_0 + w_1x_1 + w_2x_2 + w_3x_3 + \ldots \\ &= w_0 + \sum_{i=1}^{n} w_ix_i \\ &= w_0 + w^TX \\ &= b + w^TX \end{aligned} \tag{3}$$

where y is the outcome, b is the biased term ($b = w_0$), x_i are the attribute values. The weights w_i will be learned by the algorithm. In (3) w_i are the parameters that will determine the hyperplane [17].

Decision Tree is a very popular algorithm for classification in machine learning. To develop a tree, we need to choose features and conditions for splitting along with stopping criteria [18]. For classification, different split point is considered using a cost function. The split is selected that has the lowest cost. Gini score is considered as a cost function for classification.

$$G = sum(p_k * (1 - p_k)) \tag{4}$$

Here, p_k is the proportion of same class inputs present in a particular group. Gini score decides whether a split is good or not by estimating how mixed the response classes are in the groups formed by the split. When a group contains all the member from the same class then p_k is either 0 or 1 and $G = 0$. The performance of Decision Tree can be increased by pruning. It involves removing the branch that has less important features.

As we already generate label using DBSCAN, we train the dataset using SVM and Decision Tree. For model evaluation confusion matrix, accuracy and precision are used. Confusion matrix is used to decide the performance of a classification model. Accuracy, precision evaluates the

output quality of the classifier. Confusion matrix is formed using T_P, T_N, F_P, F_N. where,

T_P = *True Positive*
T_N = *True Negative*
F_P = *False Positive*
F_N = *False Negative*

For calculating accuracy (*Acc*) and precision (*Prec*), the following formulas are used.

$$Accuracy\ (Acc) = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (5)$$

$$Precision\ (Prec) = \frac{T_P}{T_P + F_P} \quad (6)$$

IV. Result Analysis

The system is implemented in the python environment using 'scikit-learn' which is a machine learning library. The synthetic dataset that we generated by using (1), it is not possible to directly apply a supervised learning algorithm on the dataset as it has no label. We standardize the features of the dataset by using (2). Then, we apply DBSCAN on the dataset with MinPts = 3, radius ε = 0.3 and choose Euclidean metric for the DBSCAN to perform its clustering.

Fig. 6 illustrates the result of applying DBSCAN on the dataset. Here, the black points represent the outlier. All the other colored points are members of different normal classes. In addition, Fig. 6 also shows that there are three normal classes (class 0, class 1 and class 2). After performing DBSCAN on the dataset, we get the label of the normal data as well as the outlier. Now, as we know the label of the dataset, we apply supervised learning algorithms i.e. SVM and Decision Tree. Fig. 7 depicts the result of applying SVM and Decision Tree on the labeled dataset.

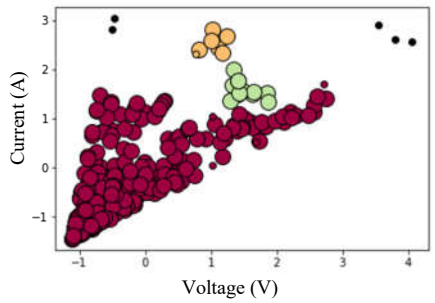
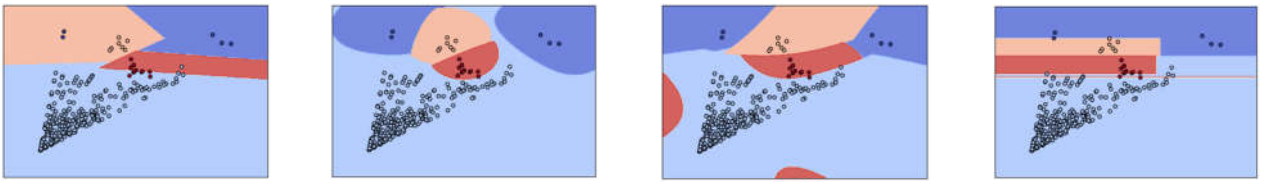


Fig. 6. Result of using DBSCAN (MinPts = 3, radius ε = 0.3, metric = 'euclidean') on the dataset.

Here, SVM and Decision Tree are used as classifier. For SVM, three different kernels i.e. linear, radial basis function ('rbf') and polynomial ('poly') are used. Different kernels are used to see which kernel can perform better. We used confusion matrix, accuracy and precision to evaluate the models. Table I and Table II represent the normalized confusion matrix for SVM and Decision Tree respectively. SVM with linear kernel predicts 50% of the outliers correctly and predict 50% as class 1.

All the other classifiers correctly predict the outlier. As our main concern is predicting the outlier, if a classifier predicts class 0 as class 1 or class 1 as class 2 it will cause no harm to the system.

The accuracy and precision are calculated using the (5) and (6) respectively. Fig. 8 illustrates the measure of accuracy and precision for training and testing phase of different classifiers. Table III contains the summary of accuracy and precision of different models. It shows that, in testing phase SVM with 'rbf' kernel has the highest accuracy and precision i.e. 98.61% and 99.60% respectively. Decision Tree classifier also performs well in the testing phase with the accuracy of 97.22% and precision of 94.59%.



(a) SVM (kernel = 'linear') (b) SVM (kernel = 'rbf', gamma = 0.5) (c) SVM (kernel = 'poly', degree = 3) (d) Decision Tree

Fig. 7. Result of applying different classifier on the labeled dataset (different colors represent different classes).

TABLE I. NORMALIZED CONFUSION MATRIX FOR SVM

	SVM (kernel = 'linear')				**SVM (kernel = 'rbf', gamma = 0.5)**				**SVM (kernel = 'poly', degree = 3)**			
	outlier	class 0	class 1	class 2	outlier	class 0	class 1	class 2	outlier	class 0	class 1	class 2
outlier	0.50	0.00	0.50	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
class 0	0.00	0.98	0.00	0.02	0.00	1.00	0.00	0.00	0.00	0.98	0.00	0.02
class 1	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00
class 2	0.00	0.80	0.00	0.20	0.00	0.20	0.00	0.80	0.00	0.20	0.00	0.80

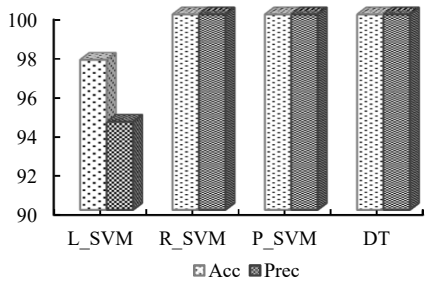
TABLE II. NORMALIZED CONFUSION MATRIX FOR DECISION TREE

	Decision Tree			
	outlier	class 0	class 1	class 2
outlier	1.00	0.00	0.00	0.00
class 0	0.00	0.98	0.00	0.02
class 1	0.00	0.00	1.00	0.00
class 2	0.00	0.00	0.20	0.80

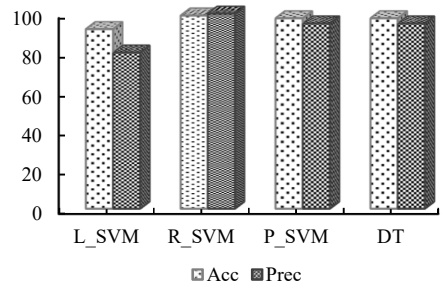
TABLE III. PERFORMANCE MEASURE INDICES OF OUTLIER DETECTION SYSTEM

Algorithm	**Kernel**	**Training Phase**		**Testing Phase**	
		Acc (%)	***Prec (%)***	***Acc (%)***	***Prec (%)***
SVM	Linear	97.68	94.51	91.67	79.71
	RBF	100	100	98.61	99.60
	Polynomial	100	100	97.22	94.59
Decision Tree	-	100	100	97.22	94.59

From Table III it is evident that if we compare the different models that have been employed on the system, SVM with 'rbf' kernel performs admirably in both training and testing phase with respect to other models.



(a) Training Phase



(b) Testing Phase

Fig. 8. Accuracy and precision for training and testing phase of different classifiers where L_SVM, R_SVM, P_SVM, DT represent SVM (kernel = 'linear'), SVM (kernel = 'rbf'), SVM (kernel = 'poly'), and Decision Tree respectively.

V. CONCLUSIONS

Protecting electronic devices from damage is very important as most of the devices are expensive. In this work, we focused on developing a system that not only detects outlier but also predicts whether a new current, voltage pair is an outlier or not. For outlier detection, a density-based clustering approach DBSCAN is used which can detect any arbitrary shaped cluster. Then, we develop models based on SVM and Decision Tree. We also provide a detailed performance measure for SVM with different kernels and Decision Tree. In the testing phase, SVM with 'rbf' kernel ends up with the highest accuracy of 98.61%. The performance shows that the proposed system will be quite helpful in protecting electronic devices.

REFERENCES

[1] V. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial intelligence review*, vol. 22, no. 2, pp. 85–126, 2004.

[2] H.-P. Kriegel, P. Kröger, and A. Zimek, "Outlier detection techniques," *Tutorial at KDD*, vol. 10, 2010.

[3] E. Eskin, A. Arnold, M. Prerau, L. Portnoy, and S. Stolfo, "A geometric framework for unsupervised anomaly detection," in *Applications of data mining in computer security*, ed: Springer, pp. 77-101, 2002.

[4] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et* al., "A density-based algorithm for discovering clusters in large spatial databases with noise." in *Kdd*, vol. 96, no. 34, pp. 226–231, 1996.

[5] K. Leung and C. Leckie, "Unsupervised anomaly detection in network intrusion detection using clusters," *in Proceedings of the Twenty-eighth Australasian conference on Computer Science-Volume 38*. Australian Computer Society, Inc., pp. 333–342, 2005.

[6] Z. A. Bakar, R. Mohemad, A. Ahmad, and M. M. Deris, "A comparative study for outlier detection techniques in data mining," in *Cybernetics and Intelligent Systems, 2006 IEEE Conference on.* IEEE, pp. 1–6, 2006.

[7] S. Ramaswamy, R. Rastogi, and K. Shim, "Efficient algorithms for mining outliers from large data sets," in *ACM Sigmod Record*, vol. 29, no. 2. ACM, pp. 427–438, 2000.

[8] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "Lof: identifying density-based local outliers," in *ACM sigmod record*, vol. 29, no. 2. ACM, pp. 93–104, 2000.

[9] A. Arning, R. Agrawal, and P. Raghavan, "A linear method for deviation detection in large databases." in *KDD*, vol. 1141, no. 50, pp. 972–981, 1996.

[10] M.-F. Jiang, S.-S. Tseng, and C.-M. Su, "Two-phase clustering process for outliers detection," *Pattern recognition letters*, vol. 22, no. 6-7, pp. 691–700, 2001.

[11] S. Hawkins, H. He, G. Williams, and R. Baxter, "Outlier detection using replicator neural networks," in *International Conference on Data Warehousing and Knowledge Discovery*. Springer, pp. 170–180, 2002.

[12] M. R. Chmielewski and J. W. Grzymala-Busse, "Global discretization of continuous attributes as preprocessing for machine learning," 1996.

[13] P. Berkhin, "A survey of clustering data mining techniques," in *Grouping multidimensional data*. Springer, pp. 25–71, 2006.

[14] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et* al., "A density-based algorithm for discovering clusters in large spatial databases with noise." in *Kdd*, vol. 96, no. 34, pp. 226–231, 1996.

[15] N. Schlitter, T. Falkowski *et* al., "Dengraph-ho: Density-based hierarchical community detection for explorative visual network analysis," in *Research and Development in Intelligent Systems XXVIII*. Springer, pp. 283–296, 2011.

[16] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.

[17] K.-j. Kim, "Financial time series forecasting using support vector machines," *Neurocomputing*, vol. 55, no. 1-2, pp. 307–319, 2003

[18] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE transactions on systems, man, and cybernetics*, vol. 21, no. 3, pp. 660–674, 1991.

Ulcer Detection in Image Converted from Video Footage of Wireless Capsule Endoscopy

Abdullah Al Mamun
Electrical and Electronic Engineering
Pabna University of Science and Technology
mamun130203@gmail.com

Md. Sohag Hossain
Electrical and Electronic Engineering
Pabna University of Science and Technology
msh22eee@gmail.com

***Abstract*—**Wireless capsule endoscopy (WCE) becoming a very useful device for detecting diseases in the human digestive tract. Using WCE Physicians diagnose the different abnormalities such as ulcer, bleeding, polyps, chronic diarrhea, small intestinal cancer/tumor and Crohn's disease in the digestive tract in invasive way. Researchers are working for the development and improvement of the performance of WCE using software to detect these diseases at a high rate of advancement automatically. This paper proposes an advance methodology for automatically discovering ulcers in the image converted from WCE video footage. The method applied for numerous frames of capsule endoscopic images with logistic regression classifier in HSV color space and found effectiveness for the physicians. We obtained the result of 87.70% accuracy and 94.00% sensitivity which provides the evidence of the accuracy of this automated computer aided system.

Keywords— CE images, Color threshold, GI tract, HSV Color space, Logistic regression classifier, Ulcer detection, Wireless capsule endoscopy.

I. INTRODUCTION

Throughout the world, a lot of people are affected by the diseases related with gastrointestinal tract specially stomach, small intestine, large intestine. In Bangladesh 27.8% people die in intestine infectious diseases [1]. In the United States, around 1.6 million Americans are now enduring from inflammatory bowel disease (IBD) and since 2011 increase to around 200,000. Each year about 70,000 new IBD are diagnosing. Among this 80,000 are children who are affected by Crohn's disease (CD) or ulcerative colitis (UC) [2]. IBD is now rising globally [3]. To cure the diseases of GI tract it is needed to detect as soon as possible. So now researchers are trying to upgrade the technique of detecting these types of disease. In this way there have used some medical diagnose system such as ultrasound, X-ray, CT scan, endoscopy, endoscopy biopsies etc. But these technologies have drawbacks and side effects for which researchers are working to discover such technique which will be the best technique and without any side effect. At last capsule endoscopy has been introduced which is painless during the procedure, easily swallowed and sedation free. At this technique a capsule shaped device containing a CMOS camera, a radio transmitter, a battery and light sources is swallowed by the patients with water like pill. This passes through the GI tract and taking caption and transmits via radio transmitter to the receiver [4]. The accruing video footage will be converted into images. These images will be processed by software for detecting the abnormalities. The abnormality for this paper is ulcer.

The WCE system consists of 3 components: (1) a capsule endoscope; (2) a sensing system with sensing pads or a sensing belt to piece together to the patient, a data recorder and a bat-terry cell; and (3) a personal computer with exclusive software for image inspection and explication. All 3 systems include handheld viewers that allow real time review of images during WCE examinations [4]. In WCE system, Different sensors are attached to the different position of abdomen and waist. A data recorder will be attached to the sensors. These sensors allow the capsule to wirelessly transmit images. Then the patient swallows the capsule with a glass of water. The capsule passes through the digestive tract and takes images. And then transmit wirelessly which is received by the receiver and stored in data recorder. The internal and experimental overview of the wireless capsule endoscopy are shown in the figure 1. Capsule endoscopy device, Internal construction of CE device, Image receiver, Attachment of sensors in human body at different positions and Real time viewer are respectively shown in the figure 1(b). By using real time viewer, the condition of imaging of the digestive tract is observed. The process continues approximately eight hours until the battery life of the capsule. There recorded approximately 50,000 images. The capsule is disposable and passes naturally within 24 – 72 hours.



Fig. 1. (a) Internal Structure of CE and (b) Imaging methods of CE

A. Related Work

Firstly, software was designed for detecting bleeding region whose sensitivity and specificity were only 21.5% and 41.8% [5]. Then a feature extractor is designed for detecting the abnormalities such as bleeding, ulcer and polyp named MPEG-7 visual descriptor [6]. According to [7] bleeding and ulcer has detected from of capsule endoscopy images which signify chromaticity moment using HSI (Hue-Saturation-Intensity) color space. In [8] statistical feature-based bleeding detection from capsule endoscopy video. In [9] bleeding detection is obtained by histogram from indexed images in capsule endoscopy videos. Curve let based local binary pattern in textural features for detecting ulcer regions from normal regions [10]. In [11] texture and color-based image segmentation and pathology detection in capsule endoscopy videos have proposed. In [12] A Color Coherence Vector (CCV) based feature is introduced for

identifying ulcer images of capsule endoscopy videos. In [4] histogram bin method for ulcer detection in capsule endoscopy was proposed.

In this paper a method is proposed using HSV color space with logistic regression classifier. At first convert the video footage into image frame. After that, enhance the image which highlights the key information. Then, color threshold is applied at HSV color space to get only the ulcer portion. After this, the images are converted into binary image, adjust intensity and sharpen the image and filter is used for removing the non-ulcer portion from the ulcer selection. Then, several features are extracted from the ulcer portion. Using these features, classifier is applied and tested on publicly available CE

II. OUR PROPOSED METHOD

The method we have used for ulcer detection is very much easy and ensured the high performance of the procedure. By using color threshold on the CE images, normal portion has been removed and separated the abnormal portion from the ideal portion which is regarded as ulcer for this paper. Normally, ulcer portion occupies small area on the ulcer image frame. Then the algorithm is only applied on that particular portion of the image frame for which the calculation is not complicated. The working chain of our proposed method is shown in Figure 2.

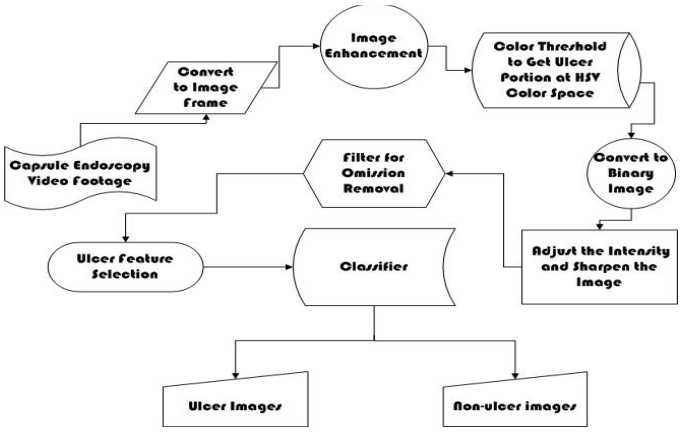


Fig. 2. Flow diagram of the proposed method

A. Video Footage to Image Frame

The initial process of this research work is to convert the acquiring video footage from wireless capsule endoscopy into image frame according to frame rate specified by the capsule specification for further processing such as image enhancement, image segmentation, feature extraction and classification. For this particular experiment, we have used capsule endoscopy of PillCam®SB whose frame rate per second is 2 and battery life is 8 hours which stored a huge amount of data for analyzing. From this huge data, we have used 220 color images from 8 capsule endoscopy videos which are available in [13].

B. Image Enhancement

Image enhancement is the method of adjusting digital images in order to make more suitable images for display or further image analysis. It removes noise, brighten an image, makes easier to identify, and sharpen as well. This is the initial step for preprocessing. The main purpose of the step is to remove noise and secondary information, enhance the contrast and high lighting the key information of an image. It also improves the quality of supervision [14]. For the image enhancement process, histogram equalization method has been employed in this paper.

C. HSV Color Threshold

Color threshold has been applied for removing the unwanted portion the images that fall within a specified color range and detect the objects of consistent color values. For this paper, the unwanted information of image is non-ulcer portion and detected object or information is ulcer. In this particular experiment, we applied color threshold in HSV color space. In the case of HSV color space, it contains the value the values of Hue, Saturation and Brightness at a particular range which only select approximately the ulcer portion from the CE images. By selecting these values **of** three channels in HSV color space we have obtained the ulcer pixels that have shown in Figure 3. This figure also displays the Hue, Saturation and Brightness value histograms which have been selected.



Fig. 3. Color threshold in HSV color space

D. Convert to Binary Images

After acquiring the color threshold images, we converted those into binary images which regard the pre-preparation for further calculation or feature extraction. Color threshold image converted to binary image has been shown in Figure 4.

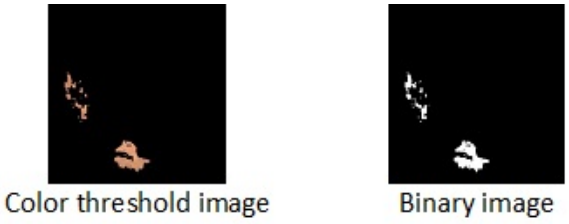


Fig. 4. Color threshold image converted to binary image

E. Filter for Omission Removal

The N-D filtering of multidimensional images has been applied that compares with a required value and truncates elements that exceed the range of the given prototype and rounds fractional values. This filter eliminates the non-ulcer portion from the pure ulcer portion which is the main requirement.

F. Convert to Binary Images

In this method, ulcer features of the capsule endoscopy images have been calculated by using mean of hue, saturation, brightness value and standard deviation of the pure ulcer portion. The image frames which contain any ulcer portion will have a particular numerical value. But image frames without any ulcer portion will contain zero numerical value. Some sample has shown in Table I.

TABLE I
Sample of ulcer and non-ulcer image features

	Mean Hue	Mean Saturation	Mean Value	Mean Std. Deviation
Ulcer	0.000308133512948861	0.00301234398048923	0.00433409447763483	0.000803345356704952
Ulcer	0.000753306172406202	0.00663013986873620	0.00979742910347761	0.001444624 87951517
Ulcer	0.000974931012444869	0.00869841504836054	0.01213699 04162840	0.002144505 84794773
Non-ulcer	0	0	0	0
Non-ulcer	0	0	0	0
Non-ulcer	0	0	0	0

G. Logistic Regression Classifier

For classifying ulcer and non-ulcer images logistic regression classifier has been used. It is a **classification** technique for different pattern recognition from the field of statistics. Logistics is the directory of the flow of sample between the point of origin and the point of haul. In this method of classification, the extracting four types of feature are measurement variables. The values of the nominal variable are "Ulcer" and "Non-ulcer" for this paper. To obtain the performance results 5-fold cross validation technique has been taken. After classification of these data ulcer and non-ulcer images will be separated.

III. Result and Discussion

The performance of logistic regression classifier in CE images is measured by accuracy, specificity and sensitivity which are extensively used to evaluate performances of classification. Here are definitions of them: In [10] definitions have been given:

$$\text{Accuracy} = \frac{Quantity\ of\ actula\ prediction}{Total\ sample\ data} \quad (1)$$

$$\text{Sensitivity} = \frac{Quantity\ of\ actula\ successful\ positive\ prediction}{Number\ of\ actula\ positives} \quad (2)$$

$$\text{Specificity} = \frac{Quantity\ of\ actula\ negative\ prediction}{Number\ of\ actula\ negatives} \quad (3)$$

Number of correct predictions is the sum of ulcer and non-ulcer images prediction is correct. Number of correct positive predictions is the number of ulcer images prediction is correct. Number of correct negative predictions is the number of non-ulcer images prediction is correct. The logistic regression classifier results obtained from the testing phase are summarized in Table II.

The Table III shows the correlation between different existing method. In addition, the table shows the evidence of higher efficiency of proposed method than the mentioned existing methods like CWC in which chromaticity is created [7], LBP which uses curvelet transformation [10], DCT-LAC in which lacunarity index is calculate to collect texture information [15] and histogram in which color plane is used [4].

TABLE II
Results obtained using logistic regression classifier for ulcer detection

Parameters	Values
Accuracy	87.70 ± 0.40%
Sensitivity	96.50±0.50%
Specificity	76.50±0.50%

TABLE III
Ulcer detection performance comparison with other existing method

Methods	Accuracy	Sensitivity	Specificity
CWC [7]	74.35%	78.46%	70.24%
LBP [10]	83.92%	86.28%	81.56%
DCT-LAC [15]	86.54%	84.51%	88.56%
Histogram bin [4]	87.27%	88.64%	85.75%
Proposed method	**87.70%**	**94.00%**	**75.00%**

IV. Conclusion

This paper proposed a method using color threshold of HSV color masking for selecting the ulcer portion in CE images that achieved from the video footage of WCE. After selecting the ulcer portion in CE images features are extracted only from the ulcer portion. Analysis of ulcer features values has been successfully prepared for providing the logistic regression classifier that offers the best separation of ulcer and non-ulcer images. The method has applied for numerous frames of capsule endoscopic images and found effectiveness. The result has obtained 87.70% accuracy and 94.00% sensitivity which proof that the proposed method is more accurate.

The results especially accuracy and specificity encourage that the method can be enhanced in the future by adding more image data base, statistical and texture features descriptor of other color spaces too.

Acknowledgement
We would like to thank Topu Biswas for his special guidance and support.

References

[1] (27 april, 2018). *Death in Different Diseases*. Available: www.healthdata.org/bangladesh

[2] G. G. Kaplan and S. C. Ng, "Understanding and preventing the global increase of inflammatory bowel disease," *Gastroenterology,* vol. 152, pp. 313-321. e2, 2017.

[3] G. G. Kaplan and T. Jess, "The changing landscape of inflammatory bowel disease: East meets West," *Gastroenterology,* vol. 150, pp. 24-26, 2016.

[4] T. Ghosh, A. Das, and R. Sayed, "Automatic small intestinal ulcer detection in capsule endoscopy images," *International Journal of Scientific and Engineering Research,* vol. 7, pp. 737-741, 2016.

[5] R. Francis, "Sensitivity and specificity of the red blood identification (RBIS) in video capsule endoscopy," in *3Rd, Int. Conf. on Capsule Endoscopy*, 2004.

[6] M. T. Coimbra and J. S. Cunha, "MPEG-7 visual descriptors—contributions for automated feature extraction in capsule endoscopy," *IEEE transactions on circuits and systems for video technology,* vol. 16, pp. 628-637, 2006.

[7] B. Li and M. Q. Meng, "Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments," *Comput Biol Med,* vol. 39, pp. 141-7, Feb 2009.

[8] T. Ghosh, S. K. Bashar, M. S. Alam, K. Wahid, and S. A. Fattah, "A statistical feature based novel method to detect bleeding in wireless capsule endoscopy images," in *Informatics, Electronics & Vision (ICIEV), 2014 International Conference on*, 2014, pp. 1-4.

[9] T. Ghosh, S. Fattah, C. Shahnaz, and K. Wahid, "An automatic bleeding detection scheme in wireless capsule endoscopy based on histogram of an RGB-indexed image," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 2014, pp. 4683-4686.

[10] B. Li and M. Q.-H. Meng, "Texture analysis for ulcer detection in capsule endoscopy images," *Image and Vision computing,* vol. 27, pp. 1336-1342, 2009.

[11] P. Szczypiński, A. Klepaczko, M. Pazurek, and P. Daniel, "Texture and color based image segmentation and pathology detection in capsule endoscopy videos," *Computer methods and programs in biomedicine,* vol. 113, pp. 396-411, 2014.

[12] J.-Y. Yeh, T.-H. Wu, and W.-J. Tsai, "Bleeding and ulcer detection using wireless capsule endoscopy images," *Journal of Software Engineering and Applications,* vol. 7, p. 422, 2014.

[13] (2016). *The capsule endoscopy website*. Available: https://www.capsuleendoscopy.org

[14] S. Suman, F. A. Hussin, W. Nicolas, and A. S. Malik, "Ulcer Detection and Classification of Wireless Capsule Endoscopy Images Using RGB Masking," *Advanced Science Letters,* vol. 22, pp. 2764-2768, 2016.

[15] A. Eid, V. S. Charisis, L. J. Hadjileontiadis, and G. D. Sergiadis, "A curvelet-based lacunarity approach for ulcer detection from wireless capsule endoscopy images," in *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*, 2013, pp. 27

Feature-Based Mobile Phone Rating Using Sentiment Analysis and Machine Learning Approaches

Abdullahil Kafi
Department of Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
kafiabdullahil96@gmail.com

M. Shaikh Ashikul Alam
Department of Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
sheikhaashik@gmail.com

Sayeed Bin Hossain
Department of Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
sb.hossain.h@gmail.com

Siam Bin Awal
Department of Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
siam.awal27@gmail.com

Hossain Arif
Department of Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
hossain.arif@bracu.ac.bd

***Abstract*— This paper proposes a model of sentiment analysis of various features of different companies' mobile phones and their overall rating. Before buying a phone, customers usually look for reviews to decide which phone to buy. The model proposed in this paper provides an optimal solution for the customer for making this decision more efficiently. In this model, each feature of a mobile phone is rated based on public opinion and an overall rating for each phone is provided. Amazon is one of the largest Internet retailers, which makes way for most public reviews on their products. These reviews are collected as a form of an open source platform and used as the dataset in this model. The gathered data is pre-processed and then separated into two different sets – Training Set and Testing Set which are used to train and test the supervised machine learning algorithms for classification. Fifteen most common features of the mobile phones based on public reviews are selected from the training data set and used as the feature set in this model. Different algorithms which include Naïve Bayes, Support Vector Machine, Logistic Regression, and Stochastic Gradient Descent algorithms are used in this model and the comparison of their performance is shown. This model provides a rating of each feature and an average rating of the mobile phone based on sentiment polarity. Thus, this research work can assist potential customers to choose the best product based on the opinion of the other users.**

Keywords—Sentiment Analysis; Machine Learning; Natural Language Processing

I. INTRODUCTION

Mobile phones have been one of the most essential components for us during the last decade. From Brick Phone to iPhone, technological advancement is seen in the sector of telecommunication. At present, mobile phones have different features according to the customer's choice and need. Among thousands of companies, offering their technology to the customers, choosing the appropriate phone for an individual has become an issue. Thus, the decision-making process becomes more challenging in recent times. In this research, a model is proposed which generates a rating of particular features and overall features, based on public reviews of different features or specification of a mobile phone. The ratings will let consumers compare their need and satisfaction and thus choose a suitable mobile phone. There are many websites, YouTube channels and also individuals on Social Media who reviews mobile phones. Professional reviewers are appointed by different companies to act as a marketing policy of the company.

However, very few reviewers are authentic and independent of their opinions. Such reviewing criteria benefits the mobile manufacturer companies to convince the customer about their product. This has major drawbacks from the consumer's point of view. However, lacking's in the software or hardware part of the product such as power consumption, battery lifetime, operating systems or operation time and many information get skipped for the sake of the company's marketing strategy. Thus, these reviews are not helpful in terms of customers point of view. Secondly, a mobile phone's technical features such as Chipset, GPU, CPU, OS's contribution to a mobile phone system are not often understood to an average user. Thus, a potential buyer cannot get suitable information that he desires about a mobile phone from the reviewers. Thirdly, the reviewer's point of explanation of a device does not match the expectation of the customer. A customer's prime concern may be about the camera while a reviewer will be emphasizing on other technical features of the device. This indicates the absence of a platform where customers can get the rating of their desired features. Additionally, a professional review never reflects the actual state of a device. Generally, reviewers don't review the device after using it for a long time, rather they do their reviews based on a test-use which doesn't always contemplate the actual experience of that particular device. Therefore, the customer's opinion about a mobile phone and its features are more reliable and authentic. This is why comparing ratings of public opinions about mobile phone to buy one is better rather than relying on professional reviews. Author's in this research also focused on public reviews rather than the professional reviewer's opinion.

The rest of the paper is organized as follows. In Section II, background study and related works have been presented. In Section III, the working principle has been explained. Furthermore, in Section IV, the result of the work has been

discussed. And finally, in Section V, the conclusion of the paper with future work has been presented.

II. Background Study

Sentiment analysis is an emerging area of research in natural language processing (NLP). In recent years, this field has gained much interest as the social context changed a lot. The internet-based economy made sentiment analysis more essential. As the number of consumers' review increases, the need for analyzing these reviews automatically is also increasing which can help today's business to significantly improve as digital marketing has become a genuine marketing concept. As a result, not only researchers but also entrepreneurs are becoming interested in sentiment analysis [1].

Sentiment analysis is a machine learning method which can be used to classify and analyze peoples' sentiments, emotions, opinion, etc. Review of online product which is expressed in terms of text, star rating, and thumbs up and down can be analyzed using sentiment analysis which can significantly help to improve certain products or services [2]. With the aid of negative prefixes, words like adjectives and adverbs are able to carry the opposite sentiment. So, in this research, the authors used the negation phrase identification algorithm to find such words and analyzed the reviews.

In some researches, twitter data is used for sentiment analysis. Similarly, in the research conducted by Rane and Kumar, the authors have classified twitter data for US airline service analysis where firstly the tweets are pre-processed to obtain a vector using deep learning concept (Doc2vec) for phrase-level analysis. And then using 7 different classification technique, the analysis is carried out. Based on the accuracy of the result, a comparison among different classification technique is given in this research paper [3]. Opinions can also be extracted based on the product's features using several techniques. In this research paper, the authors suggested a system that automatically expands opinion lexicon and based on dependency relations, the product's features are extracted. And to find this dependency relation between features and opinions, StrandfordCoreNLP dependency parser is used and then an algorithm based on double propagation is used to extract feature and opinions [4]. With the aid of Vader lexicon, polarities of extracted opinions are defined.

The author Xing Fang and Justin Zhan in their research work describe that when sentiment sentences are extracted, it contains at least one positive or negative word. Tokenization is done to separate the sentences into words and Parts-of-speech (POS) tagging is done. From the original dataset, the tokens and their scores are extracted which are known as features. So, to classify these features, they transformed into vectors [5]. Furthermore, in another paper, the authors used unsupervised learning on large twitter corpora to obtain a word embedding which uses latent contextual semantic relationships and co-occurrence statistical characteristics between words in tweets [6]. Together with n-grams features and word sentiment polarity score features, these words embedding form a sentiment feature set of tweets and this feature set is combined into a deep convolution neural network for training and predicting sentiment classification labels. This model shows a significant amount of accuracy in their result.

However, authors Roshna Rafeek and R. Remya, in their research work stated that context-dependent opinion words are challenging for sentiment analysis because of their polarity changes according to the context in which they are used. So, to overcome this challenge, they have proposed a system that uses (aspect, opinion) pairs and logistic regression (LR) model. In this model, to acquire (aspect, opinion) pairs, syntactic rules are used and for classifying the sentiment words into positive class and negative class, the LR model is used [7]. In this research paper, the authors proposed an approach for sentiment analysis using Python and its library NLTK (Natural Language Toolkit). Python, which is simple yet powerful, high level, interpreted and dynamic programming language, is very useful for processing natural language using its own library NLTK that helps to build programs and classify data. Data can be trained and different numerous classifiers can be tested using NLTK. Additionally, using NLTK, results, and patterns can be graphically represented.

In another research, the authors described different supervised machine learning approaches for sentiment analysis [8]. Various algorithms are used and their advantages and disadvantages are stated. There are two broad categories which are Lexicon based approach and machine learning approach. In their research work, they have used three algorithms. These algorithms are naïve Bayes, support vector machine and maximum entropy. According to their result, naïve Bayes algorithm provided better accuracy than the other two [9].

III. Working Principle

Sentiment analysis is a very useful tool in the modern technological era where people are frequently sharing their thoughts in microblogging social sites. Analyzing sentiments help human beings to take a decision in case of socio-economic and business aspects. In this research, we used natural language processing along with machine learning algorithms to find out the sentiment of a review. Furthermore, they also intended to find the best accuracy of the sentiment. We used different machine learning algorithms, such as Naïve Bayes Classifier, Linear Regression, Support Vector Classifier, Stochastic Gradient Descent Classifier and Random forest for analyzing sentiments. A block diagram is given below to show the proposed model of sentiment analysis,

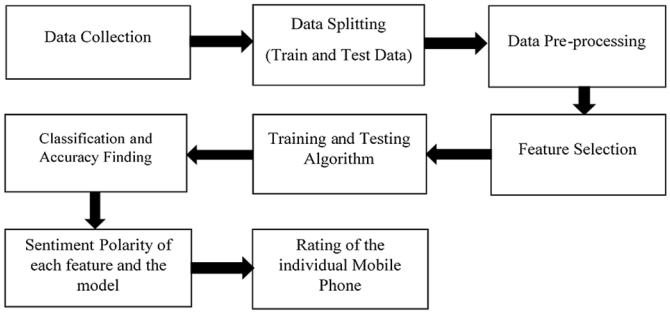


Fig. 1. Block diagram of the proposed model.

A. Data

Dataset is the most important factor in any kind of machine learning and sentiment analysis process. In the research of sentiment analysis, the most technical and one of the toughest tasks is to find out the proper dataset applicable to the model. As the aim is to find out the proper review and rating of a mobile set, we generalized searching criteria for datasets with proper reviews along with mobile set company or model name. We used Amazon's unlocked mobile phone review's dataset which was collected from an online open source platform named Kaggle. There were four hundred thousand data in the dataset. The dataset had six columns, which were 'product name', 'company name', 'review', 'rating', 'rating votes'. Three of the columns consists of textual data and the rest is numerical data. 'Product Name' column represented the model of the mobile phone whereas the 'Company Name' had the name of the company. The review column consists of reviews provided by users. These reviews are mixtures of different sentence type. Rating column is the rating of each review given by users at Amazon.

B. Text-Preprocessing

Data preprocessing is the process of cleaning the data to an extent where the machine learning algorithm will understand the content of the data. Overall the data pre-processing deals with finding the inaccurate, irrelevant and incomplete data and removing or minimizing the data. Pre-processing is a very important step in terms of textual data. This pre-processing task is done using NLTK (Natural Language Tool Kit) which is an open source language tool for textual data mining. There are several steps of text pre-processing, which are,

i. Tokenization
ii. Normalization
iii. Noise removal

Tokenization is a step which splits longer strings of text into smaller pieces or tokens. Larger chunks of text can be tokenized into sentences and sentences can be tokenized into words, etc. Further processing is generally performed after a piece of text has been appropriately tokenized. Tokenization is also referred to as text segmentation or lexical analysis. Sometimes segmentation is used to refer to the breakdown of a large chunk of text into pieces larger than words (e.g. paragraphs or sentences), while tokenization is reserved for the breakdown process which results exclusively in words.

Before further processing, the text needs to be normalized. Normalization generally refers to a series of related tasks meant to put all text on a level playing field: converting all text to the same case, upper or lower case, removing punctuation, converting numbers to their word equivalents, and so on. Normalization puts all words on equal footing and allows processing to proceed uniformly. Normalizing text can mean performing a number of tasks, but for our model, we approached normalization in three distinct steps, Stemming, Lemmatization, and POS tagging.

Stemming is the process of eliminating affixes (suffixed, prefixes, infixes, and circumfixes) from a word in order to obtain a word stem. Lemmatization is related to stemming, differing in that lemmatization is able to capture canonical forms based on a word's lemma. For example, stemming the word "cacti", "geese", "rocks" would fail to return its citation form (another word for lemma); however, lemmatization would result, "cactus", "goose", "rock". POS tagging is also a part of text pre-processing and used for detecting the parts of speech of the words in a sentence. For selecting the adjective of a sentence, we have to use this process in sentiment analysis.

C. Feature selection of a mobile phone

The model concentrates on finding out the most frequently reviewed specifications of the mobile phone, so the selection process plays a key role in the working methodology. Mobile phones have multiple features like Camera, RAM, Storage, Battery, Display, Operating system, etc. Firstly, a list of the specifications was created. As a single feature may have several names, we created NLTK synonyms for the features or specifications and stored all the strings in a single array. The features which appear mostly in the dataset are taken into account and matched with the feature set we created. We selected in total 15 features, mostly consisting of nouns which define the specification of any mobile phone.

D. Classification

We have used multiple algorithms for training the dataset. The amount of train data was increased gradually to see the accuracy of the algorithms. Classification is the task of choosing the correct class label for a given input. In the basic classification tasks, each input is considered in isolation from all other input, and the set of the labels is defined in advance. Mentioned earlier, we are using the built-in libraries to do all the classification process. To be specific, the following algorithms are used in the classification process,

- Naïve Bayes (NB), Multinomial Naïve Bayes (MNB), Bernoulli Naïve Bayes (BNB)
- Support Vector Classifier, NuSVC, and Linear SVC
- Logistic Regression and Linear Regression
- Stochastic Gradient Descent (SGD) Classifier
- Random Forest

Machine learning algorithms are used to classify the sentences based on their polarity. However, the purpose of the research is to find out the best algorithm for sentiment analysis based on the accuracy of classification. The details of the algorithms are given below,

1) Naïve Bayes Classifier: A Naive Bayes classifier is an algorithm that uses Bayes' theorem to classify data. Naive Bayes classifiers assume strong, or naive, independence between attributes of data points. The classifier uses probability theory to classify. The main key insight of Bayes' theorem is that the probability of an event can be adjusted as new data is introduced. A Naïve Bayes classifier is called naïve because it assumes that all attributes of a data point under consideration are independent of each other [10]. A Naïve Bayes model is easy to build and it has no complicated iterative parameters estimation which makes it particularly useful for very large datasets.

2) Linear Regression: In machine learning and statistics linear regression is very useful for classifying a large set of datasets. The model attempts to model the relationship between two variables by fitting a linear equation to observed data. One variable is considered as a dependent variable whilst the other variable is considered as an explanatory variable. Linear regression has the equation,

$$Y = a + bX \tag{1}$$

Here, X is the explanatory variable and Y is the dependent variable, the slope of the line is 'b' and 'a' intercepts. We determined the relationship between variables of interest and attempted to fit the linear model for the observed data.

3) Support Vector Machine: Support Vector Machine (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges. However, it is mostly used for text classification [11]. In the algorithm, each data will be plot as a point in n-dimensional space with the value of each feature. n is the number of features. Then classification is done to find the hyperplane to differentiate two classes very well. Normally this model draws lines to separate the groups according to patterns.

4) SGD Classifier: Stochastic Gradient Descent also known as SGD classifier is a very efficient approach to discriminative learning of linear classifiers under convex loss functions such as (linear) Support Vector Machines and Logistic Regression. For large scale learning, this algorithm is very efficient. Moreover, this algorithm is very successful in text classification as well as natural language processing.

5) Random Forest Classifier: Random forests, also known as random decision forests, are a popular ensemble method that can be used to build predictive models for both classification and regression problems. It is a flexible, easy to use machine learning algorithm that produces, even without hyper-parameter tuning, a great result most of the time. It is also one of the most used algorithms because it's simplicity and the fact that it can be used for both classification and regression tasks [13]. As its name, it creates a forest and makes it somehow random. The 'forest' it builds, is an ensemble of Decision Trees, most of the time trained with the "bagging" method. The general idea of the bagging method is that a combination of learning models increases the overall result. Simply random forests build multiple decision trees and merge them to get more accurate and stable predictions. There are three main choices to be made when constructing a random tree. These are (1) the method for splitting the leaf's, (2) the type of predictor to use in each leaf, and (3) the method for injecting randomness into the trees.

The classification algorithms are to classify the dataset which consists of features and finds out the accuracy of the algorithms. The classification is done in a various number of sets starting from 1000. We have to find out the accuracy of the different algorithm by gradually increasing the number of sets. However, to get the best accuracy it is vital to increase the number of sets. For this research, classification started from 1000 up to 12000 because the larger the set number the more computer power is needed for classification.

IV. Result and Discussion

In the proposed system, the Python programming language is used for implementation. To measure the accuracy of the algorithm we used an evaluation equation. Accuracy is the performance evaluation parameter and it is calculated by the number of correctly selected positive and negative words divide by the total number of words present in the corpus. The formula is given below,

$$Accuracy = \frac{\sum TruePositibve + \sum TrueNegative}{\sum TotalNumberofWords} \tag{2}$$

Here, true positive is the number of tweets recognized as positive and true negative is the number of tweets recognized as negative.

As high computing power is needed for finding the best accuracy by classifying large data sets, we classified a selected set for classification. Figure 2 demonstrates a graphical representation of the classifiers and the accuracy obtained by using the number of sets. Matplotlib, a python plotting library was used for plotting the data.

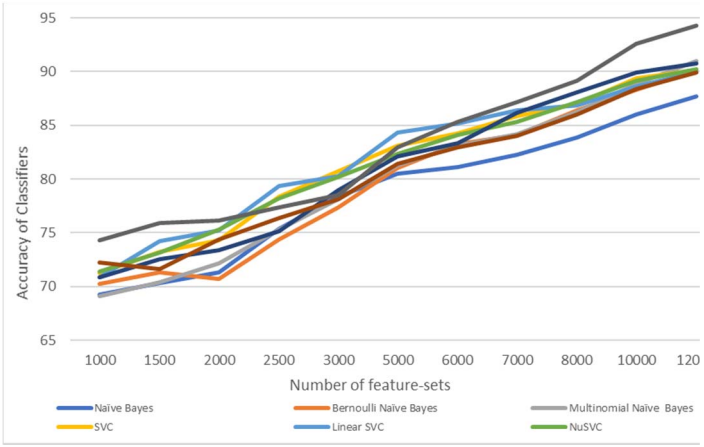


Fig. 2. Machine Learning Algorithms Accuracy

It clearly indicated the high accuracy whenever the number of sets is increased. The line chart shows the superiority of the Random forest algorithm as it shows the highest accuracy of around 92% with 12000 sets.

In this research data selection, data pre-processing, and specification or feature selection of mobile phones, training algorithms, classifying algorithms are done. Python and Scikit-learn which is a machine learning library of python and NLTK which is a language processing tool widely used for the sentiment analysis and text polarity measurement are used for the research. However, we have applied all the classifiers individually to get the best accuracy but for the best outcome from sentiment analysis, it is recommended by many researchers that a combination of multiple supervised algorithms will be better. This combined classifier provides the best possible result for the pre-processed data. As mentioned earlier, we have separated all the sentences of the review text according to the specification or feature of the mobile. As an example, if there is a review mentioning multiple features of the phones followed by their review, the sentence is picked and applied sentiment analysis on that sentence. In figure 3 selection of sentence, the process is shown in a pictorial view.

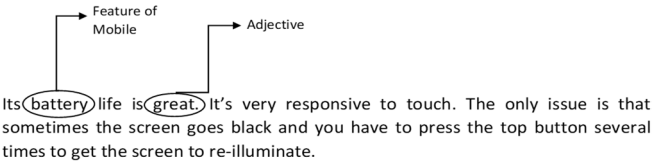


Fig. 3. Selection of Adjective and Specification (feature)

After the selection of text, we specifically applied sentiment analysis with NLTK in the adjectives or the words which can determine polarity. As an example, "Nokia is the brand that can be used in telephone, good price, excellent team, was a gift to my daughter and she liked it, easy to use and very affordable price, I recommend it." this review has one big sentence with a lot of features or specification and their classification, the result shows polarity positive and the confidence level 98.56%.

The last step of the research is to provide a rating of the mobile phone according to the review and polarity. In the previous step, we found out the polarity of each feature and the score of the polarity which starts from 0 to 100. However, there is a threshold value in sentiment score, which determines the accuracy of the sentiment analysis. We set the threshold value to 30 so whenever the score is less than 30 the sentiment analysis will indicate it as a bad sentiment.

To rate the mobile we used a special parameter. As in sentiment analysis, the polarity of each word along with the polarity percentage is shown, we used the polarity percentage and converted that numerical value into the 0-5 range. This gives an individual rating of each feature. However, to find out the rating of the product or mobile phone, the average rating of each feature is calculated which shows the rating of the desired mobile phone.

$$Average\ \ rating = \sum_i^n \frac{ratingofeachfeature}{n} \qquad (3)$$

In figure 4 the result of the example is shown, which consists of the final result of sentiment analysis.

```
C:\Users\LENOVO\Thesis\Scripts\python.exe C:/Users/LENOVO/PycharmProjects/Thesis/New.py
Review Sentence
Nokia is the brand that we can rely on, good price, excellent durability.
 I gifted my daughter and sheliked it, very affordable and long battery life
Sentiment Analysis -->
Positive
Confidence Level: 98.99


Specification/Feature based result -->

Noun --> Adjective --> Polarity

Price --> Good --> Positive
Durability --> Excellent --> Positive
Battery --> Longlife --> Positive

Process finished with exit code 0
```

Fig. 4. The final result of Sentiment analysis

The framework applied in this research seems very convenient in terms of mobile selection for customers. This methodology will help the customers to evaluate any mobile phone based on public rating. Moreover, this model also determines the rating of a mobile phone by calculating the mean value of all ratings of that particular device. However, from this research, we compiled some comparisons between different devices. Figure 5 represents the ratings of different mobile devices.



Fig. 5. Rating comparison between different models

V. CONCLUSION

In today's world, uses of sentiment analysis can play a significant role to understand people's opinion automatically and use that to improve certain product or service efficiently. This can be a field of interest for both researchers and entrepreneurs. In this research, analyzing the sentiment of the customer's review of mobile phones which is collected from Amazon and rate them according to their features. Naïve Bayes, Support Vector Machine, Logistic Regression, Stochastic Descent algorithm are used in this model and accuracy of their performance are compared which showed that SVC and Random forest perform better with greater accuracy. Using multiple algorithms helped to understand which algorithm is more suitable for this system. Ratings of the product are given based on the average polarity obtained which can help the customer to know about certain products or services and take decisions accordingly. From this research work, it is found that, the cleaner the data, the better the accuracy of the result. In the future, this model can be implemented as a web application so that while a customer is willing to buy a certain product, they can easily make a decision just by looking at the ratings.

We want to improve the algorithms that are used in this research. In the future, the goal is to apply unsupervised machine learning in the model, such as the neural network. We are looking forward to making the decision-making process easier when it comes to selecting mobile phones for users. We believe that this model will be user-friendly and accurate enough for users to use this on a regular basis for judging mobile phones and save their time from looking at different sites to find good reviews.

VI. REFERENCES

[1] K. Wójcik and J. Tuchowski, "Feature based sentiment analysis", presented at the 3rd International Scientific Conference on Contemporar Issues in Economics, Business and Management EBM, 2014.

[2] R. Safrin, K.R.Sharmila, T.S.S. Subangi and E.A. Vimal, "Sentiment Analysis on Online Product Review", International Reseach Journal of Engineering and Technology (IRJET), vol.04, issue.04. 2017.

[3] A. Rane and Dr.A.Kumar, "Sentiment Classification System of Twitter Data for US Airline Service Analysis", presented in 42nd IEEE International Conference on Computer Software & Application, 2018.

[4] M.S.Wai and S.S.Aung, "Simultaneous Opinion Lexicon Expansion and Product Feature Extraction", Third Seminar, University of Computer Studies, Mandalay, 10.13140/RG.2.2.28250.44489, 2018.

[5] X. Fang and J. Zhan, "Sentiment Analysis Using Product Review Data", Journal of Big Data, 2. 10.1186/s40537-015-0015-2, 2015.

[6] Z. Jianqiang and G. Xiaolin, "Deep Convolution Neural Networks for Twitter Sentiment Analysis", IEEE, 2017.

[7] R.Rafeek and R. Remya, "Detecting contextual word polarity using aspect based sentiment analysis and logistic regression", 2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM),2017.

[8] B.Wagh,Prof.J.V.Shinde and Prof.P.A.Kale, "A Twitter Sentiment Analysis Using NLTK and Machine Learning Techniques",International Journal of Emerging Research in Management &Technology ISSN: 2278-9359 (Volume-6, Issue-12),2017.

[9] S.S.Kamble and Prof.A.R.Itkikar, "Study of supervised machine learning approaches for sentiment analysis", International Research Journal of Engineering and Technology (IRJET) (Volume: 05 Issue: 04),2018.

[10] S.B.Kim, K.S.Han and H.C.Rim, "Some Effective Techniques for Naive Bayes Text Classification," IEEE Transactions on Knowledge and Data Engineering (Volume: 18, Issue: 11, Nov. 2006), 2006.

[11] R. Varghese and M Jayasree, "Aspect based Sentiment Analysis using support vector machine classifier," 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2013.

[12] Y.E.Cakra and B.D.Trisedya, "Stock price prediction using linear regression based on sentiment analysis," 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2015.

[13] R.Rafeek and R. Remya, "Detecting contextual word polarity using aspect based sentiment analysis and logistic regression," 2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM),2017.

[14] F.Kabir, S.Siddique, M.R.A.Kotwal and M.N.Huda, "Bangla text document categorization using Stochastic Gradient Descent (SGD) classifier," 2015 International Conference on Cognitive Computing and Information Processing(CCIP),2015.

Unsupervised Content-Based Image Retrieval Technique Using Global and Local Features

Md. Farhan Sadique
Computer Science and Engineering Discipline
Khulna University
Khulna, Bangladesh
farhan1338@cseku.ac.bd

Bishajit Kumar Biswas
Computer Science and Engineering Discipline
Khulna University
Khulna, Bangladesh
bkbkucse@gmail.com

S.M. Rafizul Haque
Computer Science and Engineering Discipline
Khulna University
Khulna, Bangladesh
rafizul@cse.ku.ac.bd

***Abstract*—Content-based image retrieval is an application of retrieving similar images of a query image. In this paper, global and local features are used to describe the content of an image. Color, texture and shape features are used as global features and SURF, FAST and BRISK are used as local features to retrieve similar images from the database. Different similarity checking methods are also tested to find the best approach. Combined approach using global features (obtained by using modified gray-level co-occurrence matrix) and local features (obtained by using SURF and color moments) produce better accuracy than some existing methods. Here, local features are obtained around the blob points detected by SURF detector. SURF descriptor and color moments which are calculated within a region of size 6*s* (*s* is the scale of the image at which the blob point is detected) around the detected blob points are used for extracting local features. Color moments are used because SURF only works on grayscale images. Thus, local color information is achieved. This combination of global and local features result in better performance in term of accuracy compared to some current methods.**

***Index Terms*—image matching, image processing, unsupervised, feature extraction, feature matching, SURF**

I. Introduction

For the last few years, similar image retrieval systems from large image database have been growing rapidly. Content-based image retrieval (CBIR) is a system which retrieves similar images of a query image from a image database without any help of caption, description of images.

During similar image retrieval, all retrieved images are not actually similar to the query image. This is because different objects can have same color, texture or shape. It is hard to differentiate objects with same color and texture. Yet, there is no universally applicable algorithm that can understand image as human does [1].

Similarity between two images is determined by checking the similarity of the contents of the two images, e.g., textures, colors, shapes etc. Feature extraction and feature matching are the two main processes of a CBIR system [2].

Any single technique usually can not extract features from an image by considering all aspects such as scale, rotation etc. There are also many good and bad aspects of similarity checking methods during similar image retrieval. No similarity checking method is perfect for matching all kinds of features. Objective of this work is to develop an efficient CBIR system which can retrieve similar images according to the query image by using the best combination of feature extraction and similarity checking methods.

II. Related Works

A CBIR system using color features (combined HSV, auto correlogram and color moments) and texture features (gabor wavelet) is proposed in [3]. It gets overall accuracy of 89.47% using a set of images as training images from a class and other images of the same class as test images. In case of another set of training images and test images it gets overall accuracy of 87.89%. In [4], color and texture features are used. Color Moments and Wavelet Packet Entropy are applied for extracting color and texture features respectively. Then, the correlation between those components of query image and images of the database are computed. The mostly correlated images with the query image are retrieved. This method shows better accuracy than the method in [3] for most of the cases by using 50 images of a class as training images and other 50 images of the same class as test images for all classes of WANG database [6], [7]. In [5], speeded-up robust features (SURF) is combined with the color features (color moments) which are calculated for a 5×5 fixed region around the SURF interest point for the RGB channel. It uses the KD-tree with the best bin first search algorithm to match the similarity between the features of the images. COIL-100 object database is used in this work which contains 72 views for 100 objects acquired by rotating the object under study about the vertical axis. This proposed method achieves 88% of average precision for 15 categories of 1080 images of COIL-100 database. A CBIR approach using SURF and fast retina keypoint (FREAK) is

proposed in [23]. This work is invariant to rotation and scale. But no color information is used in this approach since SURF and FREAK do not give color information. It uses Euclidean distance to measure the similarity between the query image and images of the dataset. Another CBIR approach based on visual words fusion of SIFT (scale invariant feature transform) and LIOP (local intensity order pattern) is proposed in [24]. This approach also uses Euclidean distance for retrieving similar images of the query image.

III. Feature Extraction

Generally, there are two kinds of features to describe the content of an image, global and local features [8]. Global features contain information of the whole image. On the other hand, local features are calculated from a local region of an image.

A. Global Feature Extraction

Generally, global features include color, texture and shape features. Color space is defined as a model for representing color in terms of intensity values [9]. There are several ways to extract color features such as, color histogram, color moments [10], color autocorrelogram [11], color layout descriptor (CLD) [12] etc. Texture is a property of image region rather than just pixel value like color [13]. There are many ways to extract texture features, such as using gray-level co-occurrence matrix (GLCM), wavelet moment, wavelet entropy [4] etc. Shape feature represents contours and shapes of objects in the image. There are many ways to extract shape features [14], such as, Hu's moment invariants [15].

1) Gray-Level Co-Occurrence Matrix (GLCM): Co-occurrence matrix describes the spatial relationships among pixels in a grayscale image. Gray-level co-occurrence matrix (GLCM) extracts texture features from an image by calculating the number of occurrence of pair of adjacent pixels with same intensity value. Similarity of pixels is calculated considering the horizontal, vertical or diagonal relation between the adjacent pixels.

B. Local Feature Detection and Extraction

Local features can be points, edges or small image patches [16]. Local feature detection means finding a pixel or a set of pixels from an image which are different with respect to color, texture from their neighbor pixels. These are called interest points. Local feature extraction means computing features around the detected interest points. Some feature detectors and descriptors are: Speeded-Up Robust Features (SURF) [17], Features from Accelerated Segment Test (FAST) [18], Binary Robust Invariant Scalable Keypoints (BRISK) [19] etc.

1) Speeded-Up Robust Features (SURF): The Speeded up robust features (SURF) is a scale and rotation-invariant blob-like interest point detector and descriptor which is computationally very fast [17]. To improve the speed of summing intensity values of an area of an image, SURF uses integral images. SURF descriptor extracts a distribution of haar-wavelet responses around the detected interest point [17].

Integral image is used for calculating the sum of intensities of a region of an image using only four memory look ups. The entry of an integral image $I_{\sum}(X)$ at a location $X = (x, y)^T$ is

$$I_{\sum}(x) = \sum_{i=0}^{i \le x} \sum_{j=0}^{j \le y} I(i,\ j). \tag{1}$$

However the region size is, calculation time for finding the sum of pixel values is independent [17].

SURF detector detects interest points (blob points) using Hessian matrix. The point where the determinant is maximum is considered as an interest point [17]. The Hessian matrix $\mathcal{H}(\mathrm{X},\ \sigma)$ for a pixel $X = (x, y)$ in an image I at scale σ is defined as follows:

$$\mathcal{H}(X, \sigma) = \begin{bmatrix} \mathrm{L_{xx}}(\mathrm{X},\ \sigma) & \mathrm{L_{xy}}(\mathrm{X},\ \sigma) \\ \mathrm{L_{xy}}(\mathrm{X},\ \sigma) & \mathrm{L_{yy}}(\mathrm{X},\ \sigma) \end{bmatrix}, \tag{2}$$

where $L_{xx}(\mathrm{X},\ \sigma) = \frac{\partial}{\partial x^2} g(\sigma)$, which is the convolution of the Gaussian second order derivative at pixel X, and similarly for $L_{xy}(\mathrm{X},\ \sigma)$ and $L_{yy}(\mathrm{X},\ \sigma)$. SURF approximates second order Gaussian derivatives with box filters [17]. The determinant of the Hessian matrix is calculated as:

$$det(\mathcal{H}_{approx}) = D_{xx}D_{yy} - (wD_{xy})^2, \tag{3}$$

where D_{xx}, D_{yy} and D_{xy} are the box filters for approximating Gaussian second order partial derivative in x, y and xy direction respectively. Here, w is kept constant at 0.9 [17]. As SURF is scale-invariant, it detects interest points at different scales. A fixed number of scales are defined for each octave by repeatedly smoothing the image with a Gaussian. In SURF, scale space is determined by increasing the box filter size. As an example, for the third octave, the filter sizes are 27-by-27, 51-by-51, 75-by-75, ... The initial scale is $s = 1.2$ which is the output of the 9×9 filter [17].

During local feature extraction, a circular region is taken around the interest point of size $6s$, where s is the scale at which interest point is detected. Haar-wavelet responses are calculated in both vertical and horizontal directions within the circular region. The dominant orientation is taken in the direction where the two summed responses are longest [17]. A squared window of size $20s$ is taken around the interest point. The window is oriented along the dominant orientation. To extract the features, for each 4×4 square of the window, the wavelet responses are computed from 5×5 samples [17]. For each 5×5 region, a four-dimensional feature vector is calculated.

$$V = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|), \tag{4}$$

where d_x is the haar-wavelet response in horizontal direction and d_y is the haar-wavelet response in vertical direction [17]. This results in a feature vector of length 64 for all subregions within the window.

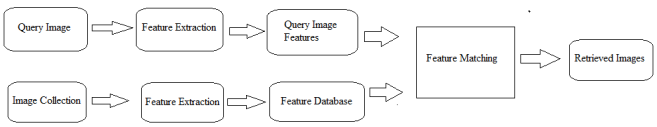


Fig. 1. The proposed CBIR system.

IV. Feature Matching

There are several ways of measuring distance between features of a query image and an image from a database. These are: euclidean distance (L2), standard euclidean distance, canberra distance, mahalanobis distance, city block/manhattan distance, relative manhattan distance (L1), minkowski distance, chebyshev/chessboard distance, cosine distance, correlation, spearman, relative deviation, fast approximate nearest neighbor search using multiple randomized kd-trees and hierarchical k-means tree [20]. Among them fast approximate nearest neighbor search is ideal for matching local features.

A. Fast Approximate Nearest Neighbor Search using Multiple Randomized Kd-trees and Hierarchical K-means Tree

An automatic process is proposed in [20] for selecting algorithm between multiple randomized kd-trees and hierarchical k-means tree for finding approximate nearest neighbor. One of these two algorithms gives the best performance depending on the dataset and desired precision.

Let, X is a high-dimensional feature vector and d_1 is a dimension of X which has largest variance. At the first stage of creating a normal kd-tree, X is split in d_1 dimension. Similarly, for each level of the tree, X is split in d_i dimension, where d_i dimension shows the largest variance. On the other hand, randomized kd-tree is built by splitting X in d_j dimension, where d_j is chosen randomly from first 5 dimensions which show largest variance [20]. The randomized kd-tree is created k times depending on the search time and tree build time. Each of the k trees is created on the rotated dataset [21]. By searching the kd-tree, similar features are found.

During creation of hierarchical k-means tree, X is split into K regions using k-means clustering for each level of the tree. This process is repeated for each regions until the number of points of a region become less than K. By searching the hierarchical k-means tree in a best-bin-first manner, similar features are found [20].

V. The Proposed System

The proposed CBIR system is shown in Fig. 1. To retrieve similar images, the feature vectors are extracted from the given query image. Similarities between the feature vectors of the query image and the feature vectors of the database images are then calculated. The retrieved images are the images with features that match with that of the query image in term of similarity.

Different combinations of feature extraction and matching techniques are investigated to find the best approach.

A. Algorithms for Feature Extraction

1) Color Feature Extraction Using Color Moments: In the proposed method, first, two color moments, mean and standard deviation are used. To do this, color moments are computed for each channel of RGB image. Thus, total 6 color moments are obtained for each image.

2) Texture Feature Extraction Using Gray-Level Co-Occurrence Matrix (GLCM): The RGB image is converted to the gray scale intensity image. Then the range of intensity values of the image are changed from 0 - 255 to 0 - 7. The GLCM is created by calculating how many times a pixel with value *i* occurs horizontally adjacent to a pixel with the value *j*. Thus, a feature vector of size (8×8) is obtained. The feature vector is simplified by resizing it to (1×64) size. Another GLCM is also obtained by calculating how many times a pixel with value *i* occurs vertically adjacent to a pixel with the value *j*. Finally, combining the 2 results, feature vector of size (1×128) is obtained. This is referred as 'modified GLCM' in this context.

3) Shape Feature Extraction Using Hu Moment Invariants: In this work, the seven Hu moment invariants are used as shape features. The Hu moment invariants are described in [15]. These seven moments are combined to make a vector of size 1×7.

4) Local Feature Detection Using SURF: At first, RGB image is converted to the gray scale intensity image. Then local feature (blob) is detected using SURF detector based on the hessian matrix as described in section III-B1. Strongest feature threshold = 500.0, number of octave = 3, number of scale levels per octave = 4 are used during SURF point detection. The number of points detected, location of the detected points, scale at which the interest points were detected, strength of detected feature, sign of the laplacian and orientation of the detected feature (specified as an angle, in radians) from the SURF points are obtained during the detection process.

5) Local Feature Extraction Using SURF: SURF algorithm derives the descriptors from pixels surrounding an interest point as described in section III-B1. The interest points can be detected using SURF, FAST or BRISK detector. Each interest point determines the center location of a neighborhood for extracting local features. If n is the number of detected interest points, then SURF descriptor extracts features of size ($n \times 64$).

B. Algorithms for Feature Matching

If F_q is the feature vector of query image and F_d is the feature vectors of the database images, then the similarity between the F_q and F_d can be computed using any similarity checking methods as described in section IV. Similarity checking methods give the distance between two feature vectors. The distance output is sorted in ascending order and the images for which the distance shows minimum value output, the corresponding images are retrieved as similar images of the query image.

1) Global Feature Matching: Euclidean distance (L2), standard euclidean distance, mahalanobis distance, city block/manhattan distance, relative manhattan distance (L1),

minkowski distance, chebyshev/chessboard distance, cosine distance, correlation, spearman and relative deviation are used for checking similarity between the global features of query image and database images.

2) Local Feature Matching: Generally, the local feature vectors are higher dimensional, large and SURF, FAST and BRISK give different number of interest points for different images. For this reason, feature matching process can be tedious. To solve this problem, an efficient approximate nearest neighbor search [20] which selects the best algorithm, between multiple randomized kd-tree and hierarchical k-means tree, based on the desired search precision, is used for local feature matching (described in section IV-A). In this approximate nearest neighbor search process, search precision 70% is chosen, meaning that 70% returned neighbors are exact nearest neighbors and 30% are just approximations. If two feature vectors are f1 and f2, then multiple features in f1 can match to one feature in f2. In this case, the best match is kept. Distance between a query point and a node of the tree is measured using sum of squared differences. After the completion of the nearest neighbor search, matching features and the distances between those features, from the two feature vectors of the images, are obtained. This nearest neighbor search is done separately between each images of the database and the query image to form a new vector:

$$f_i = \begin{matrix}[number\ of\ matching\ features\ with\ the\ query\ image,\\ sum\ of\ distances\ of\ the\ matching\ features],\end{matrix} \quad (5)$$

for each images of the database (In case of WANG database, i = 1,...,999, except the query image itself). Let, number of matching features with the query image = n and sum of distances of the matching features = d. Now the database images are sorted according to two criteria: (I) n in descending order (II) d in ascending order when $n_i = n_j$, where i and j are two different images. Here, n is much bigger number than d. So, at first, n and d are obtained by comparing query image with itself which gives, f_0 = [number of matching features (a large number), 0]. Then f_0 is compared with the vector mentioned in (5) using cityblock distance for each images of the database except the query image. The cityblock distances are sorted in ascending order. Thus, sorted similar images of the query image are obtained according to the two criteria.

VI. Experimental Results

WANG database [6], [7] is used for testing. It contains 1000 images of 10 classes, each class having 100 images. The 10 classes are African people, beaches, buildings, buses, dinosaurs, elephants, flowers, horses, mountains and food. The 10 classes are used for relevance measurement [22]. For a query image, all images from the same class of the query image are similar to that image.

To evaluate the performance of each method, the accuracy of each method is calculated. Here, accuracy is the ratio of the number of similar images retrieved to the total number of images retrieved expressed in percentage [22]. For example, if number of retrieved image for a query image is 20 and 14 of them are similar images, then the accuracy is 70%.

$$\text{Accuracy} = \frac{Number\ of\ relevant\ images}{Total\ number\ of\ images\ retrieved} \times 100, \quad (6)$$

where total number of images retrieved = number of similar images + number of dissimilar images.

In this work, accuracy is calculated by considering 20 retrieved images for each query image. Here, each of the 1000 images of WANG database is considered as query image. Thus, total number of image retrieved = (1000 $\times$ 20) = 20000.

Since this work is unsupervised, no supervised classification method is used. If the query image is from class 1 then the class 1 is the positive class and the classes except the class 1 is the negative class jointly. Therefore, the top 100 retrieved images for the query image are the images which are predicted as the members of the positive class. Because, each class has 100 images. Since the number of retrieved images which are predicted as the members of the positive class is always same, the accuracy, precision and recall values are always same. For example, top 100 images are retrieved for a query image of class 1. Among the 100 images, if 70 images are actually from class 1 then the number of true positives are 70, other 30 images are false positives and number of false negatives is also 30. So, the accuracy, precision and recall all are 70%. For this reason, precision and recall are not used here to measure the performances.

Table I shows the accuracies for each of the above described global feature extraction techniques (the one feature matching technique is included in the table which gives the best accuracy for a global feature extractor). From Table I, it is observed that, among the color feature extractors, the best accuracy is obtained using color histogram which is **60.62%** using the relative manhattan (L1) distance. Among the texture feature extractors, the best accuracy is obtained using modified GLCM which is **57.06%** using the relative manhattan (L1) distance. The best color and texture feature extractors color histogram and modified GLCM are combined and thus, a better accuracy of **63.32%** is obtained using relative manhattan (L1) distance. Then the best color and texture feature extractors are combined with the shape feature extractor invariant moment which gives accuracy of **63.32%** using the relative manhattan distance (L1). To decrease the computational expense, invariant moment is omitted as it is not upgrading the accuracy. Table I also contains the accuracy of the combinations which are used in [4]. Combination of color histogram and modified GLCM outperform them.

Table II shows the accuracies for each of the local feature detector and extractor method as described in section III-B (fast approximate nearest neighbor search [20] is used for similarity checking of local features). From Table II it is observed that, the best accuracy is obtained using the SURF detector and SURF descriptor which is **37.86%**. It is less than the accuracy achieved in [5], where SURF alone gives **65.47%** average precision. In [5], COIL-100 image database is used. On the other hand, here WANG image database is used where

TABLE I
ACCURACY OF ALL GLOBAL FEATURE EXTRACTION METHODS

Feature extraction technique	Feature matching technique	Accuracy	Feature extraction technique	Feature matching technique	Accuracy
Color Histogram	Relative Manhattan (L1)	60.62%	Wavelet Moment	Standard Euclidean	35.06%
Color Moments	Mahalanobis	57.45%	Wavelet Entropy	Mahalanobis	53.70%
Color Autocorrelogram	Relative Manhattan (L1)	23.02%	Color histogram + modified GLCM	Relative Manhattan (L1)	63.32%
Color layout descriptor (CLD)	Euclidean (L2)	52.25%	Color histogram + modified GLCM + Invariant moments	Relative Manhattan (L1)	63.32%
GLCM	Relative Manhattan (L1)	52.83%	Color Moments + Wavelet Entropy [4]	Standard Euclidean	55.77%
Modified GLCM	Relative Manhattan (L1)	57.06%			

there are 10 classes of images of different things and in each class each image is actually different images of similar things. But in COIL-100 database, each class contains same image repeatedly by just rotating it. So, images of WANG database are far more random than images in COIL-100 database. As a result, accuracy is much lower in case of WANG database.

TABLE II
ACCURACY OF ALL LOCAL FEATURE EXTRACTION METHODS.

Method Name	Accuracy
FAST detector + SURF descriptor	32.34%
BRISK detector + SURF descriptor	27.41%
SURF detector + SURF descriptor	37.86%

The second best color feature extractor color moments are also calculated within a region around the detected blob points by SURF. Color moments are calculated for a region of size 6s (orientation is also computed within a circular neighborhood of radius 6s around the blob point during the computation of SURF descriptor) around the blob points (where s is the scale at which the blob point is detected by SURF detector). So, now feature vector size is (n$\times$(64+6)) where n is the number of detected blob points by SURF. This method increases accuracy of SURF from **37.86%** to **57.34%**. The best color feature extractor color histogram is not used here, because the feature size would be (n$\times$(64+768)), which is very large.

Now SURF detector and descriptor and the best color and texture feature extractor global color histogram and modified GLCM are combined and thus, the accuracy of **63.56%** is obtained. In this case, similar images of the query image are obtained by sorting the sum of cityblock and L1 distances in ascending order. Here, the cityblock and L1 distances are obtained during matching of SURF descriptors (described in section V-B2) and during matching of color histogram and modified GLCM respectively.

If SURF detector and descriptor with color moments within the region of size 6s around the detected blob points and the best texture feature extractor global modified GLCM are combined, then the accuracy of **70.48%** is obtained. In this case, the cityblock distance obtained during matching of SURF descriptor with color moments (described in section V-B2) and L1 distance of modified GLCM are added and after sorting them in ascending order, similar images of the query image are obtained.

If SURF detector and descriptor with color moments within the region of size 6s around the detected blob points and the best color and texture feature extractor global color histogram and modified GLCM are combined, then the accuracy of **68.61%** is obtained. In this case, the cityblock distance obtained during matching of SURF descriptor with color moments and L1 distance of color histogram and modified GLCM are added and after sorting them in ascending order, similar images of the query image are obtained.

So, the best accuracy of **70.48%** is obtained by using SURF detector and descriptor with color moments within the region of size 6s around the detected blob points and global modified GLCM. Table III shows the accuracies of this method and other similar methods.

The performance of the method in this paper is compared with the method in [4] by showing an example.

TABLE III
COMPARISON OF THE APPROACHES IN THIS WORK AND SIMILAR WORKS.

Approach	Accuracy
CBIR using SURF and colour moments [5]	55.5%
CBIR using color moments and wavelet packet entropy [4]	50.13%
Feature matching technique: standard euclidean distance instead of correlation in the approach [4]	55.77%
Approach in this proposed work	70.48%

Fig. 2 and 3 show 20 most similar retrieved images for a query image by this proposed method and method in [4]. It is observed that, this proposed method retrieves all of the 20 images correctly. On the other hand, the method in [4] retrieves 18 images correctly.

From the experimental results it is observed that, the best result is obtained by combining global modified GLCM and

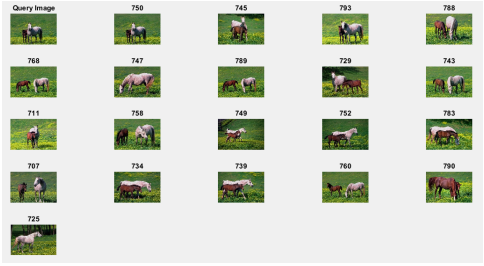


Fig. 2. Retrieved images for a query image by this proposed method.

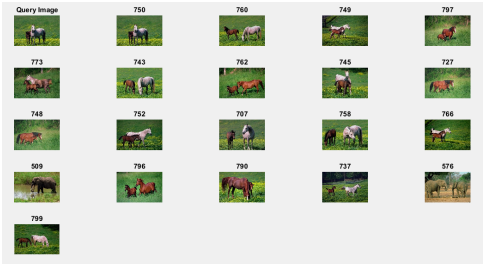


Fig. 3. Retrieved images for a query image by method in [4].

SURF detector and descriptor with color moments which are calculated within a region of size $6s$ (s = scale at which blob point is detected) around the detected blob points where L1 distance is used for matching modified GLCM features and efficient approximate nearest neighbor search [20] is used for matching SURF features with color moments. This approach gives better accuracy than the approaches in [4] and [5], where the approach in [4] gives better accuracy than the approach in [3].

VII. Conclusion

In this paper, various global and local feature extraction methods based image retrieval techniques are investigated. The performances of these feature extraction methods based on their similar image retrieval accuracies are calculated on WANG dataset of 1000 images. This approach uses combination of SURF detector and descriptor with color moments as local features and modified GLCM as global features. Color moments are calculated in the region around the SURF blob points which gives the local color information. Both local and global features are used because only local features are not suitable when the variety of images is large. As an example, bus and building both have many straight edges and corners. In this case, local feature detectors would detect lots of interest points on the edges and corners, which might make it difficult to differentiate between the two images. On the other hand, global features are biased towards the background of the images. As an example, if two images of horse and elephant both have natural background, then similar features are likely to be made by global feature extractors because it calculates features involving all pixels of image. Therefore, both global and local features are important in case of similar image retrieval from large image database. The future scopes of CBIR are very large. Detecting round, square, triangle objects and extracting their features might bring better performance.

References

[1] Datta, R., Li, J., & Wang, J. Z., "Content-based image retrieval: approaches and trends of the new age", In Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval, 2005, November (pp. 253-262). ACM.

[2] Kumar, D. A., & Esther, J., "Comparative study on CBIR based by color histogram, Gabor and wavelet transform", International Journal of Computer Applications, 2011, 17(3), 37-44.

[3] Singla, A., & Garg, M., "CBIR approach based on combined HSV auto correlogram color moments and Gabor wavelet", International Journal of Engineering and Computer Science, 2014, 3(10), 9007-9012.

[4] Islam, S. M. Mohidul & Debnath, R., "An RST invariant image retrieval approach using color moments and wavelet packet entropy", In Informatics, Electronics and Vision (ICIEV), 2016 5th International Conference on. IEEE, 2016, May, pp. 747-752.

[5] Velmurugan, K., & Baboo, L. D. S. S., "Content-based image retrieval using SURF and colour moments", Global Journal of Computer Science and Technology, 2011.

[6] Li, J., & Wang, J. Z., "Automatic linguistic indexing of pictures by a statistical modeling approach", IEEE Transactions on pattern analysis and machine intelligence, 2003, 25(9), 1075-1088.

[7] James Z. Wang, Jia Li, Gio Wiederhold, "SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol 23, no.9, pp. 947-963, 2001.

[8] Hassaballah, M., Abdelmgeid, A. A., & Alshazly, H. A., "Image Features Detection, Description and Matching", In Image Feature Detectors and Descriptors. Springer International Publishing, 2016, pp. 11-45

[9] Wang, J. Z., & Du, Y., "Scalable integrated region-based image retrieval using IRM and statistical clustering", In Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries. ACM, 2001, January, pp. 268-277

[10] Keen, N., "Color moments", School Of Informatics, University Of Edinburgh, 2005

[11] Hazra, D., "Retrieval of color images using Color Correlogram and wavelet filters", In Proc. of Int. Conf. on Advances in Computer Engineering, 2011

[12] Laencina Verdaguer, S., "Color based image classification and description", Master's thesis, Universitat Politecnica de Catalunya, 2009

[13] Howarth, P., & Ruger, S. M., "Evaluation of texture features for content-based image retrieval", In CIVR, 2004, July, Vol. 3115, pp. 326-334

[14] Latecki, L. J., Lakamper, R., & Wolter, D., "Shape similarity and visual parts", In DGCI, 2003, November, pp. 34-51

[15] M. K. Hu, "Visual Pattern Recognition by Moment Invariants", IRE Trans. Info. Theory, vol. IT-8, pp.179187, 1962

[16] Tuytelaars, T., & Mikolajczyk, K., "Local invariant feature detectors: a survey", Foundations and trends in computer graphics and vision, 2008, 3(3), 177-280.

[17] Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool., "SURF:Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), 2008, Vol. 110, No. 3, pp. 346 - 359

[18] Rosten, E., & Drummond, T., "Machine learning for high-speed corner detection", Computer Vision-ECCV 2006, 430-443.

[19] Leutenegger, S., Chli, M., & Siegwart, R. Y., "BRISK: Binary robust invariant scalable keypoints", In Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011, November, pp. 2548-2555

[20] Muja, M., & Lowe, D. G., "Fast approximate nearest neighbors with automatic algorithm configuration", VISAPP (1), 2009, 2(331-340), 2.

[21] Silpa-Anan, C.,& Hartley, R., "Optimised KD-trees for fast image descriptor matching", In Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008, June, pp. 1-8

[22] Chadha, A., Mallik, S., & Johar, R., "Comparative study and optimization of feature-extraction techniques for content based image retrieval", arXiv preprint arXiv:1208.6335, 2012

[23] Jabeen, Safia, et al. "An effective content-based image retrieval technique for image visuals representation based on the bag-of-visual-words model", PloS one 13.4 (2018): e0194526.

[24] Yousuf, Muhammad, et al. "A novel technique based on visual words fusion analysis of sparse features for effective content-based image retrieval", Mathematical Problems in Engineering 2018 (2018).

Bangla Handwritten Character Recognition: an overview of the state of the art classification algorithm with new dataset

Md Nazmul Hoq
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
salim.cse@diu.edu.bd

Nadira Anjum Nipa
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
nadira.cse@diu.edu.bd

Mohammad Mohaiminul Islam
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
niazoys94@gmail.com

Sadat Shahriar
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
sadat.cse@diu.edu.bd

***Abstract*— Recognition of handwritten characters from Bangla handwritten texts is of immense importance considering the complexity of the task. Researchers have explored the task of recognizing Bangla handwritten digits, but a few numbers of published works are available for Bangla Handwritten Character Recognition (BHCR). In this paper, we present a comparative overview of classification algorithms for BHCR, which may help the researcher to decide an appropriate classification algorithm for their work. We have created a new dataset of Bangla handwritten characters from 150 volunteers at different levels. We extracted around 2500 samples of Bangla characters which consist of Bangla Vowels only. Histogram adjustment and other image preprocessing techniques are applied in handwritten characters before their classification. We compare the performance of seven commonly used classification algorithms for BHCR in terms of Sensitivity, Miss Rate, Specificity, Precision, Fall-out, F-score, and Overall Accuracy. This result shows that among the seven algorithms, the best performance is provided by Artificial Neural Network (ANN). Logistic Regression performed well compared to others in terms of the sensitivity, specificity and error rate. This comparative overview will help scientists, especially the new researchers to give a quick start with Bangla handwritten character recognition.**

Keywords*— Handwritten character recognition, Pattern recognition, Histogram of Oriented Gradient (HOG)

I. INTRODUCTION

Bengali script is a writing system which originates from Indian Subcontinent for the Bengali language and with the Assamese alphabet. Bangla script is the second most popular script in the Indian subcontinent and mostly used for Bangla, Ahamia and Manipuri languages. Among these languages, Bangla is the national language of Bangladesh and is the fifth most popular language of the world [1]. Around 210 million people use Bangla as their speaking and writing purpose as a first or second language. Therefore, automatic recognition of Bangla characters has massive importance.

There are many applications of Bangla handwritten character recognition, such as, Bangla Optical Character Recognition (OCR), National ID number recognition system, automatic license plate recognition system for vehicle and parking lot management system, post office automation, online banking, writer identification, searching in historical manuscripts, and others[11] .

The Bengali script has a total of 50 characters consisting of 11 vowels and 39 consonants. Figure 1 depicts the vowels. This script is written from left to right, and there are no cases like uppercase or lowercase letters. Bangla character has a unique feature which is called MATRA- a distinctive horizontal line running along the tops of the letters that links them together. Also, the Bengali script is less blocky and presents a more curvy shape that facilitates classification.

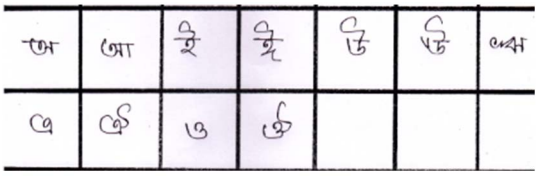


Fig. 1. Vowels in Bangla Script

BHCR is a classification task, which composes of several steps like handwritten image processing, feature extraction and applying classification algorithms. However, it is tough to determine which algorithm is the best compared to the others, as it depends on various factors, such as, classification accuracy, algorithm performance, computational resources and so on [4]. There is a trade-off between time complexity, resource constraints, and accuracy. For example, Artificial Neural Network (ANN) provides good result but takes much more time as compared to the others. Moreover, the accuracy of algorithms also depends on the quality and robustness of training dataset, the selection of the dataset used and different environmental settings. A sufficient number of training samples and their representativeness are critical for image classifications [5]. Kamavisdar, Pooja, et al. [6] described the advantages and disadvantages of classification algorithms like ANN, Support Vector Machine (SVM), Decision tree. However, they only described from the theoretical perspective and did not show any experimental results. Arun and Vidushi [7] compared different classification algorithms on multiclass dataset extracted from the multispectral satellite image. They

compared the result based on the average result of all classes and showed that KNN gives the best results over the others for satellite image. Although much of the previous research concerned explicitly with image classification, a comprehensive, up-to-date overview of BHCR is not available which inspired us to conduct a study on classification algorithms performance on BHCR.

II. Related work

Researchers have explored the area of recognition and classification of Bangla handwritten digit and character. Several machine learning algorithms for handwritten character recognition, such as Multilayer Perceptron, Support Vector Machine, Naïve Bayes, Bayes Net, Random Forest, J48 and Random Tree have been presented in the literature [8]. Most of these algorithms rendered a reasonably good result, but they do not focus on a wide range of evaluation matrices and visualizations of the results. Shyla Afroge et al. proposed a model which is based on Discrete Fréchet Distance for Bangla Optical Character Recognition which recognizes both characters and digits with 90-95% accuracy [9]. Some articles show that classic KNN provides excellent accuracy (95%) for BHCR [12]. Researchers also used Genetic algorithm (GA) to sample the optimal set of local regions of Bangla digits and then applied SVM classifier with popular optimization techniques like simulated annealing (SA) and hill climbing (HC) on the same data set to maximize recognition accuracies [13]. The geometric convex hull-based feature set for effective recognition of isolated handwritten Bangla basic characters and digits also maximize the success rate of 99.45% for Bangla handwritten digits [14].

Recently, ANN, especially deep learning along with Convolutional Neural Network used extensively in BHCR. Cheng-Lin Liua, Ching Y. Suenb worked with MLP neural network, modified quadratic discriminant function (MQDF), discriminative learning quadratic discriminant function (DLQDF), polynomial network classifier (PNC), class-specific feature polynomial classifier (CFPC), SVM classifiers and found that DLQDF, PNC, CFPC, SVM gave higher Accuracies than MLP and MQDF [15]. Deep Belief Network (DBN), CNN, CNN with dropout, CNN with dropout and Gaussian filters, and CNN with dropout and Gabor filters are also used in BHCR where CNN with dropout and Gabor features exceeds the state-of-the-art algorithms for HDBR with 98.78% accuracy [16]. Shopon et al. get even higher accuracy which is almost perfect (99.50%) using Convolutional neural networks (ConvNets) [17]. ConvNet used19,313 images from ISI handwritten character dataset for training and for testing, they used CMATERDB dataset. All of the approaches achieved good accuracy, but no comparative study is found in the literature which provides a clear idea on different classification algorithms performance evaluation. We have provided a comparative overview which provides a clear idea with different evaluation metrics and graphical visualizations.

III. METHODOLOGIES

BHCR is basically a multiclass classification problem. In Bangla script, there are total eleven vowels which are not case sensitive like English script. For multiclass classification, we have divided the whole task into several parts. The first part is feature extraction from the handwritten image and second part is implementing classification algorithms. This also includes post-processing and visualizations of the final results. Finally, we have assessed the accuracy of the whole system based on different evaluation metrics.

A. Feature Extraction

We have used Histogram of Oriented Gradient (HOG) for extracting the features from the image. HOG uses the histogram as distribution and directions of gradients which is called oriented gradients as the feature. Gradients are x and y derivatives of an image. HOG feature detector encodes local shape information from regions within an image. Figure 2 shows the gradients with direction. The first image is the original image and the second image is with the oriented gradient in red color. We have experimented with different cell size and find out that 4x4 provides reasonable results. We have got a total of 4356 features from HOG for each image which has been used as the input for the classification algorithms model.



Fig. 2. Feature Extraction from the image.

B. Classification algorithms

Classification is a subarea of artificial intelligence, which is used in many practical applications. In this article, different classification algorithms are discussed in the context of BHCR.

1) Feedforward Artificial Neural Network (ANN): Artificial neural network resembles the human brain system. Like the human brain, ANN is the collection of interconnected neurons which transfer information among them. Each connection of the neurons has numeric weights which will be updated based on adaptive learning. This weight is used to approximate function which uses inputs to estimate the output class. The multilayer perceptron is the most popular type of neural network in image classification [18]. Figure 3 depicts the used architecture of ANN in our system. However, it is challenging to decide the structure (no of hidden layers) in the network.

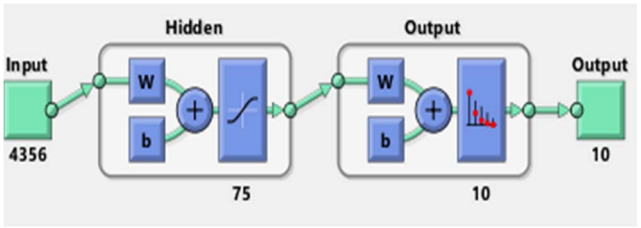


Fig. 3. Architecture of Applied ANN Model

2) Decision Tree (DT): Decision Tree calculates class membership by iteratively dividing a dataset into uniform subsets. DT is a kind of hierarchical classifier which permits the acceptations and rejection of class labels at each intermediary stage or nodes. DT consists of 3 parts: dividing the nodes, find the terminal nodes or leaf and allocation of the class label to leaf. DT is simple, computationally efficient and non-parametric methods which do not require any extensive design and training. But DT becomes complex when various

value is undecided or class label are missing and when various outcomes are correlated.

3) Random Forest (RF): Random Forests (RF) or random decision forests are an ensemble learning method for classification, regression and other tasks. RF constitutes of multiple DT which are used for training purpose. RF output the mode or mean class of all classes predicted from the collection of DT[19,20]. The number of the decision tree that will be combined to form the forest can be determined through a hyper-parameter at most implementation. MATLAB has one function named 'TreeBagger' which performs as Random Forest classifier

*4) Linear Discriminant Analysis (LDA):*The classifiers that assign data into two or more groups and predict categorical or grouping variable by one or more predictive variable is called as discriminant classifiers. They follow a set of rules in which maximum likelihood used for classification linearly and nonlinearly. It uses the multivariate normal distribution of the data depending on whether it is linear or nonlinear. Discriminant analysis can be linear and quadratic. LDA has average classification accuracy but this method is poor in object recognition.

5) Support Vector Machine (SVM): SVM is a supervised learning algorithm which is used for classification and regression. For classification, SVM builds a hyperplane or set of hyperplanes in a high or infinite dimensional space. Largest distance to the nearest training data point of any class (functional margin) gives good separation in the hyperplane that results in a good classifier and lowers the generalization error of the classifier by removing overfitting problems. Performance of SVM depends on the selection of hyperplane and kernel parameter. But determination of the optimal parameter is a little bit difficult for multiclass classification which is normally nonlinearly separable.

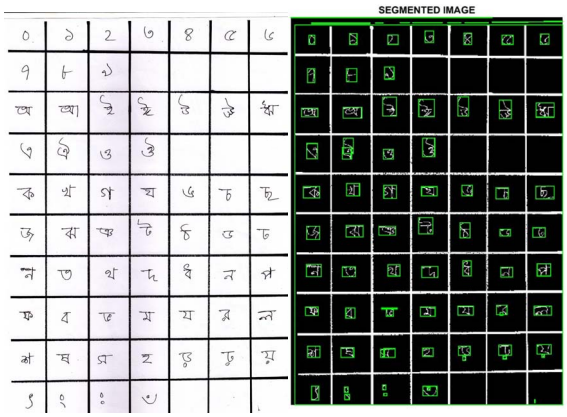
6) Logistic Regression (LogReg): Logistic Regression is one of the most commonly used classification algorithms in machine learning and data science. Logistic regression mainly used for the binary classification task. But logistic regression can also be used with multiple class classification tasks. For multi-class classification, the one-vs-all paradigm is used in logistic regression. Fine tuning of learning parameters like alpha, number of iterations in gradient descent gives good results. At the same time, poor selection of these parameter gives poor results and the different implementations of the gradient descent algorithm affect the results of the algorithm.

7) K Nearest Neighbor(KNN): KNN is also a distance-based classification algorithm. KNN classify an object based on the information of closest or most similar data points in the feature space. KNN use distance measure to compare the distance or similarity between two data points i.e. the Euclidean distance, Manhattan distance, and City Block distance etc. KNN select the label of the test data based on the label of surrounding data points or neighbor. KNN is the simplest but widely used classification method that has been applied in many fields like big data, biology, satellite image classification, medical image classification etc.

IV. EXPERIMENTAL SETUP AND RESULTS

A. Data Collection and Description

We created a new dataset of Bangla Handwritten Characters. We experimented on 150 volunteers. We collected data in a plain white page and then extracted the individual characters. Sample of individual handwriting is given in figure 4(a).



(a) (b)

Fig. 4. (a) Collected Handwritten Sample (b) Object Detection Sample

After that, we segmented the characters from the individual handwritten image. First we detected the characters using object detection algorithm and then segmented the individual's characters. Sample is shown in figure 4(b).

Finally, image preprocessing technique has been applied to the character image. Then we divided the dataset at a ratio of 70% and 30% for training and testing set respectively. We preprocessed and resize all handwritten characters (vowels) images into a fixed size of 50x50 pixels. The dataset was pre-labeled according to their containing image class. We use this label in the training phase of our classification algorithms. After that, we use the test set to measure the performances of the classification algorithm

B. Evaluation Criteria

We evaluated the classification algorithm performance for individual class rather than overall measurement. This will help us and others to decide which classification algorithm will be useful for characters (vowels) recognition and provide an insight for individual handwritten characters (vowels). We provided both the confusion matrix and Receiver operating characteristic (ROC) curve for the mentioned 8 image classification algorithms.

On the confusion matrix plot, the rows or x-axes correspond to the predicted class (Output Class) and the columns or y-axes correspond to the true class (Target Class) [22]. The second plot shows the receiver operating characteristic (ROC) for each of the output class. The quality of classifiers checked by the receiver operating characteristic [23]. Figure 05-12 shows the confusion matrix and ROC curve of the classification algorithm performances on test data. It has been found that ANN performs best with an

overall accuracy of 93.8% (Figure 5) and then LR performs with an overall accuracy of 86.8% (Figure 10). Similarly, RF (Figure 6) and SVM (Figure 12) provides more than 79% accuracy. Whereas, LDA (Figure 8) provides more than 73.1% accurate result. But KNN (Figure 12) and DT (Figure 7) perform poorly which gives only 68.3% 66.6% accuracy for the given dataset.

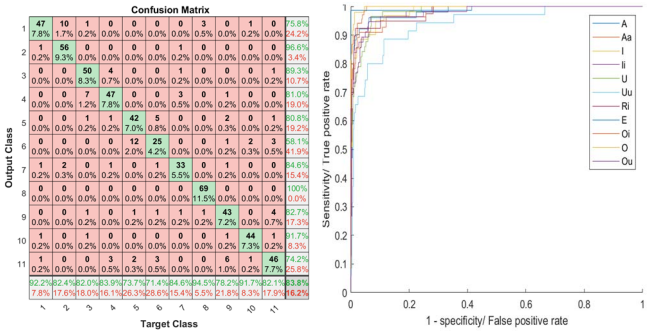


Fig. 5. Performance of ANN (Confusion Matrix and ROC curve)

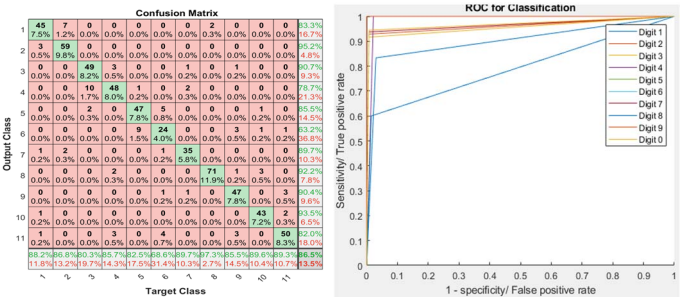


Fig. 6. Performance of RF (Confusion Matrix and ROC curve)

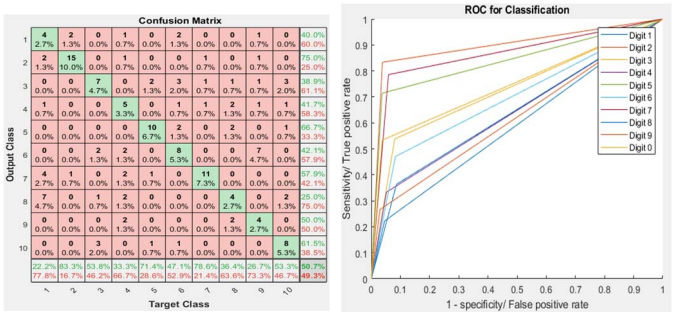


Fig. 7. Performance of DT (Confusion Matrix and ROC curve)

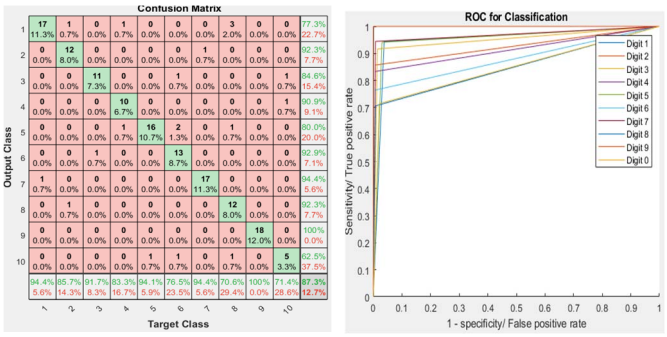


Fig. 8. Performance of LDA (Confusion Matrix and ROC curve)

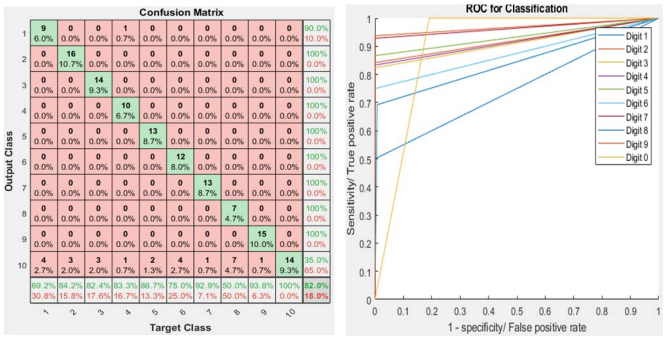


Fig. 9. Performance of SVM (Confusion Matrix and ROC curve)

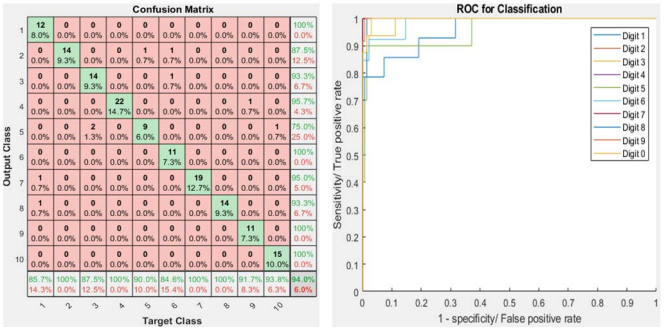


Fig. 10. Performance of LR (Confusion Matrix and ROC curve)

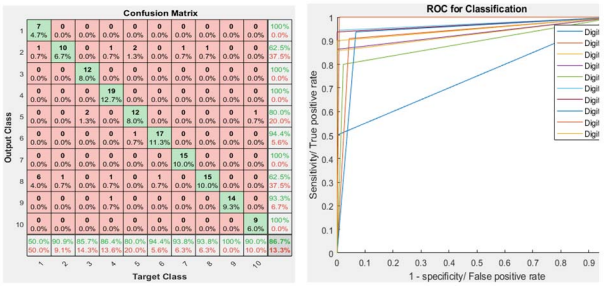


Fig. 11. Performance of NB (Confusion Matrix and ROC curve)

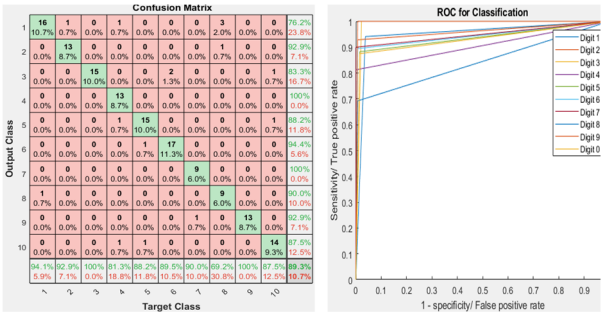


Fig. 12.Performance of KNN (Confusion Matrix and ROC curve)

As we measure class wise performance for individual characters (vowels) for each classification algorithms, we averaged the results of individual class or characters (vowels) to show an overview of the performance of whole classification algorithms. Table 1 summarizes the results. As expected ANN provides the best sensitivity, specificity, F score value.

TABLE I. COMPARISON OF AVERAGE PERFORMANCE OF INDIVIDUAL CHARACTERS (VOWELS) ON CLASSIFICATION ALGORITHMS

Algorithm	AS[a]	AMS[b]	ASP[c]	AP[d]	AF[e]	AFS[f]	AOA[g]
ANN	0.833	0.166	0.983	0.016	0.832	0.829	0.938
RF	0.857	0.142	0.986	0.013	0.858	0.856	0.865
DT	0.650	0.349	0.966	0.033	0.662	0.652	0.666
LDA	0.722	0.277	0.973	0.026	0.724	0.720	0.731
SVM	0.788	0.211	0.979	0.020	0.822	0.797	0.793
LR	0.864	0.135	0.986	0.013	0.865	0.861	0.868
KNN	0.680	0.319	0.968	0.031	0.687	0.676	0.683

[a.] Average Sensitivity, [b.] Average Miss Rate [c.] Average Specificity, [d.] Average Precision, [e.] Average Fall-out, [f.] Average F-score, [g.] Average Overall Accuracy

Finally, we graphically show the relative performance of different classification algorithm on the Bangla Handwritten characters (vowels) recognition. We have used the grouped bar graph which made it easy to understand the performance differences among different classification algorithms. X-axes of the bar graph represents grouped data to show the performance of individual characters (vowels). Y axes of the bar graph represents the metrics based on which the comparison is made. We have used total 08 metrics for understanding the difference in the performance of the classification algorithm. Error rate has been also measured Miss Rate (Type-II error). The result shown that ANN performs better and DT performs in a poor manner. In Figure 13, we showed Type-II error, performance on different classification algorithm with individual characters (vowels) performances. We got almost zero error for ANN but for DT the error is pretty high.

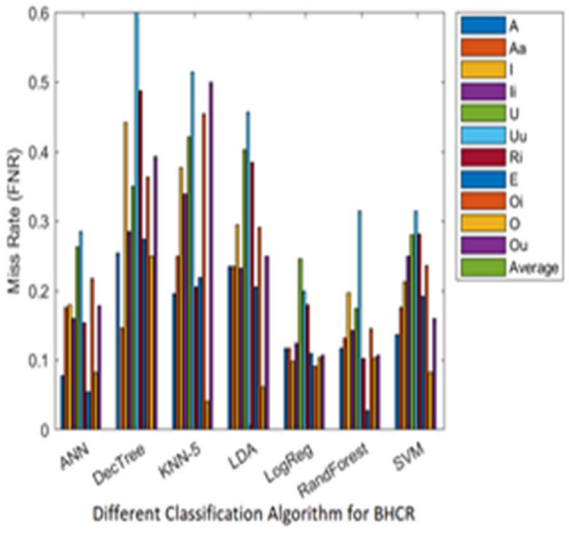


Fig. 13. ype-II error of discussed classification algorithms

V. DISCUSSION

We present a totally new dataset for Bangla handwritten characters. More than 150 students of Daffodil International University, a private university of Bangladesh has participated in the experiment with their verbal consent to use this data in research purpose. After that data set has been prepared for using in classification algorithm by thorough preprocessing described in the experiment section. Finally, Different supervised classification algorithms performance has been compared to find out the best classification algorithm. This result shows that among the seven algorithms, the best performance is provided by Artificial Neural Network (ANN). Among the old classic algorithm Logistic Regression performed surprisingly well compared to others. We consider each character as a class. For each class we measure the performance of the classification algorithm. Finally, we calculated the average performance of all classes (characters) for a single classification algorithm. ANN provides 93.8% average overall accuracy and logistic regression provides 86.8% average overall accuracy. Individual character performance has similar performance for all classification algorithms. For example vowel Uu (ঊ) and vowel U (ঊ) error rate is high in all classification algorithms. This will help ones to identify each character separately. Most of the algorithms presented in section-II do not focus on a wide range of evaluation matrices and visualizations of the results. Most of them used the metrics that provides the best result. We have shown 08 metrics. Moreover, we have shown graphical view of our outputs which enables one to capture the result easily.

VI. CONCLUSION

Our study has discussed the performance comparisons of mostly used classification algorithms for BHCR. We have found that Logistic Regression (86.8%) and ANN (93.8%) outperform all other classification algorithms. Besides, we have shown that individual handwritten characters (vowels) perform differently on the same classification algorithms. We considered only the supervised classification algorithms in the context of BHCR but did not consider the different approaches of classification like [4, 24, 25] which can be done in the future. In the future, we can extend our study on semi-supervised learning. We have also used only one feature extraction method to emphasize the performance of a classification algorithm on BHCR. In the future, multiple feature extraction method like zoning, template matching, deformable templates, unitary transforms etc can be used to show which feature extraction perform best. The performance of a different combination of feature extraction methods and classification algorithms can also be analyzed. We get the result after experimenting only on the vowels of Bengali alphabet. We can extend our works on consonant and especially on joint alphabet. Finally, this work can be extended to recognize Bangla handwritten text like handwritten letters, applications, exam scripts, different forms etc.

REFERENCES

[1] Basu, S., Sarkar, R., Das, N., Kundu, M., Nasipuri, M., & Basu, D. K. (2005, December). Handwritten Bangla digit recognition using classifier combination through DS technique. In *International*

Conference on Pattern Recognition and Machine Intelligence (pp. 236-241). Springer, Berlin, Heidelberg.

[2] S. Basu, N. Das, R. Sarkar, M. Kundu, M. Nasipuri, D. Basu, Recognition of numeric postal codes from multi-script postal address blocks, in: S. Chaudhury, S. Mitra, C. Murthy, P. Sastry, S. Pal (Eds.), Pattern Recognition and Machine Intelligence, Springer, Berlin, Heidelberg, 2009, pp. 381–386. [2] S.N. Srihari, Recognition of handwritten and machine-printed text for postal address interpretation, Pattern Recognit. Lett. 14 (1993) 291–302.

[3] S.N. Srihari, Recognition of handwritten and machine-printed text for postal address interpretation, Pattern Recognit. Lett. 14 (1993) 291–302.

[4] DeFries, R. S., & Chan, J. C. W. (2000). Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. Remote Sensing of Environment, 74(3), 503-515.

[5] Nath, S. S., Mishra, G., Kar, J., Chakraborty, S., & Dey, N. (2014, July). A survey of image classification methods and techniques. In Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014 International Conference on (pp. 554-557). IEEE.

[6] Kamavisdar, P., Saluja, S., & Agrawal, S. (2013). A survey on image classification approaches and techniques. International Journal of Advanced Research in Computer and Communication Engineering, 2(1), 1005-1009.

[7] Sharma, A., & Sharma, V. An Empirical Study of Supervised Learning Techniques on Multispectral Dataset.

[8] Shamim, S. M., Miah, M. B. A., Sarker, A., Rana, M., & Al Jobair, A. (2018). Handwritten Digit Recognition Using Machine Learning Algorithms. Global Journal of Computer Science and Technology.

[9] Afroge, S., Ahmed, B., & Hossain, A. (2017, February). Bangla optical character recognition through segmentation using curvature distance and multilayer perceptron algorithm. In *Electrical, Computer and Communication Engineering (ECCE), International Conference on* (pp. 253-257). IEEE.

[10] Basu, S., Sarkar, R., Das, N., Kundu, M., Nasipuri, M., & Basu, D. K. (2005, December). Handwritten Bangla digit recognition using classifier combination through DS technique. In International Conference on Pattern Recognition and Machine Intelligence (pp. 236-241). Springer, Berlin, Heidelberg.

[11] Alom, M. Z., Sidike, P., Hasan, M., Taha, T. M., & Asari, V. K. (2018). Handwritten Bangla Character Recognition Using the State-of-the-Art Deep Convolutional Neural Networks. Computational intelligence and neuroscience, 2018.

[12] Al-Amin, M., Islam, M. S., & Uzzal, S. D. (2017, February). Sentiment analysis of Bengali comments with Word2Vec and sentiment information of words. In Electrical, Computer and Communication Engineering (ECCE), International Conference on (pp. 186-190). IEEE.

[13] Surinta, O., Karaaba, M. F., Schomaker, L. R., & Wiering, M. A. (2015). Recognition of handwritten characters using local gradient feature descriptors. Engineering Applications of Artificial Intelligence, 45, 405-414.

[14] Das, N., Sarkar, R., Basu, S., Kundu, M., Nasipuri, M., & Basu, D. K. (2012). A genetic algorithm based region sampling for selection of local features in handwritten digit recognition application. Applied Soft Computing, 12(5), 1592-1606.

[15] Das, N., Pramanik, S., Basu, S., Saha, P. K., Sarkar, R., Kundu, M., & Nasipuri, M. (2014). Recognition of handwritten Bangla basic characters and digits using convex hull based feature set. arXiv preprint arXiv:1410.0478.

[16] Liu, C. L., & Suen, C. Y. (2009). A new benchmark on the recognition of handwritten Bangla and Farsi numeral characters. Pattern Recognition, 42(12), 3287-3295.

[17] Alom, M. Z., Sidike, P., Taha, T. M., & Asari, V. K. (2017). Handwritten bangla digit recognition using deep learning. arXiv preprint arXiv:1705.02680.

[18] Shopon, M., Mohammed, N., & Abedin, M. A. (2016, December). Bangla handwritten digit recognition using autoencoder and deep convolutional neural network. In Computational Intelligence (IWCI), International Workshop on(pp. 64-68). IEEE.

[19] Atkinson, P. M., & Tatnall, A. R. L. (1997). Introduction neural networks in remote sensing. International Journal of remote sensing, 18(4), 699-709.

[20] Ho, T. K. (1995, August). Random decision forests. In Document analysis and recognition, 1995., proceedings of the third international conference on (Vol. 1, pp. 278-282). IEEE.

[21] Barandiaran, I. (1998). The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence*, *20*(8).

[22] N. Das, J. M. Reddy, R. Sarkar, S. Basu, M. Kundu, M. Nasipuri, and D. K. Basu, "A statistical–topological feature combination for recognition of handwritten numerals," Applied Soft Computing, vol. 12, pp. 2486-2495, 2012.

[23] https://uk.mathworks.com/help/nnet/ref/plotconfusion.html last accessed on 27th June 2018 at 13:15

[24] https://uk.mathworks.com/help/nnet/ref/roc.html last accessed on 27th June 2018 at 13:20

[25] Nath, S. S., Mishra, G., Kar, J., Chakraborty, S., & Dey, N. (2014, July). A survey of image classification methods and techniques. In Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014 International Conference on (pp. 554-557). IEEE.

[26] Kamavisdar, P., Saluja, S., & Agrawal, S. (2013). A survey on image classification approaches and techniques. International Journal of Advanced Research in Computer and Communication Engineering, 2(1), 1005-1009.

Personality Detection from Texts using Convolutional Neural Network

1st Md. Abdur Rahman, 2nd Asif Al Faisal, 3rd Tayeba Khanam, 4th Mahfida Amjad, 5th Md Saeed Siddik
[1]Center for Advanced Research in Sciences, [2,5]Institute of Information Technology
[1,2,5]University of Dhaka, [3]AIUB, [4]Stamford University Bangladesh, Dhaka, Bangladesh
{[1]mukul.arahman,[2]asif.faisal.eee, [3]tko.ocean95, [4]mahfidaamjad}@gmail.com, [5]saeed.siddik@iit.du.ac.bd

***Abstract*—The distinguishing characteristics belong to a person are personality traits, which is predicted from person's behavioral pattern. As most of the people provide lots of information knowingly or unknowingly into their writings, it is possible to extract personality traits from those texts. Individual personality traits detection from texts, yields enormous possibilities toward various applications named as forensic department, mental health diagnosis, etc. Meanwhile, deep learning algorithm performs fairly well in text personality detection; however, its performance may vary with activation functions. Hence, this paper proposed an empirical approach to find the best personality detection performance by comparing several activation functions named sigmoid, $tanh$, and leaky relu. Here, text documents were pre-processed and vectorized for input in convolutional neural network. The input size was multiple to length of word, sentence, documents, and feature vectors. Five personality traits named as EXT, NEU, AGR, CON, and OPN have been used for experimental analysis. The result showed that $tanh$ and leaky relu performs over sigmoid in all datasets. The average F1-score of sigmoid, $tanh$ and leaky relu showed 33.11%, 47.25%, and 49.07% respectively. However, F1-score of Leaky relu was high only for CON, $tanh$ showed better result for others datasets. The overall performance showed by $tanh$ is better than sigmoid and leaky relu for personality detection from texts.**

***Index Terms*—Personality Traits, Sentiment Analysis, Convolutional Neural network, Deep Learning, Text Analysis.**

I. Introduction

Personality traits are distinguishing quality or characteristics typically belong to a person. The way that an individual talk and write is reflect upon their personality. A persons thought pattern is built upon the behavioral, emotional, mental and motivational attributes which collectively called personality. According to J. Digman, personality is typically assessed along big five personality traits which are the results of (i) Extroversion (EXT): is the person outgoing, talkative, and energetic or reserved? (ii) Neuroticism (NEU): is the person sensitive and nervous versus secure and confident? (iii) Agreeableness (AGR): is the person trustworthy, straightforward, generous, and modest versus unreliable, complicated, meager, and boastful? (iv) Conscientiousness (CON): is the person efficient and organized versus sloppy and careless? and (v) Openness (OPN): is the person inventive and curious versus dogmatic and cautious? [1].

Individual's personality detection from texts yields enormous possibilities toward various practical applications. For example, companies can establish their strategies based on positive sentiment of the people towards their products. Forensic department can exploit personality traits to reduce the circle of suspects. In the context of mental health diagnosis, certain diagnoses correlate with certain personality traits. In human resource management, detection of personality traits can graph the thought pattern which affect one's suitability for certain jobs. To exploit personality traits in such different applications, detection of an individual's personality from texts is essential. However, manual text analysis for personality traits identification is not feasible in terms of time, cost, and accuracy. Therefore, automatic personality traits detection system is required to exploit its strength effectively.

To this end, researchers introduced few works for personality traits detection using textual content analysis, which is a established field of Natural Language Processing (NLP). In 1999, a personality trait analysis dataset has been developed experimenting 2,468 peoples [2]. On the other hand, a document-level feature set for personality detection has been developed considering 84 features such as linguistic inquiry, word count features, medical research council features, etc. [6]. Using large text corpus continuous vector representations of words has been developed to strengthen machine learning based textual analysis research work [3]. To take the advantages of deep learning in text analysis, several sentence classification techniques have been applied using convolutional neural network [7]. Recently, document modeling has been carried out using Convolutional Neural Network (CNN) feature extractor for personality detection [5].

The existing personality traits detection strategies exploited deep learning algorithms to improve classification accuracy. However, performances of different activation functions in deep learning based textual data classification are yet to be explored. To this end, this research proposed an empirical approach to find the best personality traits extraction framework. Where, textual data analysis performance of sigmoid, hyperbolic tangent (tanh), and leaky rectified linear unit (leaky relu) activation functions has been compared. This method classified five different personality traits named as EXT, NEU, AGR, CON, and OPN using exclusive five special CNN networks considering same architecture. Where, each network acts as a binary classifier for the respective traits to be positive or negative. In this framework, first text documents were pre-processed to remove unwanted characters. Then, n-gram feature vectors obtained by feeding pre-processed data as

input in convolutional filters. These feature vectors aggregated and concatenated with Mairesse features to represent each document. Also, emotional features containing sentences were filtered for classification accuracy improvement. Finally, the document vectors were used as input to neural network with one hidden layer.

The empirical results showed that $tanh$ and leaky relu performed better over sigmoid activation function. The average training performance of sigmoid, $tanh$ and leaky relu are 51.19%, 56.15%, and 57.07% respectively. Where, three function's average validation performance are 49.30%, 53.90%, 53.52% and test performance are 49.48%, 54.29%, 54.24%. On the other hand, average F1-score of sigmoid, $tanh$ and leaky relu functions are 33.11%, 47.25%, and 49.07% respectively. Even though, average F1-score is high for Leaky relu, this function outperformed only for CON personality trait. However, for all other personality traits named as EXT, NEU, AGR, and OPN $tanh$ over performed. In a nutshell, the overall performance of $tanh$ function is better than sigmoid and leaky relu in personality detection using texts.

The rest of this paper is organized as follows; Section II discussed the existing work related to this thesis. Section III and IV illustrated the proposed method and result analysis respectively. Finally section V concludes this paper with future research direction.

II. Related Works

Automatic personality detection from text is relatively new concept in the field of Natural Language Processing (NLP). There have been few works implemented on this domain using NLP algorithm which are discussed in this section.

There have been many theories regarding the structure of personality. How a personality should be classified was a bigger challenge in research community. However, in 1990, John M. Digman, a researcher from University of Hawaii, has published a journal which shaped a persons personality into five different categories [1]. After that, it becomes the most reliable unified theory for personality till date. According to J. Digman, personality comes as a five factor model. The first of those five factors is extraversion or introversion, which means, a person talkative or not, more outgoing or stays inside more etc. The second factor is, agreeableness and defined by a person's paranoid, friendly or unfriendly, social or not etc. As for the third factor, conscientiousness defined as the prudence, judgment of a person. The fourth factor, neuroticism is defined as the mental or emotional stability of a person. Final factor is openness, which involves a persons intellect, creativity etc.

Experimenting 2,468 peoples, Pennebaker et al. has developed a dataset that is used for personality trait analysis research work [2]. Each of them had given a task to write whatever comes in their mind in a very short period of time in a closed environment. So, when people do not have enough time to think they tend to write something that reflects upon their personality. After that they had developed a text analysis application Linguistic Inquiry and Word Count to get better insight from the collected data.

On the other hand, Francois Mairesse and his colleagues developed a document-level feature set for personality detection, consisting of 84 features [6]. It comprised the linguistic inquiry and word count features, medical research council features, utterance-type features, and prosodic features. Examples of the features included in this set are the word count, average number of words per sentence, total number of pronouns, past tense verbs, present tense verbs, future tense verbs, letters, phonemes, syllables, and assertions in the document.

Tomas Mikolov et. al proposed two model for computing continuous vector representations of words using large data sets [3]. The quality of resulting vector representations was measured using word similarity and results were compared with existing best techniques. In this study, they have proposed two approaches to define word embedding. One is Continuous Bag of Words (CBOW) model and another one is Continuous Skip Gram model. In continuous bag of words model, target word is predicted from the context whereas, in skip gram model, a context is predicted from a target word.

In 2014, Yoon Kim showed several sentence classification based on convolutional neural network [7]. This work trained a simple CNN model using one layer of convolution and Googles pre-trained word vectors. Initially, keeping word vectors static this research learned only other parameters of the model. This model achieves excellent results on multiple datasets with hyper-parameters tuning. Finally, this paper concluded that word vectors is a key element for deep learning applications in NLP. However, efficiency of this sentence classification model could be improved using multiple convolution layer and extensive hyper parameters tuning.

Majumder et. al. proposed a novel document modeling approach using convolutional neural network feature extractor [5]. In this study, sentences from the stream-of-consciousness essay dataset were fed as input in convolution filter to obtain the sentence model in the form of n-gram feature vectors. The individual essay was represented by aggregating the vectors of its sentences. Then the obtained vectors were concatenated with the mairesse features. For final classification, this document vector is being fed into a fully connected CNN, which act as a binary classifier that predicted the corresponding trait to be positive or negative.

The analysis of existing natural language processing approaches shows that different topic modeling strategies have been proposed for personality trait analysis such as personality classification, personality dataset development, CNN based n-gram feature extraction, etc. However, no guideline has been found yet to compare different deep learning activation functions such as sigmoid vs tanh or tanh vs leaky relu etc. in personality analysis.

III. METHODOLOGY

Document modeling framework to extract personality traits using Convolutional Neural Network (CNN) is proposed to improve classification accuracy. This approach classified five different personality traits named as EXT, NEU, AGR, CON, and OPN using exclusive five special CNN networks in same

architecture. Each network acts as a binary classifier for the respective traits to be positive or negative. In this architecture, sentences are preprocessed to remove unwanted characters then feed in convolutional filters to obtain sentence model as n-gram feature vectors. These feature vectors are aggregated and concatenated with Mairesse features to represent each document [6]. Emotional features containing sentences are filtered for result improvement. The processed document vector is used as input to neural network with one hidden layer in the final classification. The proposed method is broadly divided into following 2 steps.

- Step 1: Data Preprocessing
- Step 2: Classification

The details regarding the steps are elaborated in following sections and sub-sections.

A. Step 1: Data Preprocessing

Initially standard text preprocessing steps named as sentence splitting, tokenizing, stemming, and case-folding applied for text preparation. Preprocessed document used for document level feature extraction using convolutional filter and Mairesse baseline feature-set. Extracted features from non-neutral sentence considered for training and others were discarded to reduce noise as well as input size. This process was considerably affected on training time. The word level features extracted using word2vec algorithm for each word in documents as word vectors [8]. Theses word vectors used to represent words of each sentence in documents. Thus, variable length words became fixed-length word feature vectors. Those generated word vectors were used for input of CNN network.

B. Step 2: Classification

Deep CNN used as the classification network, where initial layer process the text in a hierarchical manner. Initially, words and sentenced are represented as fixed length feature vector and variable length word vectors. However, at some network layer, the variable length vectors of each sentence reduced to fixed length vectors. At that level, documents are represented as a variable number of such fixed length sentence embedding. Similarly, variable length document vector also reduced to fixed length document vector at deep network layer. Finally, this fixed length feature vectors are concatenated with document level features for final classification. The whole classification process using neural network model are divided into following six layers:

- Layer 1: Word Vectorization
- Layer 2: Convolution (sentence vectorization)
- Layer 3: 1-max pooling (document vectorization)
- Layer 4: Concatenation (document vectorization)
- Layer 5: Linear with tanh and leaky relu activation (classification)
- Layer 6: Two neuron softmax output (classification)

The functions of each layer is described in the following sections.

1) Layer 1: Input (Word Vectorization: The text document in dataset was preprocessed then vectorized to feed as input in CNN. The dataset is represented as a set of documents where each document is a sequence of sentences, each sentence is a sequence of words, and each word is a real valued word vector. Thus, the input layer is a four dimensional real values array named as document, sentence, word, and embedding. All the documents are padded to keep uniform size of sentences. Where, shorter documents and sentences are padded with dummy sentences and words respectively for uniformities.

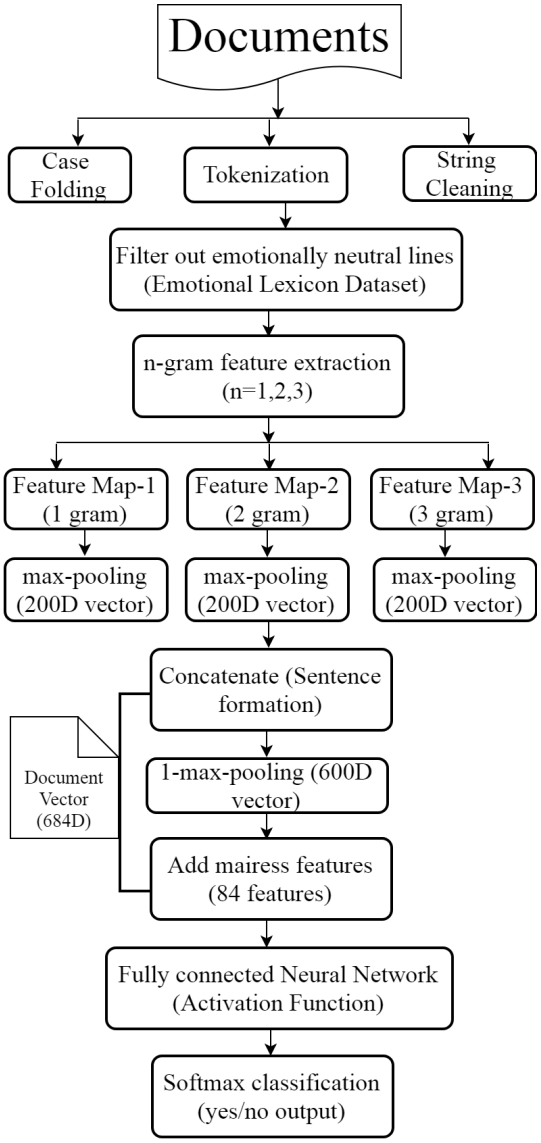


Fig. 1. Architectural overview of proposed system

2) Layer 2: Convolution (Sentence Vectorization): Convolutional filters with max pooling were applied used to extract n-gram features from each sentence; where n belongs to unigram, bigram and trigram. The size of convolutional filter is n-gram multiplied with length of word embedding on each sentence. A bias added with output filtering, which generated three feature maps for a given sentence. Then, max pooling

was applied on each feature map to further reduce sample dimension similar to n-gram feature dimension in convolution. The obtained feature vectors for unigram, bigram and trigram were concatenated to represent sentence vector. This process is applied to all sentences in the document.

3) Layer 3: Max Pooling (Document Vectorization): The sentence vectors obtained in previous step were concatenated to get document vectors, where dimension of each sentence vector is multiple of n-grams. If the document belongs to at least one feature, it considered for vectorization otherwise discarded. Thus featured documents converted to fixed size real valued vector. The maximum number of sentences across all documents was considered as the document vector size.

4) Layer 4: Concatenation (Document Vectorization): In this layer, document level features are added to document vector [6]. The considered features are words per sentences, type/token ratio, words captured, negations, negative emotions, anxiety, personal concerns, concreteness rating etc. Among those, 84 features are concatenated with document vector. The dimension of resultant vector changed to the summation of embedding word vector size and selected features size.

5) Layer 5: Non-linear with tanh and Leaky ReLU Activation (Classification): A fully connected two layer perceptron is used for activation in the final classification, where the layer size is r (real number). In this activation process, the document d is multiplied by a matrix W^{fc}. A bias B^{fc} is also added to it for calculating the document vector d^{fc}. Where, d, W, B, fc, denotes document, weight, bias value and fully connected network respectively. The dense layer output gives leads to linear value which affect on classification accuracy. Therefore, $tanh$ activation function used to ensure non linearity which improves classification result. Equation (1) denotes the $tanh$ activation function.

$$d^{fc} = tanh(dW^{fc} + B^{fc}) \tag{1}$$

where $tanh(x) = \frac{e^x - e^-x}{e^x + e^-x}$.

Another activation function Leaky Rectified Linear Unit (ReLU) also used to improve efficiency of the model. Equation (2) denotes this activation function.

$$d^{fc} = lrelu(dW^{fc} + B^{fc} \tag{2}$$

Where, $d^{fc} = 1(x < 0)(ax) + 1(x \geq 0)(x)$ and a is a small constant.

6) Layer 6: Two-Neuron Softmax Output (Classification): The two-neuron softmax function was used to measure probability of each document. The softmax produce either 0 or 1 as output result, this result indicates the class to which document belongs. Following vector was used for softmax probability determination.

$$x_{yes}, x_{no} = dW^{sm} + B^{sm} \tag{3}$$

Where, weight $W^{sm} \in R^{200\times2}$ and bias $B^{sm} \in R^2$. After that, the class probabilities calculated as,

$$P\frac{i}{network_{prs}} = \frac{exp(x_i)}{exp(x_{yes}} + exp(x_{no})for \quad i \in \{1, 0\} \tag{4}$$

In this research, the documents were reprocessed and vectorized to fed into CNN model. The N-gram feature extraction technique and max pooling were applied in this personality traits analysis framework. Finally, softmax classifier has been used to categorise each documents into five personality traits.

IV. Experimental Result

This section discusses the experiment of empirical approach for the proposed personality trait identification using convolutional neural network. The classification performance of used activation functions is compared with other significant activation functions. Details regarding the implementation setup of those functions are also provided in this section.

A. Environment Setup

The research work was performed on a single workstation having 3.6 GHz Intel Core i7 CPU and 32GB RAM running the Ubuntu 16.04 LTS version operating system. To implement this work python 2.7, python numerical computational library theano (version 0.7), and pandas 18.0 were installed.

B. Datasets

A popular collection of essays collectively called stream of consciousness has used for evaluation of proposed technique [2]. The dataset contains 2,468 random essays marked with the writer's personality traits which are, EXT, AGR, NEU, CON and OPN [2]. The rest of the datasets are subsidiary which were built upon these datasets. Mairesse listed a number of features based on these 2468 essays [6]. NRC emotion lexicon datasets has used to remove those words that are not emotionally significant [10]. Finally, Google's Pre trained Word Vectors and Phrases dataset has used to obtain word vectors of the essay dataset.

C. Stream of Consciousness Essays

This dataset is developed by James W. Pennebaker and his colleagues at University of Texas at Austin experimenting 2,468 people [2]. Each people of them had given a task to write whatever comes in their mind in a very short period of time in a strict environment. So, when people do not have enough time to think they will tend to write something that will reflect upon their personality. This was the main goal of this particular experiment.

D. NRC Emotion Lexicons

NRC word emotion lexicon aka EmoLex is a list of English words and their associations with eight basic emotions named as anger, fear, anticipation, trust, surprise, sadness, joy, and disgust with two sentiments (negative and positive) [10]. The annotations were manually given by crowdsourcing and this dataset consists of 14182 words that are identified to be emotionally charged or not.

E. Googles Pre-trained Word and Phrase Vectors

These are the pre-trained vectors trained on part of Google News dataset that contains about 100 billion words [8]. The model contains 300-dimensional vectors for 3 million words and phrases. The phrases were obtained using a simple data-driven approach that is skip-gram model of Word2Vec.

F. Mairesses Baseline Feature Set

Based on stream of consciousness essay dataset, Francois Mairesse and his colleagues have developed a document level feature set for personality detection that consists of 84 features [6]. Examples of the features included in this set are words per sentences, type/token ratio, words captured, words longer than 6 letters, prepositions, negative emotions, anxiety, etc.

G. Training

The dataset contains total 2,468 number of random essays among which 60%, 10%, and 30% data has taken for training, validation and testing respectively. Network parameters are randomly initialized and stochastic gradient descent with Adaleta update rules have used to adjust network parameters which reduced the negative log likelihood error [12]. Due to limitation of the hardware specifications the network is trained up to 6 epochs for each personality trait.

H. Experimental Data Preprocessing

The document preprocessing is important to remove unwanted text, symbols, and splitting documents into a list of words. This process consists of document tokenization, stop word removal, and word stemming. First, for tokenization, texts were splitted into sentences using period or question marks. Second, using apostrophe or white space character, sentences were splitted into words. Third, words were case-folded by reducing all letters to lowercase and ASCII characters also removed. The sentences longer than 150 words were splitted into sentence of 20 words.

After getting list of words, document level features extracted using Mairesses Baseline feature set [6]. The extracted document features used filtering out emotionally neutral words using NRC Emotion Lexicon dataset [11]. At last, word level features extracted, where Googles pre-trained vectors used to convert words into 300 dimensional vectors. The words assigned randomly to 300 coordinates with a uniform distribution from negative 0.25 to positive 0.25, when they are not found in the list.

I. Classification

This experiment has done with CNN and Mairesse setting, where a Multiple Layer Perceptron with one hidden layer is used for train the data. For final classification three activation functions namely Sigmoid Function, Hyperbolic Tangent Function (TanH), Leaky Rectified Linear Unit (Leaky RELU) have experimented.

J. Result Analysis

The prepared dataset is trained and tested to evaluate classifying accuracy of proposed model. The labeled data was used for training and individual written texts were used for testing. The experimental results of Personality Traits (PT) for different Activation Functions (AF) have shown in Table 1. The result includes training Performance (TrP), Value Performance (VP), Test Result (TR), and F-Score (F1).

TABLE I
EXPERIMENTAL RESULT FOR 2 AF

PT	AF	TrP	VP	Test P	F1(%)
EXT	Sigmoid	51.65%	56.50%	49.26%	33.00%
	Tanh	54.20%	56.50%	50.00%	39.40%
	Leaky ReLU	49.90%	44.09%	49.63%	35.02%
NEU	Sigmoid	50.65%	47.50%	47.40%	32.20%
	Tanh	57.75%	53.00%	58.14%	58.14%
	Leaky ReLU	56.60%	51.50%	55.20%	54.50%
AGR	Sigmoid	53.05%	53.50%	53.70%	35.00%
	Tanh	55.80%	50.50%	55.90%	46.16%
	Leaky ReLU	56.00%	49.50%	56.00%	45.70%
CON	Sigmoid	51.65%	46.50%	47.40%	32.16%
	Tanh	52.05%	47.00%	47.41%	32.76%
	Leaky ReLU	58.85%	55.50%	52.22%	52.11%
OPN	Sigmoid	48.95%	42.50%	49.63%	33.20%
	Tanh	60.95%	62.50%	60.00%	59.80%
	Leaky ReLU	64.00%	67.00%	58.15%	58.00%

From Table I, it is observed that overall performance of hyperbolic tangent function is better than sigmoid or Leaky ReLU. However, for some cases Leaky ReLU outperformed than other two functions. For example, agreeableness personality trait test result has found 56.00% for Leaky ReLU where 53.70% and 55.90% is for Sigmoid and $tanh$ respectively. Similarly, for Conscientiousness, test result has found 52.22% for Leaky ReLU where 47.40% and 47.41% is for Sigmoid and $tanh$ respectively.

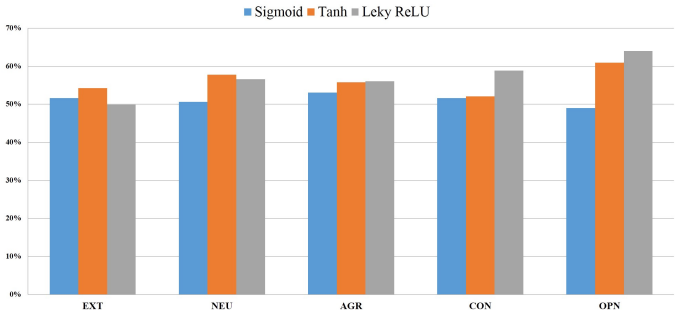


Fig. 2. Training Performance Comparison Chart

Figure 2, is showing a comparison chart for training performance using sigmoid, $tanh$ and leaky relu activation function. Tanh and leaky relu performed better for all the cases except for the extroversion trait where sigmoid performed better than leaky relu.

As for the test results shown in Figure 2, $tanh$ and leaky relu outperformed the sigmoid activation for all of the personality

traits. Moreover to this, $tanh$ is seems like the best performer amongh these three.

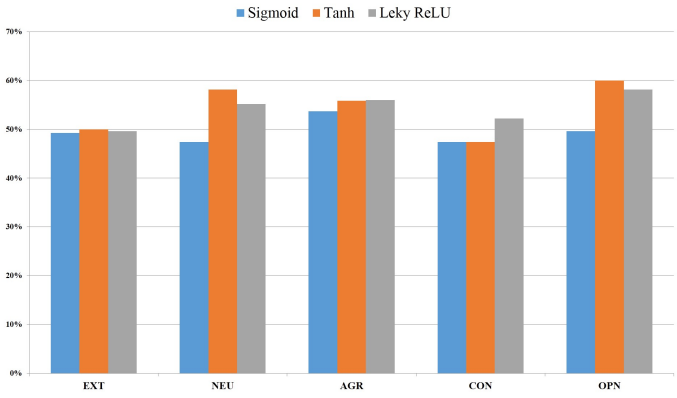


Fig. 3. Test Result Comparision Chart

Proposed model's classification accuracy was estimated with F1-Score to evaluate how accurately this model can classify unknown samples. F1 score is the weighted average of precision and recall considering both false positive and false negative. This score conveys the balance between precision and recall. In this strategy, the F1 score is calculated by equation (5).

$$F1 - score = 2 * (\frac{Precision * Recall}{Precision + Recall}) \tag{5}$$

In Table I and Figure 4, F1-score of sigmoid activation is very poor for all the personality traits whereas hyperbolic tangent and leaky relu performed better. Particularly, leaky relu outperformed on conscientiousness data set, however, $tanh$ performs better on all other datasets.

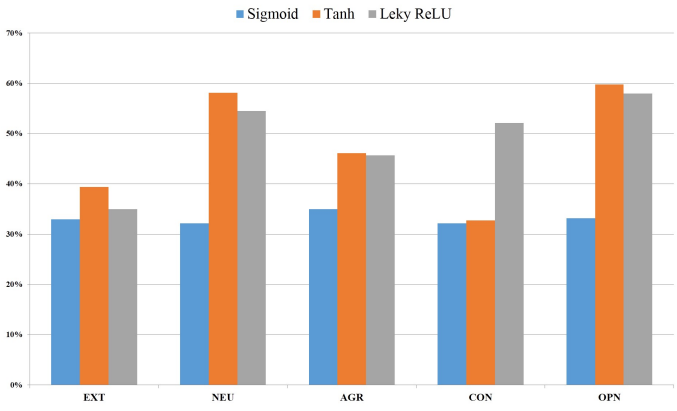


Fig. 4. F1-Score Comparison Chart

All the empirical results showed that leaky relu outperformed on conscientiousness data set only, where $tanh$ performs best over other datasets. Therefore, this research concluded that $tanh$ activation function is suitable for personality traits analysis.

V. CONCLUSION

This research proposed an empirical approach for finding the best personality traits detection approach using CNN model. The main contribution of this work is to compare personality traits detection performance of sigmoid, $tanh$ and leaky relu activation functions for selecting the best one. This method consists of data preprocessing, data filtering, feature extraction, and classification phases. First, the text documents have been preprocessed to remove any unwanted characters. Then, the processed data was vectorized using word2vec algorithm. The vectorized document was fed as input in CNN where features were extracted in hidden layers. Finally, deep CNN has used for classification, and soft max function is applied as output layer to find probability of each document.

Five personality traits named as EXT, NEU, AGR, CON, and OPN have been used for comparative experimental analysis. The comparison results showed $tanh$ and leaky relu performs over sigmoid in all personality traits datasets. The average F1-score of sigmoid, $tanh$ and leaky relu are 33.11%, 47.25%, and 49.07% respectively. However, F1-score of leaky relu was high only for CON dataset and for all other $tanh$ showed better results. In a nutshell, the overall performance of hyperbolic tangent function is better than sigmoid and leaky relu. Changing in the activation functions i.e. using $tanh$ and leaky relu showed surprising improvement at the both training accuracy and testing accuracy over the sigmoid function. Hence, it implies that, for text analysis using Convolutional Neural Network other activation functions may improve the result in a big scale. Incorporating RNN for personality traits will be the future research direction of this work.

REFERENCES

[1] J. M. Digman, "Personality structure: Emergence of the five-factor model," *Annual review of psychology*, vol. 41, no. 1, pp. 417–440, 1990.
[2] J. W. Pennebaker and L. A. King, "Linguistic styles: Language use as an individual difference.," *Journal of personality and social psychology*, vol. 77, no. 6, p. 1296, 1999.
[3] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.
[4] J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. Warde-Farley, and Y. Bengio, "Theano: A cpu and gpu math compiler in python," in *Proc. 9th Python in Science Conf*, vol. 1, 2010.
[5] N. Majumder, S. Poria, A. Gelbukh, and E. Cambria, "Deep learning-based document modeling for personality detection from text," *IEEE Intelligent Systems*, vol. 32, no. 2, pp. 74–79, 2017.
[6] F. Mairesse, M. A. Walker, M. R. Mehl, and R. K. Moore, "Using linguistic cues for the automatic recognition of personality in conversation and text," *Journal of artificial intelligence research*, vol. 30, pp. 457–500, 2007.
[7] Y. Kim, "Convolutional neural networks for sentence classification," *arXiv preprint arXiv:1408.5882*, 2014.
[8] T. Mikolov et al., "Word2Vec - Google Code Archive," Google, 2013. [Online]. Available: https://code.google.com/archive/p/word2vec/.
[9] M. Coltheart, "The mrc psycholinguistic database," *The Quarterly Journal of Experimental Psychology*, vol. 33, no. 4, pp. 497–505, 1981.
[10] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word–emotion association lexicon," *Computational Intelligence*, vol. 29, no. 3, pp. 436–465, 2013.
[11] M. D. Zeiler, "Adadelta: an adaptive learning rate method," *arXiv preprint arXiv:1212.5701*, 2012.

An Empirical Study of Cervical Cancer Diagnosis using Ensemble Methods

Enamul Karim
Department of CSE
Daffodil International University
Dhaka, Bangladesh
enamul.cse@diu.edu.bd

Nafis Neehal
Department of CSE
Daffodil International University
Dhaka, Bangladesh
nafis.cse@diu.edu.bd

***Abstract*—Cervical Cancer, being one of the most pressing issues now-a-days, needs to be addressed properly. With a view to achieving an accurate diagnosis method for Cervical Cancer by screening the risk factors, different machine learning approaches have been taken over time. But by analyzing the performances of most of state-of-the-art approaches, it was inferred that there is still room for improvement by developing a more accurate model. Hence, in this paper an approach using ensemble methods with SVM as the base classifier has been taken. The ensemble method with Bagging technique achieved an accuracy of 98.12% with very high precision, recall and f-measure value.**

***Index Terms*—Ensemble methods, Bagging, Machine learning, cervical cancer, risk factors**

I. Introduction

Cervical cancer has turned out to be a major concern around the world. According to the World Health Organization (WHO), it is the fourth most common cancer in women. Cervical cancer represents almost 7.9% of the total female cancer that occurred in 2012 [1]. It arises in the cervix cells in uterus and has the potential to invade other parts of the body [2]. A number of risk factors are associated with this type of cancer. It is very pivotal to make a prediction based on the risk factors for the detection of the disease. An early identification can play a crucial role in the treatment process. Proper medical care, after an early detection can increase the chances of survival to a great extent [1]. Different stages of Cervical Cancer has been shown in Fig 1.

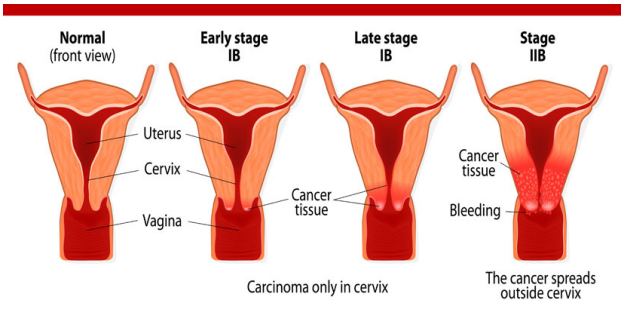


Fig. 1. Cervical Cancer [3]

In this study, different machine learning algorithms were carried out based on the risk factors in order to make a comparative analysis in the intervention of the disease. Four classification algorithms such as SVM, Decision tree, Multilayer Perceptron and K-nearest neighbors were used for the purpose in **WEKA** data mining tool. Classification accuracy was then improved by applying different ensemble methods. Bagging and boosting were performed on the data-set to acquire better accuracy.

The organization of the remaining paper is as follows - Earlier works related to cervical cancer is presented in Section II, Section III represents the overall methodology. Section IV shows the performances of the experiment done and Section V summarizes the results with a short discussion. Finally, Section VI concludes the paper with an overall discussion.

II. Literature Review

Cervical Cancer has been a red-flagged issue for a very long time now. Proper diagnosis of cervical cancer has become a very challenging and demanding domain of research. Various machine learning based attempts has been made in order to achieving a classification model with high accuracy for screening cervical cancer.

TABLE I
Literature Review

Author	Accuracy	Classifier
Kelwin et. al. (2017) [4]	68.30%	SVM
Hayder et. al. (2017) [5]	42.9%	Decision Tree
Wesabi et. al. (2018) [6]	97.5% 49.0% 88.2% 87.6%	Decision Tree Logistic Regression SVM KNN
Wen et. al. (2017) [7]	94.13% 94.03% 94.03%	SVM SVM-RFE SVM-PCA
Ramit et. al. (2017) [8]	94.92%	Random Forest
Fahri et. al. (2018) [9]	95.89% 95.89% 97.26%	kNN MLP Bayes Net

All the above mentioned papers in TABLE I have used the same dataset. Our empirical result is also based on the same dataset with 36 attributes containing the medical records of 858 patients. In [4], a transfer learning based approach has been taken. The authors proposed a strategy that facilitates target and source models to share the same coefficient signs for reducing the amount of labeled data from each expert/modality. In [5], a 3-staged cost sensitive classifier for cervical cancer screening was proposed. The classifier was a decision tree with cost selectivity and the model was evaluated with 10-fold cross validation.

In [6], diverse classification techniques on cervical cancer dataset have been applied. The result shows the advantage of feature selection approaches to the best predicting of cervical cancer disease. The results also showed that first sexual intercourse, age, hormonal contraceptives, number of pregnancies, smokes, and STDs:genital herpes are the main predictive features for cervical cancer. In [7] three SVM based approaches, namely - Simple SVM, SVM-RFE and SVM-PCA were introduced for cervical cancer classification. According to their result, SVM-PCA method was the most accurate method.

In [8] a penalty function was introduced to the already existing fitness function for Binary Firefly Algorithm. This addition radically reduced to an optimal subset which gained an increased classification accuracy compared to the traditional deep learning and information gain methods. In [9] a comparative study of the performance of popular classification algorithms like Multilayer Perceptron, BayesNet and k-Nearest Neighbor was done.

All the results from the aforementioned papers are presented in Table 1.

III. Methodology

The overall workflow had been divided into some major steps. Each and every steps are shown in Figure 2 and discussed in details afterwards.

A. Data Collection

The data-set used for the classification was collected from the UCI machine learning repository. This data-set was originally collected at 'Hospital Universitario de Caracas' in Caracas, Venezuela. The data-set comprises demographic information, habits, and historic medical records of 858 patients. Several patients decided not to answer some of the questions because of privacy concerns. So definitely there were some missing values present in the dataset which was handled in the later step.

The classifications were carried out on the medical information of 858 different patients with 36 features. Table I shows the information of the available attributes. Four attributes such as **Schiller**, **Hinselmann**, **Biopsy** and **Cytology** were selected as the class attributes for classification.

TABLE II
ATTRIBUTES [10]

Attribute Name	Type	Attribute Name	Type
Age	Integer	First sexual intercourse (age)	Integer
Number of sexual partners	Integer	STDs:molluscum contagiosum	Boolean
STDs:pelvic inflammatory disease	Boolean	STDs:genital herpes	Boolean
Number of pregnancies	Integer	STDs:AIDS	Boolean
Smokes	Boolean	Smokes (years)	Integer
Smokes (packs/year)	Integer	STDs:HPV	Boolean
STDs:HIV	Boolean	STDs:Hepatitis B	Boolean
Hormonal Contraceptives	Boolean	Hormonal Contraceptives (years)	Integer
STDs: Time since first diagnosis	Integer	STDs: Number of diagnosis	Integer
Intrauterine Devices (IUDs)	Boolean	STDs: Time since last diagnosis	Integer
IUDs (years)	Integer	Dx:Cancer	Boolean
Sexually Transmitted Diseases (STDs)	Boolean	Dx:CIN	Boolean
STDs:condylomatosis	Boolean	STDs:pelvic inflammatory Dx	Boolean
STDs (number)	Integer	Dx:HPV	Boolean
STDs:vaginal condylomatosis	Boolean	STDs:vulvo-perineal condylomatosis	Boolean
STDs:cervical condylomatosis	Boolean	STDs:syphilis	Boolean
Hinselmann	Boolean	**Biopsy**	Boolean
Schiller	Boolean	**Cytology**	Boolean

Each patient can be classified into four different classes at the same time since it is a multilabel and multiclass problem. The classifications were done on WEKA data mining tool. A cross-validation with 10 folds were conducted on the dataset to evaluate the performances of the different algorithms used for comparison.

B. Pre-processing

The dataset collected contained a lot of missing values. After analysis, it was seen that two attributes named ***STDs: Time since last diagnosis*** and ***STDs: Time since first diagnosis*** contained the highest number of missing data. So, they were removed from the dataset and classifications were done based on the other 34 attributes. The missing values of these 34 attributes were filled up by the median of the corresponding column in java.

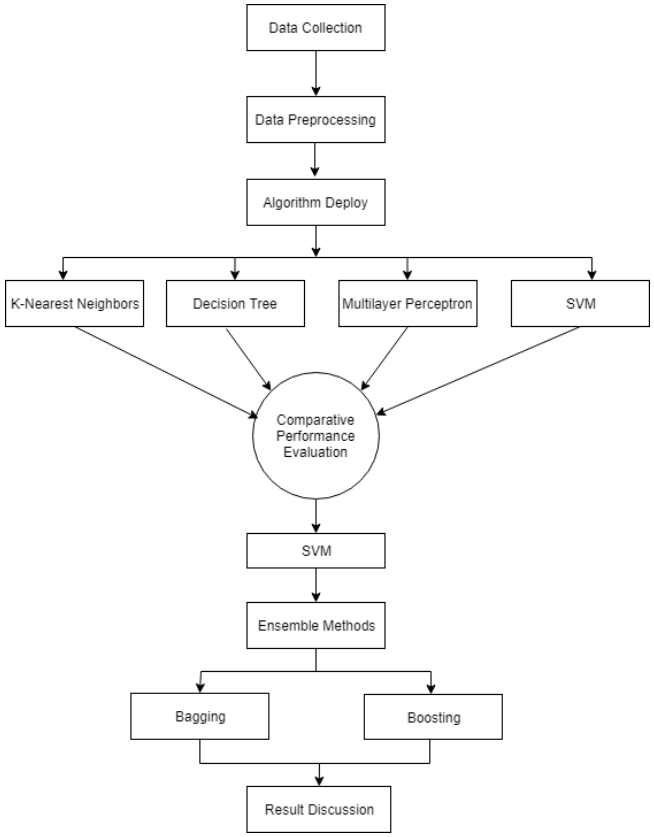


Fig. 2. Overall Work flow.

C. Classification Algorithms

Decision Tree:

Decision tree is a fast learning algorithm that enables us to classify and predict a target variable. It is a tree like structure in which every non-terminal nodes represent a condition and the terminal nodes contain class labels [5].

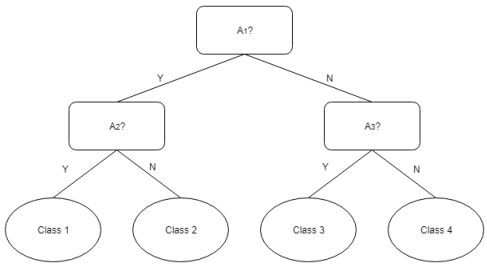


Fig. 3. Decision Tree Induction.

Multilayer Perceptron:

Multilayer Perceptron (MLP) is a type of neural network model that makes the use of supervised learning technique where the input sets are mapped into proper output sets. It consists of different layers of nodes in which one layer is connected to the next layer. An MLP typically has 3 types of layers, namely - Input Layer, Hidden Layer and Output Layer. Every node, other than the input nodes, is a processing unit called neuron that has a nonlinear activation function. In order to train the network, back propagation is typically used in this algorithm. [9]

SVM:

SUPPORT vector machine is a strong classifier for solving regression problems [11]. It has a very easy implementation. A hyperplane is chosen to separate all the points in the input variable with respect to their classes. Sequential Minimal Optimization (SMO) is one of the ways to solve the SVM problem. According to [12], SMO is fast and has good scaling properties.

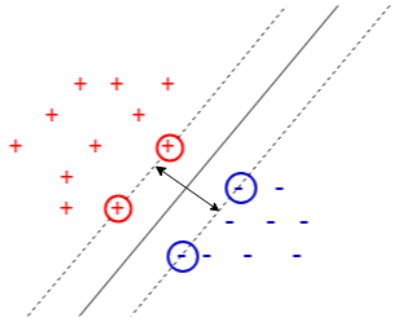


Fig. 4. Optimal hyperplane separating two classes.

K-Nearest Neighbor:

The k-NN is one of the supervised learning techniques for solving classification problems. One crucial point is that the features are to be determined in advance. According to this algorithm, the distance between the desired item to be classified and the previous items is determined and the nearest k class is taken. Identifying the number of neighbors is a very important optimization problem in this algorithm [9].

Ensemble Methods Ensemble methods are hybrid algorithms that combine different machine learning algorithms into one predictive model and decrease bias (boosting), variance (bagging) or improve predictions (stacking). Most ensemble methods use a single base learning algorithm to produce homogeneous base learners.

1) Boosting: Boosting is one of the ensemble methods which is used for improving the performances of a weak classifier. Sequential learning of the predictors is used here. Initially, the whole data set is used for learning. Subsequently, learning occurs through the training sets based on the previous performances. The weights of the misclassified instances are increased so that the possibility of appearing in the training set of the following predictor gets higher [13]

2) Bagging: Bagging is an ensemble methods which is used to improve unstable classification problems. If

the position of a training point changes marginally, weak classifiers can become unstable. Bagging can be applied to different classification algorithms. It is a very useful technique for large data sets in which detecting a good model is difficult due to the complexity and scale of the problem [13].

IV. Performance Evaluation

Performances were determined based on accuracy, precision, recall and f-measure. Table II shows the equations used for calculating the desired performance measures.

TABLE III
PERFORMANCE EVALUATION MATRIX [14]

Measure	Equation
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Recall	$\frac{TP}{TP+FN}$
Precision	$\frac{TP}{TP+FP}$
F-Measure	$\frac{2X\,Recall X\,Precision}{Recall+Precision}$

Using 10 folds cross validation, different classification algorithms were conducted on the data-set to obtain the performance measures which is represented by the following graphs.

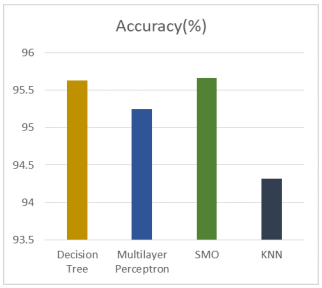


Fig. 5. Accuracy.

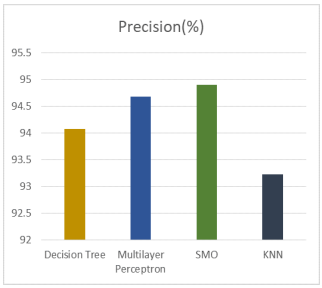


Fig. 6. Precision.

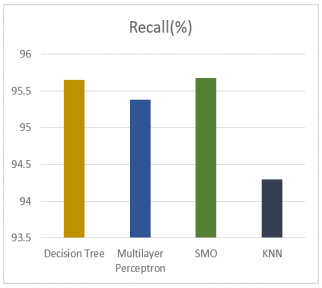


Fig. 7. Recall.

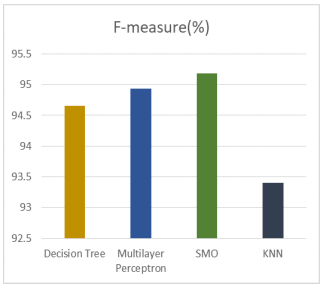


Fig. 8. F-measure.

After analyzing the results obtained, it is seen that SMO offers us the best performance. SMO is superior to all other classification algorithms used in respect of accuracy, precision, recall and f-measure. Two ensemble methods namely AdaBoostM1 and bagging were conducted using 10 fold cross validation and SMO as the classifier.

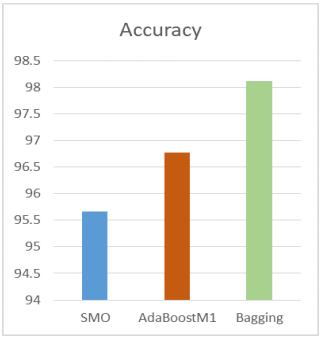


Fig. 9. Accuracy.

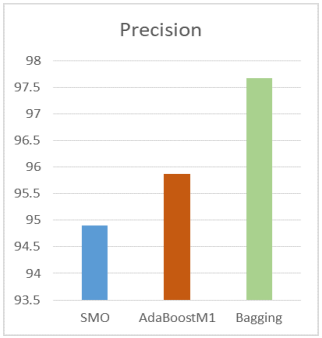


Fig. 10. Precision.

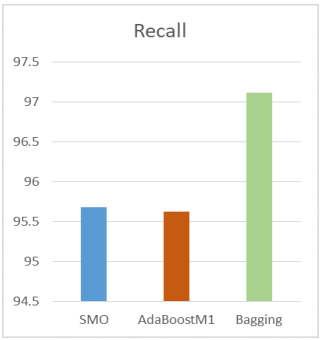


Fig. 11. Recall.

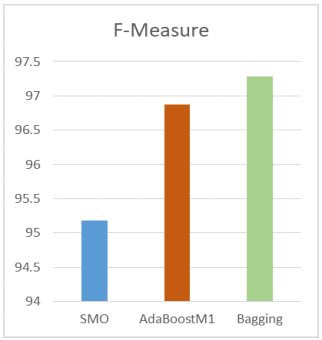


Fig. 12. F-measure.

V. Results and Discussion

Four classification algorithms were used in our study. After analyzing the performances, it was seen that SMO had a better performance than the others. So, SMO was selected as the base classifier and two ensemble methods namely AdaboostM1 and bagging were carried out to enhance the performance. AdaboostM1 had a slight increase in performance but bagging had the bigger impact.

Machine learning techniques aim to find a single model that best predicts our desired result. Instead of using a single model as the predictor, ensemble methods take into account different models and average those models to generate a final stable model.

Bagging is a powerful ensemble method which uses multiple models of the same algorithm. It reduces variance by building more advanced models of data sets. It takes different subsets from our training data sets randomly and then uses bootstrap as a subroutine of its learners. Bagging trains the learning models and then take the votes on their output. Thus, by reducing variance, it can reduce the overfitting problem leading to higher stability.

VI. Conclusion

Machine learning classification has received a great attention in the past days. Some of the popular classification algorithms were used in our study. The study was conducted among 858 patients having 36 different attributes. The performances were measured in respect of accuracy, precision, recall and f-measure. The algorithm with the best performance was selected and the ensemble methods were carried out on them to improve the overall accuracy. This accuracy can further be improved with nature inspired optimization algorithms and therefore will remain as a possible future extension for our work. Also, we could try out other non-ensemble based classification models to reach higher accuracy. Finally, this study turns out to be a very crucial one considering its health and social impact and research on this arena should be kept going on in a continued manner.

References

[1] "Cancer," Available:http://www.who.int/cancer/prevention/diagnosis-screening/cervical-cancer/en/, [Online; accessed 29-August-2018].

[2] "About cervical cancer," Available:https://www.cancer.org/cancer/cervical-cancer/about/what-is-cervical-cancer.html, [Online; accessed 29-August-2018].

[3] "Cervical cancer," https://www.womenworking.com/cervical-cancer-5-signs-never-ignore/, online; accessed: 2018-09-03.

[4] K. Fernandes, J. S. Cardoso, and J. Fernandes, "Transfer learning with partial observability applied to cervical cancer screening," in *Iberian conference on pattern recognition and image analysis*. Springer, 2017, pp. 243–250.

[5] H. K. Fatlawi, "Enhanced classification model for cervical cancer dataset based on cost sensitive classifier," *Int. J. Comput. Tech*, vol. 4, pp. 115–120, 2007.

[6] K. Barker, D. Berry, and C. Rainwater, "Classification of cervical cancer dataset."

[7] W. Wu and H. Zhou, "Data-driven diagnosis of cervical cancer with support vector machine-based approaches," *IEEE Access*, vol. 5, pp. 25 189–25 195, 2017.

[8] R. Sawhney, P. Mathur, and R. Shankar, "A firefly algorithm based wrapper-penalty feature selection method for cancer diagnosis," in *International Conference on Computational Science and Its Applications*. Springer, 2018, pp. 438–449.

[9] M. F. Unlersen, K. Sabanci, and M. Özcan, "Determining cervical cancer possibility by using machine learning methods."

[10] Z. Ceylan and E. Pekel, "Comparison of multi-label classification methods for prediagnosis of cervical cancer," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 5, no. 4, pp. 232–236, 2017.

[11] S. K. Shevade, S. S. Keerthi, C. Bhattacharyya, and K. R. K. Murthy, "Improvements to the smo algorithm for svm regression," *IEEE transactions on neural networks*, vol. 11, no. 5, pp. 1188–1193, 2000.

[12] S. S. Keerthi, S. K. Shevade, C. Bhattacharyya, and K. R. K. Murthy, "Improvements to platt's smo algorithm for svm classifier design," *Neural computation*, vol. 13, no. 3, pp. 637–649, 2001.

[13] I. Syarif, E. Zaluska, A. Prugel-Bennett, and G. Wills, "Application of bagging, boosting and stacking to intrusion detection," in *International Workshop on Machine Learning and Data Mining in Pattern Recognition*. Springer, 2012, pp. 593–602.

[14] Y. J. Huang, R. Powers, and G. T. Montelione, "Protein nmr recall, precision, and f-measure scores (rpf scores): structure quality assessment measures based on information retrieval statistics," *Journal of the American Chemical Society*, vol. 127, no. 6, pp. 1665–1674, 2005.

Bangla Document Categorisation using Multilayer Dense Neural Network with TF-IDF.

Manisha Chakraborty
dept. of Computer Science and Engineering
United International University
Dhaka, Bangladesh
mnshchakraborty19@gmail.com

Dr. Mohammad Nurul Huda
dept. of Computer Science and Engineering
United International University
Dhaka, Bangladesh
mnh@cse.uiu.ac.bd

***Abstract*—Document categorization is a classic NLP task which includes allocating documents based on their content into one or more predefined classes. In this paper, we proposed a model using multilayer dense neural network with TF-IDF as feature selection technique in terms of Bangla text document categorization. The proposed system is distributed into three consecutive steps: i) preprocessing raw text data and extracting feature using TF-IDF, ii) designing the model architecture and fitting the model to train set, and iii) evaluating model performance on test set based on accuracy and F1 score. It is observed from experiments that the proposed method exhibits higher accuracy (84.58%) and F1 score (0.84) compared to the other well-known classification algorithms (Support Vector Machine, Multinomial Naive Bayes and so on) for Bangla text document classification.**

***Index Terms*—Document categorization, TF-IDF, Neural Network, Classification.**

I. Introduction

Document categorization is a classification task in natural language processing (NLP) with a variety of real-life applications that range from sentiment analysis to spam detection. In terms of document classification tasks, supervised classification algorithms have been a popular choice among NLP practitioners. [1], [2] Furthermore, neural networks(NN) with word embeddings have accomplished remarkable results regarding text classifications in recent times. [3] However, ample amount of studies has been implemented on English in contrast to Bengali. Bangla, also known as Bengali, is one of the most widely spoken language throughout the world. Hence, with the rapid growth of internet, almost every form of digital text documents such as blogposts, newspaper articles, social media posts are noticeably escalating all over the web. Consequently, the necessity of automated document categorisation has become a critical issue to address. Various methods of Bangla document classification have been proposed in recent studies. [4]–[7] In this paper, we designed a dense neural network model with TF-IDF as feature selection technique for Bangla text document categorization. However, a comparative analysis among our proposed model and other well-established methods can aid to the better comprehension of our model's applicability in terms of Bangla document categorization. In this paper, we designed total nine experiments with two different feature selection method (TF-IDF and word embeddings) for performance evaluation.

- Six classification models with TF-IDF: Logistic Regression (LR) , Stochastic Gradient Descent classifier(SGD), K-Nearest Neighbor (KNN), Decision Tree(DT), Support Vector Machine (SVM) and Multinomial Naive Bayes (MNB).We used Multinomial Naive Bayes (MNB) instead of Naive Bayes(NB) since study [8] suggests that MNB performs better than other Naive Bayes models.
- Two neural network models with word embedding layer: Dense NN (embed +NN) and Convolutional NN (embed +CNN).
- Our proposed model: one dense neural network model with TF-IDF(TF-IDF +NN)

The purpose of these experiments is to evaluate and compare the efficiency of our model and the other models in terms of document categorization. The rest of this paper is as organised as follows: section-II discusses about related early works. Section-III depicts the design and methodology of our study. In section-IV, experimental analysis is discussed. Section-V provides the concluding remark of this study.

II. Related works

Literacy review indicates that extensive research works are conducted on English document categorization. Well-known supervised algorithms have been implemented for English text document categorization tasks. [1], [2], [8], [9] TF-IDF has been commonly used for extracting feature from text data in considerable amount of studies. [10], [11] However, word embedding has shown to achieve outstanding results in recent years. In terms of Bangla document classification, several classifiers such as Stochastic Gradient Descent classifier (SGD), K-Nearest Neighbor (KNN), Naive Bayes (NB), Decision Tree (DT), Support Vector Machine (SVM) has been implemented with TF-IDF and chi-square as feature selection methods [4], [5], [7], [12].Word embeddings with Word2Vec model has shown to accomplish excellent result in document categorisation. [4], [6]

III. Methodology

In this section, the procedure of our designed experiments is described. This process is composed of three vital steps. Firstly, we preprocess raw text data and divide the entire

dataset into train set and test set. Secondly, classifier algorithms and neural networks are applied to test set. Finally, prediction is made on test data and models are evaluated. The details of these experiments are provided in the following subsections:

A. Data Preprocessing

We collected 1600 articles from two online Bangla newspaper known as ProthomAlo (http://www.prothom-alo.com) and BDnews24 (https://bdnews24.com) by parsing Bangla texts from these newspapers. 16 types of documents priorly labeled in the mentioned newspapers are obtained. These documents are arranged into 16 classes of each having 100 articles. Afterwards, this entire dataset is divided into two sets: train set containing 1120 articles and test set with 480 articles where each class has 70 training documents and 30 test documents. Our dataset is stratified as train set should contain data from all possible classes which is not certain in random sampling. Punctuation marks, english letters, mail addresses, html tags are removed from each raw document while tokenizing. Data set description is illutrated at table-I

TABLE I
DATA SET DESCRIPTION

Classes	Total documents	Train set	Test set
Commerce	100	70	30
Science	100	70	30
Art and Literature	100	70	30
Economy	100	70	30
Education	100	70	30
Entertainment	100	70	30
Immigrant	100	70	30
Kids	100	70	30
Lifestyle	100	70	30
Movie	100	70	30
National	100	70	30
Politics	100	70	30
Sports	100	70	30
Stock	100	70	30
Technology	100	70	30
World	100	70	30

For classifier algorithms, TF-IDF vectorizer is used to extract feature. Top 1500 features with the highest term frequency and document frequency less than 0.7 that appear in more than four documents is selected during the process.

For neural networks, 5000 words with the highest word frequencies in train set are selected for feature extraction where each word is mapped to an integer number (known as token). These tokens are used in convolutional neural network(CNN) as CNN works with sequential data and TF-IDF does not interpret neighbourhood among the words. However, for multilayer dense neural network, tokens are transformed into TF-IDF vectors and both TF-IDF and word vectors are utilised with two different architectures.

Maximum number of words in an article of the train set is 2037. To work with embedding layer, it is essential for each document to have the same length. For this purpose, each document size is fixed to 2000 words by padding and truncating as needed.

B. Model Fitting

a) Classifier algorithms: Six classifier algorithms(K-NN,DT,SVM,SGD,MNB,LR) have been used on train set for hyper parameter tuning and better model adeptness. For each model, grid search is applied on train set with different sets of parameters for hyper parameter tuning. Hyper-parameter optimisation for a model is a crucial task as the performance of a model can highly depend on it's hyper-parameters. Optimal parameters are determined based on the best score of 10 fold cross validation. During the process, train set is employed to fit each model with optimal set of parameters. The diagram for classifier models is illustrated in figure-1.

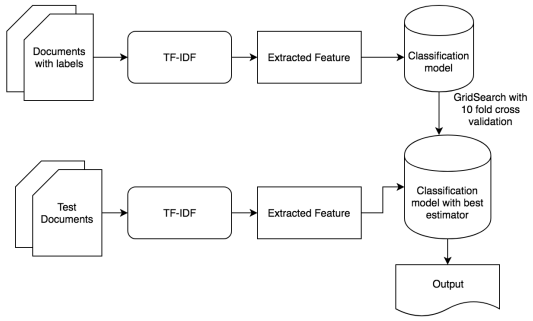


Fig. 1. Diagram for classifier models.

b) Neural Networks: In CNN model, we used embedding layer as a part of neural network in order to create word embeddings. This embedding layer learns word embeddings along with the model itself while being fitted on training data. An input document is provided as a sequence of N tokens (N=2000) to the model's embedding layer that transforms the tokens into n-dimensional word vectors (n=128). Subsequently, a one dimensional spatial dropout layer is added for regularisation in stead of dropout layer as suggested [13]. Next layer is a 1-D convolution layer followed by a GlobalMaxPool1D layer as pooling layer. Finally, a dense layer with 16 units and softmax as activation function is added as output layer.

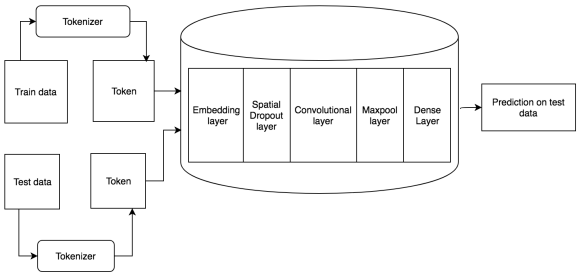


Fig. 2. Diagram for CNN with embedding layer.

In terms of dense NN, which contains embedding layer as the first layer, takes input as sequences like CNN. Next layer is a flatten layer that outputs a one dimensional tensor which

can be utilised by the following dense layer. Afterwards, a dropout layer is added. The output layer has the same structure as CNN.

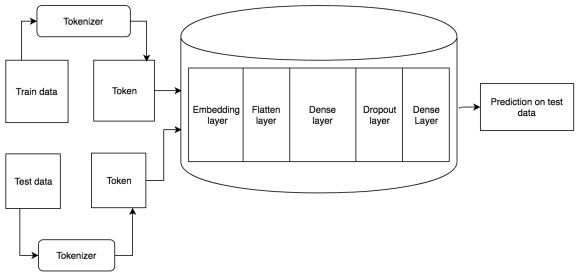


Fig. 3. Diagram for embedded Dense Neural Network

In terms of our proposed model which uses TF-IDF feature selection methods, first layer of this model is a dense layer which uses TF-IDF features as input. It is worth mentioning that this model does not include any flatten layer or spatial dropout layer due to the absence of embedding layer. Instead dropout layer is added after the first two dense layer to work as regularizer. The purpose of a dropout layer is to refine generalisation in performance by inhibiting strong correlation among activation functions. Finally, the output layer is a dense layer with 16 units and uses softmax as activation function. The diagrams of NN models are depicted in figure 2-4.

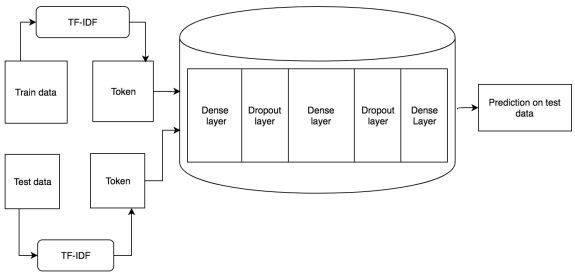


Fig. 4. Diagram for our proposed model: Dense NN with TF-IDF

C. Model evaluation

To evaluate the performance of our experimental models, we used accuracy and F1 score as performance measures. Test set is utilized for prediction on each fitted model of all 9 experiments. Predicted labels and true labels of test set are used to quantify accuracy and F1 score.

IV. Empirical analysis

a) Data collection and preprocessing: 1600 articles of 16 priorly labeled categories are parsed from two online Bangla newspaper (ProthomAlo, BDnews24) and grouped into 16 classes with 100 articles per class. The dataset is divided into train set and test set where test size parameter is 0.3. For TF-IDF used in classifier models, max df is 0.7 and min df is 4. For NN, we used keras.preprocessing.text.tokenizer tool for feature selection where number of maximum word is set to 5000. For NN models with embedding layer, we fixed the size of each article's word length to 2000 words by pre-padding or pre-truncating.

b) Experimental Setup: For classifier models, we used sets of parameters for hyper-parameter tuning. For DT, entropy criterion, maximum depth of 10 and minimum sample leaf 1 are selected as the best parameters through grid search with 10 fold cross validation. In SVM, Rbf kernel, C=10 and gamma value 0.1 are selected as the optimal parameters. Multinomial LR with newton-cg solver is chosen through grid-search. KNN with 15 neighbours, p value 2 and distance weight function is used for prediction. In terms of MNB, alpha value 0.1 is chosen. With regards to SGD, alpha value 0.0001 and log as loss function is selected through hyper parameter optimization. For neural network models, we used keras for all NN models and built them with sequential model which allows to create models layer-by-layer. For our proposed model, first dense layer has 128 node and relu as activation function. Second dense layer has 64 nodes with sigmoid activation function. After each dense layer, a dropout layer is added with value 0.5 as recommended [3]. For CNN and embed + Dense models, embedding layer takes 5000 word tokens as vocabulary and each token is represented in 128 dimensions (n=128) and maximum length of each token sequence is 2000. For CNN, Spatial dropout layer has been added with drop out value of 0.2. In 1 dimensional convolution layer, 256 different filters of kernel size 3 is added with relu activation function followed by a 1-D global max pool layer. For dense NN, flatten layer is added followed by the embedding layer. Next layer is a dense layer with 128 unit and sigmoid as activation function followed by a dropout layer of 0.5 to reduce overfitting. For all NN models, output layer is a dense layer with 16 units and softmax activation function. During model compilation, adam optimizer is used [14] and categorical cross entropy as loss function with accuracy and F1 score as metric. While fitting the model, validation split=0.3 and batch size=128 is used and all models are trained for 25 epochs.

c) Result: For all models, prediction on test dataset is conducted and confusion matrix is evaluated. Confusion matrix provides information about how many documents of a class on test set is correctly classified or misclassified. Afterwards, accuracy and F1 score is calculated from confusion matrix.The Confusion matrix of our proposed model is illustrated below in figure-5. Accuracy score and F1 score are depicted for all models in table -I.

From table-I, it is evident that our proposed model has the highest accuracy (84.580%) and F1 score (0.84). Afterwards, there is SGDclassifier and SVM with the same F1 score(0. 82) depicting slight difference in accuracy (82.290 % and 81.870 % respectively). LR and MNB performances are almost similar with accuracy 80.830% (LR) and 80.620% (MNB), and their F1 scores are 0.81 and 0.80 respectively. Next is KNN with F1 score of 0.76 which is slightly higher than the F1 score of CNN (0.75) and their accuracy is 74.375% and 75.0% respectively. Lastly, DenseNN + embedding and DT has lower

TABLE II
EVALUATION METRICS

Algorithms	F1 Score	Precision	Recall	Accuracy (%)
TF-IDF+DenseNN	0.84	0.86	0.85	84.580
SVM	0.82	0.82	0.82	81.870
SGD	0.82	0.83	0.82	82.890
LR	0.81	0.81	0.81	80.830
MNB	0.80	0.81	0.81	80.620
KNN	0.76	0.82	0.74	74.375
Embedding+CNN	0.75	0.76	0.75	75.000
Embedding+DenseNN	0.57	0.62	0.59	58.750
DT	0.53	0.59	0.52	51.670

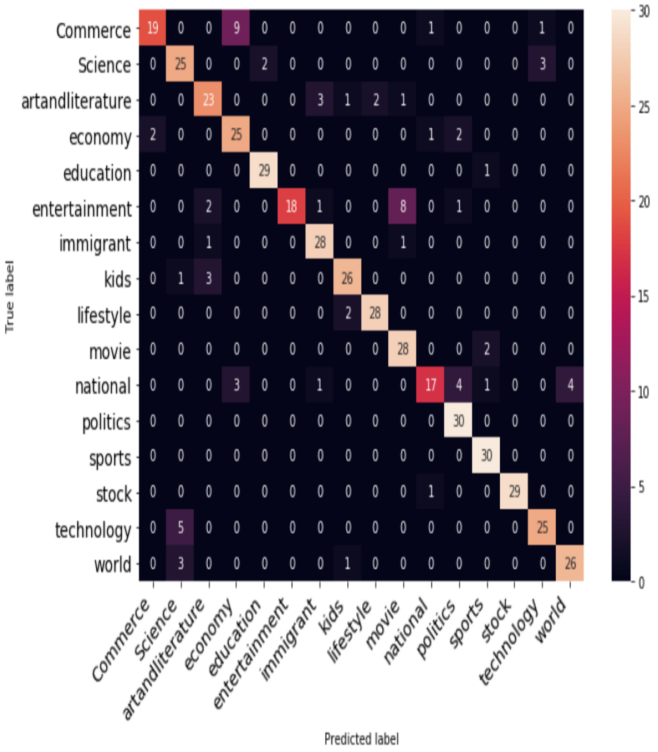


Fig. 5. Confusion matrix of our proposed model: TF-IDF+DenseNN

accuracy and F1 score than the other models where accuracy is 58.750% and 51.670% respectively and F1 score is 0.57 and 0.53 respectively.

V. CONCLUSION

In this paper, we represent a method to categorise Bangla text document by integrating TF-IDF feature selection technique with dense neural network. This paper concludes with the following verdicts:

- Our proposed system has accomplished higher accuracy (84.580%) and F1 score (0.84).
- Our proposed method which incorporates TF-IDF has superior performance than word embedding based NN models.

In future, the authors would like to implement hybrid approach with deep learning on larger datasets regarding Bangla text document categorization.

REFERENCES

[1] S. Tong and D. Koller, "Support vector machine active learning with applications to text classification," *The Journal of Machine Learning Research*, vol. 2, pp. 45–66, 12 2001.

[2] D. D. Lewis and M. Ringuette, "A comparison of two learning algorithms for text categorization," in *Third annual symposium on document analysis and information retrieval*, vol. 33, 1994, pp. 81–93.

[3] Y. Kim, "Convolutional neural networks for sentence classification," *arXiv preprint arXiv:1408.5882*, 2014.

[4] M. R. Hossain and M. M. Hoque, "Automatic bengali document categorization based on word embedding and statistical learning approaches," in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*. IEEE, 2018, pp. 1–6.

[5] F. Kabir, S. Siddique, M. R. A. Kotwal, and M. N. Huda, "Bangla text document categorization using stochastic gradient descent (sgd) classifier," in *2015 International Conference on Cognitive Computing and Information Processing (CCIP)*. IEEE, 2015, pp. 1–4.

[6] A. Ahmad and M. R. Amin, "Bengali word embeddings and it's application in solving document classification problem," in *2016 19th International Conference on Computer and Information Technology (ICCIT)*. IEEE, 2016, pp. 425–430.

[7] M. Islam, F. E. M. Jubayer, S. I. Ahmed *et al.*, "A comparative study on different types of approaches to bengali document categorization," *arXiv preprint arXiv:1701.08694*, 2017.

[8] A. McCallum, K. Nigam *et al.*, "A comparison of event models for naive bayes text classification," in *AAAI-98 workshop on learning for text categorization*, vol. 752, no. 1. Citeseer, 1998, pp. 41–48.

[9] and A. K. Jain, "Classification of text documents," in *Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No.98EX170)*, vol. 2, Aug 1998, pp. 1295–1297 vol.2.

[10] T. Bruno, M. Sasa, and D. Donko, "Knn with tf-idf based framework for text categorization," vol. 69, 11 2013.

[11] M. Lan, C.-L. Tan, H.-B. Low, and S.-Y. Sung, "A comprehensive comparative study on term weighting schemes for text categorization with support vector machines," in *Special Interest Tracks and Posters of the 14th International Conference on World Wide Web*, ser. WWW '05. New York, NY, USA: ACM, 2005, pp. 1032–1033. [Online]. Available: http://doi.acm.org/10.1145/1062745.1062854

[12] A. K. Mandal and R. Sen, "Supervised learning methods for bangla web document categorization," *arXiv preprint arXiv:1410.2045*, 2014.

[13] B. Shu, F. Ren, and Y. Bao, "Investigating lstm with k-max pooling for text classification," in *2018 11th International Conference on Intelligent Computation Technology and Automation (ICICTA)*. IEEE, 2018, pp. 31–34.

[14] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

Neural Network based Software Defect Prediction using Genetic Algorithm and Particle Swarm Optimization

Safial Islam Ayon
Department of Computer Science and Engineering
Khulna University of Engineering & Technology, Khulna-9203, Bangladesh
safialislam302@gmail.com

Abstract — **In the arena of software engineering, software defects prediction is one of the most attractive research topics. Here the main task is to predict if there is any bug in the software or not. For software testing, software defect detection is important for reducing the time and resources consumed. Accurate estimate of defect software prediction process enables effective discovery and identification of the defects. Such prediction methods are important for the big scale systems, where verification specialists need to focus their attention. In this paper, we proposed a method where the features are selected using Genetic Algorithm (GA). Secondly, make cluster of the selected features using Particle Swarm Optimization (PSO) and then train the model with different Neural Network (NN) methods such as: Feedforward Neural Network (FNN), Recurrent Neural Network (RNN), Artificial Neural Network (ANN), and Deep Neural Network (DNN) and finally calculate accuracy, sensitivity, specificity, precision, negative prediction value, F1 score, and Matthews correlation coefficient. We use five different datasets from NASA promise software engineering repository for our study. From our study, we get the best accuracy result using deep neural network. Experimental consequences show that proposed strategy is a decent technique to predict the software defects.**

Index Terms*— *Software Defect Prediction, Neural Network, Genetic Algorithm, Particle Swarm Optimization

I. INTRODUCTION

As software company grows in a number nowadays, the complexity and prediction of software defect plays an important role in the software industry. Build high-quality software, defect prediction plays an important role to consider the quality level of the software. Also, it helps for software testing and debugging. For large software, manually find the defects is very complex and time-consuming. Because there are several source files and many logic connections among those files. Quality of the software is very important and different issues which are related to testing software are becoming crucial when the defects are prediction. Code defects consider the key reasons for failures to build a software [1], so to improve consistency many companies like Google, employ code review and test the code to find bugs [2]. So, an automatic system is needed to find out the defect in a quick possible time [3]. This not only saves time but also ensure to build high-quality software. This type of system also gives information about the software like code complexity, time complexity. It helps the code reviewers to solve the defect in minimum time.

Software defect prediction is very challenging tasks in the software engineering sector. To detect the defect in manually a lot of manpower and times are needed. Also handle this huge data in manually is also very difficult. So, different types of computer-based automatic system are very effective to predict the defect in the software. In recent years, different methods have been proposed to detect a software contains any defect or not. The traditional approaches focus on designing structures of programs [4 – 6]. Then different machine learning techniques are proposed to build the models. Some of them use Neural Networks (NN) [7], Support Vector Machine [8], Transfer Learning [9] and etc. methods.

Artificial Neural Network (ANN) makes using the basic structure of the human brain, where the neurons receive input, process the input, and finally produce the output according to the processing steps. Feedforward Neural Network (FNN) is the first and simple type of ANN devised. Here the information moves only one direction i.e. forward. Here also have input state, one hidden layer, and output state like ANN. Recurrent Neural Network (RNN) is also a modification of ANN, where the connections of graphs between the nodes are temporary. RNN is used in two broad network classes of similar structure. Deep Neural Network (DNN) is a fast-growing method and it works like a human brain. Many studies work to prove that DNN provides a better outcome, decline the classification error rate and more robust to noise than other strategies.

In this paper, we proposed a method to predict the defect of software using Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). GA is a method which impersonators the way biological evolution works. Mainly GA is an optimization algorithm but nowadays GA uses in different sectors in engineering. PSO is also one kind of optimization technique which is inspired by nature. In recent year it becomes very powerful and interesting search technique. It also modified to solve different engineering problems. In the proposed method, GA is used to select the features and PSO is used for clustering. Then the datasets are trained used four neural network techniques artificial neural network, feedforward neural network, recurrent neural network, and deep neural network. Then using the test set measure different performance measure parameters like accuracy, sensitivity, specificity, precision, negative prediction value, F1 score, and Matthews correlation coefficient.

Rest of the paper has been prepared as follows. Section II provides a short view of the researchers that have been done in this field. Section III provides the proposed methodology. Experimental outcomes of the proposed strategy are described in Section IV. Finally, Section V presents a brief conclusion of the paper.

II. Related works

X. Yang et al. [10] used deep learning methods to predict the defects. They use Deep Belief Network to extract the features from the initial set of linear features. Then they use Logistic Regression technique to classify the software contain any bug or not. S. Wang et al. [11] proposed a method to extract the features from the given set of data using deep learning algorithm and then applied Deep Belief Network to classify the defect and non-defect software.

T. Menzies et al. [12] focus on how the attributes are used for predicting defect rather than the attributes are used. They proved that this type of prediction is useful. They got 71% of mean probability detection and 25% of mean false alarms rates. Zhong et al. [13] make the cluster using k-means and Neural-Gas techniques. Neural-Gas method works better in the terms of mean squared error and the k-means method works much faster compare to other methods. N. Sharma et al [14] compare various algorithm for clustering the ISBSG and PROMISE dataset using WEKA. There are also some traditional statically methods [15 – 17] to predict the defect of software.

III. Methodology

This section discusses the proposed methodology for predicting the defect in the software. In this paper, we proposed a method using Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) to predict the defect of software. We follow 5 steps, (1) Dataset Collection, (2) Feature selection, (3) Build Cluster, (4) Train the model, and (5) Performance measurement. Figure 1 describes the overall system for software defect. The following subsections explain all the steps in details.

A. Dataset Collection

We consider 5 different datasets to predict the defect of software. The dataset is taken from NASA promise software engineering repository [18] namely PC1, CM1, JM1, KC1, and KC2. There are no missing values and have 22 attributes in all the datasets. If the software contains defect then the final attributes were true or 1and have not any defect then the value was false or 0.

B. Feature Selection

Using Genetic Algorithm (GA) the features are selected for this paper. There are two operators namely crossover and mutation help to make a new population from the current population. Crossover exchanges a part of chromosomes to another chromosomes and mutation randomly change some allele values in the chromosome. GA calculate the fitness of every individual chromosome and chose the best fitness value for the next generation. The process of crossover and mutation are continuing until the requirement are not meet.

C. Build Cluster

To build cluster using Particle Swarm Optimization (PSO), a single particle represents the cluster centroid point. Each particle represents a clustering solution. Each particle maintains a matrix $X_i = (C_1, C_2, \ldots, C_i, \ldots C_k)$, where C_i in the *ith* cluster centroid vector and k is the number of clusters. Then randomly taken n centroid points. For each centroid point calculate the Euclidean distance, here the distance is considering the fitness value. For this paper we take the value of $n = 5$, i.e. for the given data, there are 5 clusters.

D. Train the Model

In the proposed method, we train the model 4 using different neural network methods: Feedforward Neural Network (FNN), Recurrent Neural Network (RNN), Artificial Neural Network (ANN), and Deep Neural Network (DNN). For training the dataset we divide the dataset into two parts. (1) train set and (2) test set. The ratio of divide the dataset in all the dataset is 75% for training and 25% for testing.

E. Performance Measurement

After the training the model, using the test set we measure some performance measurement parameters like Accuracy, Sensitivity, Specificity, Precision, Negative Prediction Value (NPV), F1 score, and Matthews Correlation Coefficient (MCC). In the confusion matrix, there are four classification performance indices are present. The indices are (1) True Positive (TP), (2) True Negative (TN), (3) False Positive (FP), and (4) False Negative (FN).

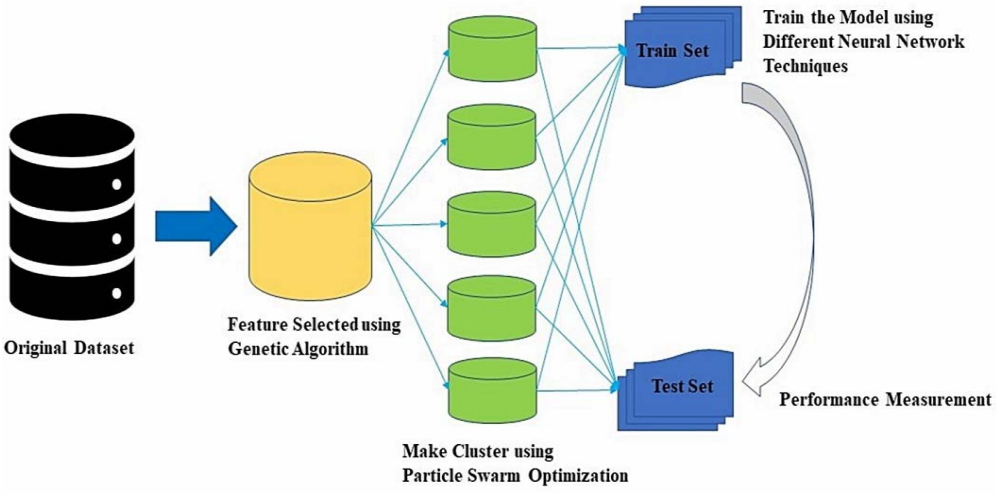


Figure 1. Full System for predicting software defect.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FN+FP)} \quad (1)$$

$$Sensitivity = \frac{TP}{(TP+FN)} \quad (2)$$

$$Specificity = \frac{TN}{(TN+FP)} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Negative\ Predictive\ Value\ (NPV) = \frac{TN}{FN + TN} \quad (5)$$

$$F1\ Score = \frac{2TP}{(2TP+FP+FN)} \quad (6)$$

$$Matthews\ Correlation\ Coefficient\ (MCC) = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (7)$$

IV. Result And Analysis

This section discusses the result and compares the performance with other existing methods.

A. Experimental Setup

All the experiment was implemented on Python 3.5.4. All the program was run on HP ProBook 450G laptop computer with configuration Intel® Core ™ i5-4200M, CPU @ 3.60 GHz, RAM 8 GB with Windows10.

B. Results Analysis

In this paper, first the features were selected using Genetic Algorithm (GA) from the given dataset. Then using Particle Swarm Optimization (PSO) made cluster. Finally train the model using different neural network techniques: Feedforward Neural Network (FNN), Recurrent Neural Network (RNN), Artificial Neural Network (ANN), and Deep Neural Network (DNN). We used five datasets namely: PC1, CM1, JM1, KC1, and KC2. All the datasets were partitioned into 10-fold cross-validation. The advantage of the cross-validation method was to observations for both training and testing, and each observation was used for the test set exactly once. The ratio of divide the dataset in all the cases was 75% for training and 25% for testing.

Using the confusion matrix accuracy, sensitivity, specificity, precision, negative predictive value, F1 score, and matthews correlation coefficient were measured. Results were shown in Table I, Table II, Table III, Table IV, and Table V for PC1, CM1, JM1, KC1, and KC2 dataset respectively. Best accuracy result found in Table I, Table II, Table III, Table IV, and Table V was 98.27%, 97.59%, 98.47%, 92.65%, and 91.54% respectively. All the cases deep neural network gave the best accuracy result. Other performance measurement parameters: sensitivity, specificity, precision, negative predictive value, F1 score, and matthews correlation coefficient were also shown in Table I, Table II, Table III, Table IV, and Table V. The result of matthews correlation coefficient was above 93% for JM1 dataset, which was very near to the best value of matthews correlation coefficient. Also, the F1 score was above 98% for PC1 and JM1 dataset, which was also near the best value 1.

TABLE I: EVALUATION METRICS FOR PC1 DATASET

Evaluation Metrics	FNN	RNN	ANN	DNN
Accuracy (%)	97.11	97.57	97.20	98.27
Sensitivity (%)	97.87	98.26	98.16	98.89
Specificity (%)	87.01	88.31	84.42	89.61
Precision (%)	99.02	99.12	98.83	99.25
Negative Predictive Value (%)	75.28	79.07	77.38	85.19
F1 Score	0.98	0.99	0.98	0.99
Matthews Correlation Coefficient	0.79	0.82	0.79	0.86

TABLE II: EVALUATION METRICS FOR CM1 DATASET

Evaluation Metrics	FNN	RNN	ANN	DNN
Accuracy (%)	95.38	95.78	96.79	97.59
Sensitivity (%)	79.59	81.63	85.71	89.80
Specificity (%)	97.10	97.33	97.99	98.44
Precision (%)	75.00	76.92	82.35	86.27
Negative Predictive Value (%)	97.76	97.98	98.43	98.88
F1 Score	0.77	0.79	0.84	0.88
Matthews Correlation Coefficient	0.75	0.77	0.82	0.87

TABLE III: EVALUATION METRICS FOR JM1 DATASET

Evaluation Metrics	FNN	RNN	ANN	DNN
Accuracy (%)	97.93	98.39	98.12	98.47
Sensitivity (%)	98.85	99.20	98.99	99.25
Specificity (%)	94.11	95.01	94.49	95.20
Precision (%)	98.59	98.81	98.68	98.85
Negative Predictive Value (%)	95.15	96.62	95.77	96.81
F1 Score	0.99	0.99	0.99	0.99
Matthews Correlation Coefficient	0.93	0.95	0.94	0.95

TABLE IV: EVALUATION METRICS FOR KC1 DATASET

Evaluation Metrics	FNN	RNN	ANN	DNN
Accuracy (%)	91.51	92.41	92.56	92.65
Sensitivity (%)	92.02	97.24	96.32	94.48
Specificity (%)	91.42	91.53	91.87	92.32
Precision (%)	66.23	67.73	68.41	69.21
Negative Predictive Value (%)	98.43	99.45	99.27	98.92
F1 Score	0.77	0.80	0.80	0.80
Matthews Correlation Coefficient	0.73	0.77	0.77	0.77

TABLE V: EVALUATION METRICS FOR KC2 DATASET

Evaluation Metrics	FNN	RNN	ANN	DNN
Accuracy (%)	90.38	89.23	89.62	91.54
Sensitivity (%)	78.10	75.24	71.43	79.05
Specificity (%)	93.50	92.77	94.22	94.70
Precision (%)	75.23	72.48	75.76	79.05
Negative Predictive Value (%)	94.40	93.67	92.87	94.70
F1 Score	0.77	0.74	0.74	0.79
Matthews Correlation Coefficient	0.71	0.67	0.67	0.74

TABLE VI: COMPARISON OF THE PROPOSED METHOD TO OTHER EXISTING METHODS

Authors	Methods	Accuracy (%)
G. Abaei and A. Selamat [19]	Decision Tree	79.50
	Random forest	81.14
	Naïve Bayes	80.42
	Immunos2	80.65
	CSCA	80.17
A. Chug and A. Dhall [20]	Decision Tree	99.36
	Naïve Bayes	94.11
	Random Forest	99.55
M.M. R. Henein et al. [21]	K-means + ANN	80.40
A.A. Abdou and N.R. Darwish [22]	Naïve Bayes	82.44
	SVM	91.74
	Random forest	93.40
	Logistic	87.40
	Decision tree	89.84
Proposed Method	**FNN**	**97.93**
	RNN	**98.39**
	ANN	**98.12**
	DNN	**98.47**

C. Comparative Analysis

Table VI shows the comparative study to our model and other existing models. G. Abaei and A. Selamat [19] survey the software defect prediction using many techniques like decision tree, random forest, naïve bayes, neural network etc. methods. A. Chug and A. Dhall [20] using supervised and unsupervised algorithm to predict the defect software. They used simple k-means cluster, hierarchical cluster, and make density-based cluster to make clustered. M.M. R. Henein et al. [21] proposed a method namely cluster based under sampling. They used k-means algorithm to make cluster and artificial neural network to predict the model. A.A. Abdou and N.R. Darwish [22] use different machine learning techniques to predict the software defects.

V. CONCLUSIONS

Software defects prediction is an important task in the sector of software engineering. Early software defect prediction will help the developer to build the software in a short time. This paper proposed a new technique to predict defects in software. In the proposed method the features are selected using Genetic Algorithm and further using this data cluster are make using Particle Swarm Optimization. Finally used different neural network methods to predict the defects. To predict the defect, we use various dataset namely PC1, CM1, JM1, KC1, and KC2. From the experimental result, we see that deep neural network works better than the other methods. We get the accuracy of above 90% in all the dataset using deep neural network. Other three methods feedforward neural network, recurrent neural network, and artificial neural network also works well and give good accuracy. In the future, the same techniques will be applied to other sectors to solve different engineering problems.

REFERENCES

[1] G. Liu, E. Musial, and M.H. Chen, "Progressive Reliability Forecasting of Service Oriented Software," Proc. of the International Conference on Web Services, 2011.

[2] Bug prediction at google. [Online]. Available: http://googleengtools.blogspot.hk/

[3] T. Menzies, Z. Milton, B. Turhan, B. Cukic, Y. Jiang, and A. Bener, "Defect Prediction from Static Code Features: Current Results, Limitations, New Approaches," Automated Software Engineering, vol. 17, pp. 375–407, 2010.

[4] S.R. Chidamber and C.F Kemerer, "A metrics suite for object-oriented design," IEEE Transactions on software engineering, vol. 20, pp. 476–493, 1994.

[5] M.H. Halstead, "Elements of Software Science," vol. 7, Elsevier, 1977.

[6] T.J. McCabe, "A complexity measures," IEEE Transactions on software Engineering, vol. 4, pp. 308–320, 1976.

[7] I. Gondra, "Applying Machine Learning to Software Fault-proneness Prediction," Journal of Systems and Software, vol. 81, pp. 186–195, 2008.

[8] K.O. Elish and M.O. Elish, "Predicting defect-prone software modules using support vector machines," Journal of Systems and Software, vol. 81, pp. 649–660, 2008.

[9] Y. Ma, G. Luo, X. Zeng, and A. Chen, "Transfer learning for cross-company software defect prediction," Information and Software Technology, vol. 54, pp. 248–256, 2012.

[10] X. Yang, D. Lo, X. Xia, Y. Zhang, and J. Sun, "Deep learning for just-in-time defect prediction," Proc. of the International Conference on Software Quality, Reliability and Security, pp. 17 – 26, 2015.

[11] S. Wang, T. Liu, and L. Tan, "Automatically learning semantic features for defect prediction," Proc. of the International Conference on Software Engineering, pp. 297-308, May 14 - 22, 2016.

[12] T. Menzies, J. Greenwald, and A. Frank, "Data mining static code attributes to learn defect predictors," IEEE Transactions on Software Engineering, vol. 33, pp. 2–13, January 2007.

[13] S. Zhong, T.M. Khoshgoftaar and N. Seliya, "Analyzing Software Measurement Data with Clustering Techniques," IEEE Intelligent Systems, vol.19, pp.20-27, April 2004.

[14] N. Sharma, A.Bajpai , R. Litoriya "Comparison the various clustering algorithms of WEKA tools," International Journal of Emerging Technology and Advanced Engineering, vol. 2, issue 5, May 2012.

[15] N. Ohlsson, M. Zhao, and M. Helander, M, "Application of multivariate analysis for software defect prediction," Software Quality Journal, vol.7, pp.51-66, 1998.

[16] K.E. Emam and W. Melo, "The Prediction of Defect Classes Using Object-Oriented Design Metrics," Technical report: NRC 43609, 1999.

[17] M.H. Tang, M.H. Kao, and M.H. Chen, "An empirical study on object-oriented metrics," In Proceedings of Metrics, pp. 242-249.

[18] Promise Software Engineering Repository, Public dataset, http://promise.site.uottawa.ca/SERepository/datasetspage.html

[19] G. Abaei and A. Selamat, "A survey on software fault detection based on different prediction approaches," Vietnam Journal of Computer Science, vol. 1, pp. 79 – 95, May 2014.

[20] A. Chug and A. Dhall, "Software Defect Prediction Using Supervised Learning Algorithm and Unsupervised Learning Algorithm," Next Generation Information Technology Summit, January 2013.

[21] M. M. R. Henein, D. M. Shawky, and S. K. Abd-El-Hafiz, "Clustering-based Under-sampling for Software Defect Prediction," 13th International Conference on Software Technologies (ICSOFT), pp. 185 – 193, 2018.

[22] A.A. Abdou and N.R. Darwish, "Early Prediction of Software Defect using Ensemble Learning: A Comparative Study" International Journal of Computer Applications, vol. 179, no. 46, June 2018.

Artificial Neural Network Inducement for Enhancement of Cloud Computing Security

Md. Raihan Uddin
Dept. of Electronics and Telecommunication Engineering
Daffodil International University
Dhaka, Bangladesh
mkraihan13@gmail.com

Kh. Mohaimenul Kabir
Dept. of Computer Science and Engineering
Dhaka International University
Dhaka, Bangladesh
skabir.diu@gmail.com

Md. Taslim Arefin
Dept. of Electronics and Telecommunication Engineering
Daffodil International University
Dhaka, Bangladesh
arefin@daffodilvarsity.edu.bd

***Abstract*— Cloud computing is a spirited topic in this epoch where IOT introducing the modern lifestyle. These embrace cost-effectiveness, time savings and actual arrangement of computing properties. Conversely, in Cloud Computing privacy and security is a big issue where there are many kinds of thread and attack are working rapidly. In the context, the best cryptographic method AES algorithm and AI-based method ANN have chosen to work for the enhancement of the security where cloud computing demands the highest priority. In this case, this paper is proposing the highly advanced two-step security layer for cloud computing. Primarily, AES provides for the first layer security of the first stratum. In the AES algorithm, the execution is dependent on the key size of the algorithm when the number of rounds to be accomplished. A MATLAB code is developed for plaintext encryption and ciphertext decryption. Experiments are conducted to measure execution times. ANN is a method of computation which is biologically stimulated. These only look like the parallel computation generated which are the basics of human learning by the biological neural network. Iris and Finger recognition is implemented using MATLAB through ANN. The paper demonstrated the great proficiency of the proposed system for the user purpose.**

Keywords— Cloud Computing Security, Artificial Neural Network, Advanced Encryption Standard, Iris Recognition, Fingerprint Recognition

I. INTRODUCTION

The world is more conscious about enriching the security issue. Cloud computing is the most burning topic to ensure the security of it. To do that many algorithms applied to the platform. Many of them are secured but not fast, many of them fast but not secured. So the research is continued till now on this security system. The proposed method of the paper is to increase both fast and secure algorithm.

In 1997, the National Institute of Standards and Technology decided to organize a competition to find out the successor of DES algorithm. Where the Rijndael algorithm (designed by Vincent Rijmen and Joan Daemen in 1977) had won the competition in April 2000 and would, later on, become the AES. The cryptographic community was assured the algorithm has been thoroughly tested and screened by the NIST to still be secure even 2030 [1].

AES (Advanced Encryption Standard) is the advanced, fast, latest and most secure system at the current time. The Fast Encryption system is too far to meet with the encryption time of conventional AES algorithm. Like as, most valuable research topic become the high-performance computing capability of graphics processing unit [2]. The security system used on the first layer of the proposed method is to encryption the traditional login system with a password. AES operates on a 4 × 4 column-major order matrix of bytes, designated by the state, although some versions of Rijndael have a larger block size and have extra columns in the state. Most AES calculations are done in a particular finite field.

The second layer is based on Iris recognition and Fingerprint based on biometric tools. Iris recognition is a preset method of biometric identification. It uses mathematical pattern-recognition techniques on images of one or both of the Irises of an entity's eyes. where, patterns are unique, stable, and can be understood from some distance. A biometric system is designed for automatic authentication over the special characteristic option of the human body. The most reliable and proficient Biometric system is Iris recognition. There are some parts; spherical Iris and pupil region, occluding eyelids and eyelashes and reflections of Iris are extremely automatic segmentation where the Hough transform and is capable to localize. [3]. So this will help the system to make more efficient to the best security system. In the method firstly the system will capture the Iris to register for the future login. In the second verification of login, there will the first option for verifying the identity.

On the other hand, if anyone is unable to do Iris scan, there is another option to verify is Biometric Fingerprint scanning. The Biometric authentication of the user provides some facility like as fast, precise, easy to use, trustworthy and economical where¬ this is conventional knowledge-based and token-based methods. [4]

II. RELATED WORKS

In this proposed work we introduce the security enhancement of user side access of cloud computing. There are well known and developed algorithms AES for the initial step to encrypt and decrypt password and another one is Artificial Neural Network (ANN) using for implementing the Fingerprint and Irish recognition.

In [5], proposed a method based on neural networks for detection and classification of underwater mines using sonar images is proposed. The paper shows the success of classification rate using neural networks.

Likewise, in [6] shows neural network system with supervised and unsupervised learning is used, so that the presentations among these are associated, shown same spectral data in input.

In Y. LeCun et al [7] they recognize handwrite using convolution neural network, which provides a good result. And another one is similarly P. Vincent [8], they proposed a method for show the advantage and stacked de-noising auto encoders model in handwriting recognition. Although there are a variety of deep structure models [9], they all gain very

good results in many fields, such as visual [10], natural language processing [11]and signal processing and so on. So we believe that the method based on depth neural network can also achieve good results for the fingerprint classification.

Bodade and Talbar [12] proposed a method for recognizing different kinds of iris images, spoofs and eyes on video. This paper shows 3 categories to spoof detection are the reflectance method, the pupil dynamics method and the frequency spectrum method. They achieve a mishmash of the methods is needed and shows some outcomes of iris images recognition.

Daugman approaches cosmetic contact lenses recognition [13]. This paper shows the working procedure for lenses where dot-matrix type technology is used for manufacturing. However, there are different types of approaches for manufacture the contact lenses and dot-matrix type lenses also can be failed as if they printed resolution of the iris sensor more than the camera resolution.

III. METHODOLOGY

Here, the paper will design to express the real-life strong method to secure cloud computing technology. This system will narrate in several sections includes encryption, the process of registration and verification with the provided database. This system is shown in some flowcharts and diagram.

A. Initial Encryption and Decryption:

The basic level of the providing of cloud computing security is designed with Advanced Encryption Security (AES). Which one is the existing conventional method for encryption and decryption work on the first layer of our proposed method? Password and Username will be stored in the database when the user will provide.

B. User Registration procedure:

The user will get in this process after completing the initial step. This layer is designed with 2 option where the biometric system Iris recognition and Fingerprint recognition will be worked for registration. The explanation and working procedure are briefly narrating with each step include Figure-1:

Step-1: The Encrypted password will be stored in the database. Successfully completing this, there will be 2 option to choose as a process of registration.

Option-1: The Neural Network algorithm will be applied to recognize the Iris. Here, our system will take some different position, a movement image of Iris. Which will be saved in the Database of the root user.

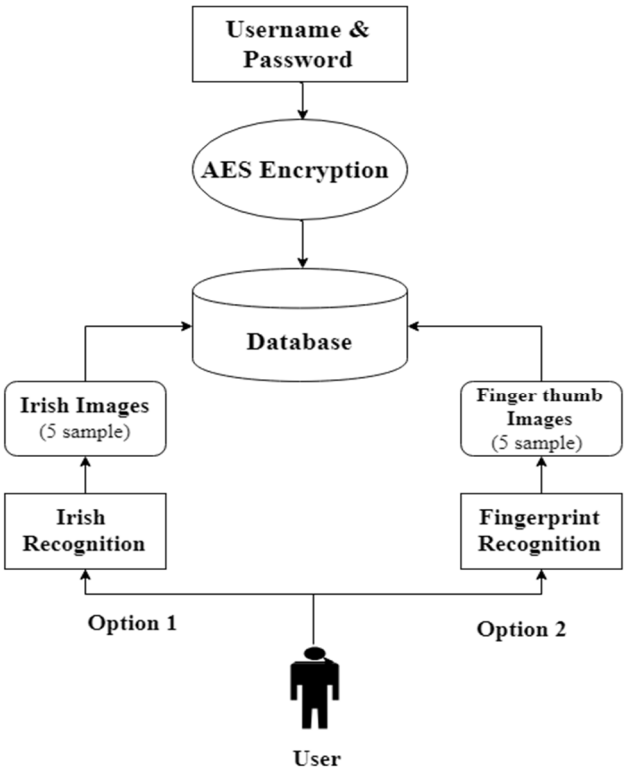


Figure-1: Registration Cycle

Option-2: The Neural Network algorithm will be applied for this option for the user registration. Where this process will take some sample of Finger to execute this method. This sample also saves in the database.

C. Data Matching:

Matching is the very crucial parts of the proposed model where the design requirements to verify with stored data from the database with the input data by comparing. This procedure will show in the figure and it will explain several steps.

Step-1: The username and password will be the first step of the Data-matching process where AES will be executed for encryption and decryption. The user will give his password and username which will be cross match through the Database. When the particular password verified with the database, the user gets access from the method and step-2 begins. But, if not verified with the database, then "Login Error" will be shown.

Step-2: When successfully complete the authentication, the user will able to get through the Biometric verification with one option out of two.

Option-1: This portion works on user Iris recognition where the user's Iris picture will be captured using the Digital camera. here, there are a number of pictures will be reserved. The Neural Network is applied as a pattern recognition with the database. If the Iris matched enough, then the user will get the confirmation to get access to "Login". On the other hand, if the algorithm gives a big number of false position rate on Iris. It will not verify which will indicate "Login Error".

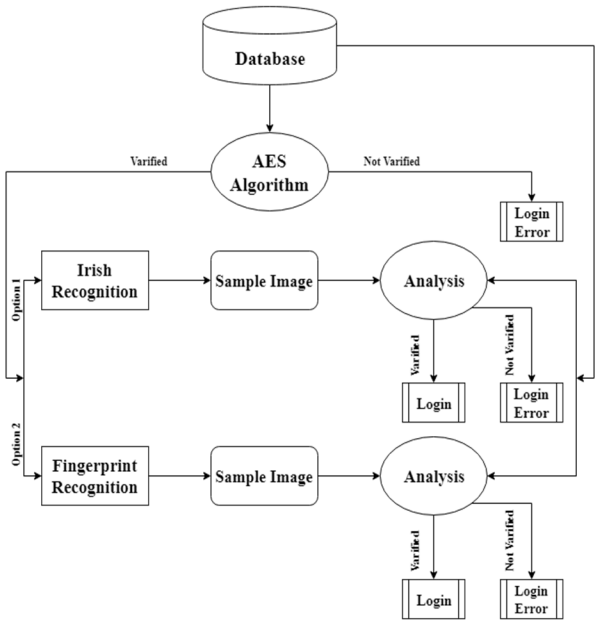


Figure-2: Data Matching process.

Option-2: This step is another Biometric step where the Neural Network algorithm will apply for the pattern recognition on the Fingerprint of the user. If the Iris matched enough, then the user will get the confirmation to get access to "Login". On the other hand, if the algorithm gives a big number of false position rate on Iris. It will not verify which will indicate "Login Error".

IV. Implementation

This part will represent the full proposed method where we designed a system where it will find out the user by using three cited the most secure option. These methods will be shown for AES, Neural Network for Iris recognition and Fingerprint recognition. Figure-1,2,3 defines how the implementation process is verified in the actual object.

A. Advance Security System (AES)

AES is a Symmetric Algorithm, which uses encryption and decryption a data or text by a similar key for both sender and receiver. This is the most proficient algorithm aimed at software application. [14]

AES algorithm is a Block cipher and works on a fixed-length set of bits. It proceeds an input block of 128 bits and yields a matching output block of the similar size. This system calls for a second input, which called a secret key. Moreover, AES has three different key sizes: 128, 192 and 256 bits. This time, Algorithm works on around function for 4 different byte-oriented layers through cipher and inverse cipher:

1) Byte substitution using a substitution table (Sbox)

2) Shifting rows of the State array by different offsets

3) Mixing the data within each column of the State array

4) Adding a Round Key to the State. AES algorithm uses this round function

for ten rounds, but the first and last round of the AES algorithm is differing from other rounds. That is depicted in figure 1 below. This template was designed for two affiliation

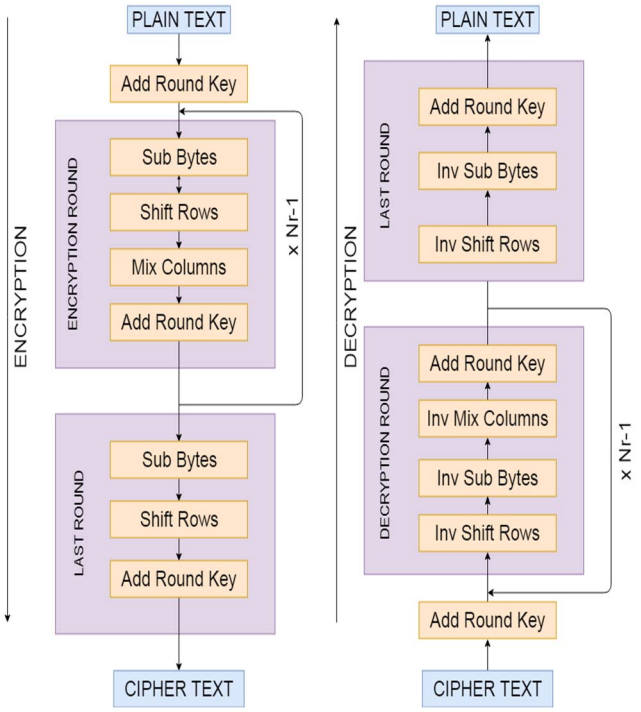


Figure-3: Advanced Encryption Standard (AES) working flowchart.

B. Neural Network

The artificial neural network works on neuron which is taken for input and changes it's some of the internal state names as Activation based on the input, besides create the output dependent on input and Activation. Basically, a directed weighted graph consists of the network which customs by linking that input of additional neurons to the neuron's output. The weights can be improved by a system known as learning that computes the activation and directed by a learning rule. [15]

ANN is capable of learning and it needs to be trained. This system runs on several strategies:

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

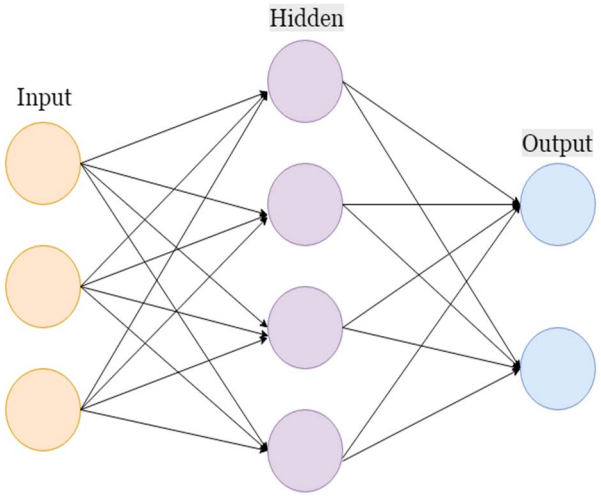


Figure-4: Working process of Artificial Neural Network.

Backpropagation: The backpropagation's weight updates operated by stochastic gradient descent over the equation:

$$\omega_{ij}(t+1) = \omega_{ij}(t) + \eta \frac{\partial C}{\partial \omega_{ij}} + \xi(t)$$

where,

η = the learning rate

C = the cost (loss) function

$\xi(t)$ = stochastic term.

The cost function dependent on influence and the activation. Influence is generated by supervised, reinforcement etc. Such as, if executing a function on a multi-class arrangement of the problem, cost function and common selection of activation are defining consistently, cross entropy function and softmax function. The softmax is shown as

$$\rho_j = \frac{\exp(x_j)}{\sum_j d_j \log(x_k)}$$

where ρ_j is the class probability (output of the unit j and signify the total input to units j and k of the same level correspondingly. [16]

C. Finger Recognition

The fingerprint is the most common and exclusive method in Biometric options which is transnationally known as a legal method to identify a person. The reproductions of the minute ridge which called as Dermal of the Finger. Ridges and Valleys are distinctive and unchangeable. [17] Finger ridge ending as a ridge point is established ends sharply. Ridge splits are mention as ridge point which is connected with the branch ridges. The minutiae location, attributes and time like a direction can be shown in a Fingerprint. There are 40 to 100 minutiae in a standard quality fingerprint sample, on the other hand, dozens of minutiae cannot identify the pattern of the fingerprint. [18]

Implementation of the Fingerprint verification system is broken down into four distinct stages as follows:

- Image acquisition.
- Edge detection.
- Comparison of images.
- Decision making.

D. Iris Recognition

Iris biometric technology was established by John Daugman [19]. Iris biometric is a texture of the Iris can be used as an underlying idea where the code is generated based on the image condition. Iris authentication becomes a progressively spirited research topic in the new era [20]. The exhausting of different types of contact lenses present defies for Iris biometric systems [21]. "Cosmetic contact lenses" used for the change Iris behavior of others in any way. Challenges to identify the image of eye uncovered with cosmetic contact lenses with the image of the similar eye covered with cosmetic contact lens outcome in an extremely wrong nonmatch rate [21]. That similarly imaginable, an attacker trying to imitate a specific user possibly will have cosmetic contact lenses completed to achieve the victim person's Iris grain. This is called "spoofing" of the Iris biometric system. The cornea is the outermost (anterior) layer of the eye. It is normally clear; this system provides some portions of the Iris sample. The aqueous humor is following layer of the cornea. normally clear; this system affords some percentages of the Iris sample. The following layer is Iris (and pupil). Where the quality of the Iris is imaged over the cornea and aqueous humor. [22]

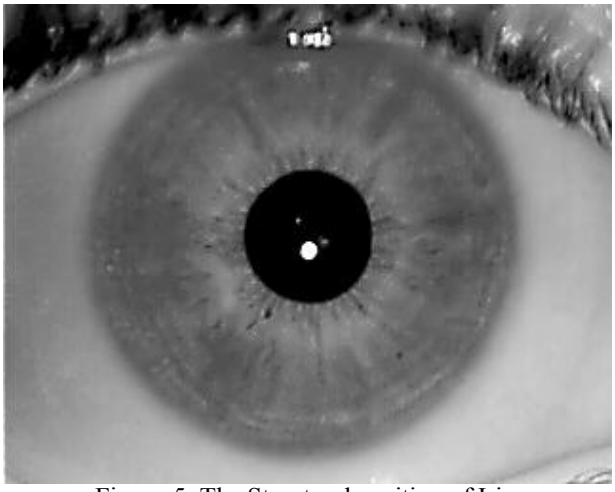


Figure-5: The Structural position of Iris.

Implementation of the Fingerprint verification system is broken down into four distinct stages as follows:

- Image acquisition.
- Edge detection.
- Comparison of images.
- Decision making.

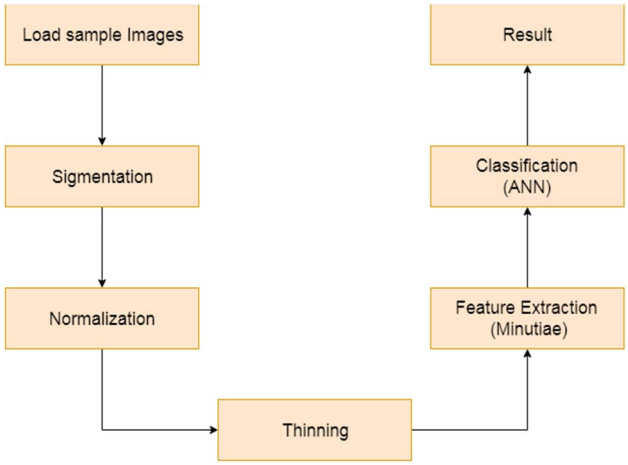


Figure-6: The Working procedure of Finger and Iris Recognition by using ANN.

In MATLAB the iteration is conducted by our own created tow dataset of separately Fingerprints and Irises. Here, we established the iteration is conducted by the sample image matrix binary values. Figure-7 depicts the whole procedure of Fingerprint and Iris recognition mechanism using Artificial Neural Network with some stepwise operations.

The Experimental dataset for the experiment is created our won using camera and fingerprint scanner. Our working station configuration is 7th gen core i3-7020u processor (3m cache,2.30 GHz); 4GB DDR4 RAM, 1TB HDD with

windows 10 setup. In the dataset, our data samples are in .png format images. We collected different peoples' data use for make the data set.

Following steps (1-4) establish how MATLAB works to recognize the user's iris and finger samples from the database. For the successful experiment, we have developed two personal datasets to effectively iris and fingerprint in MATLAB. These steps are provided for fingerprint recognitions where will be similar steps for iris recognition too.

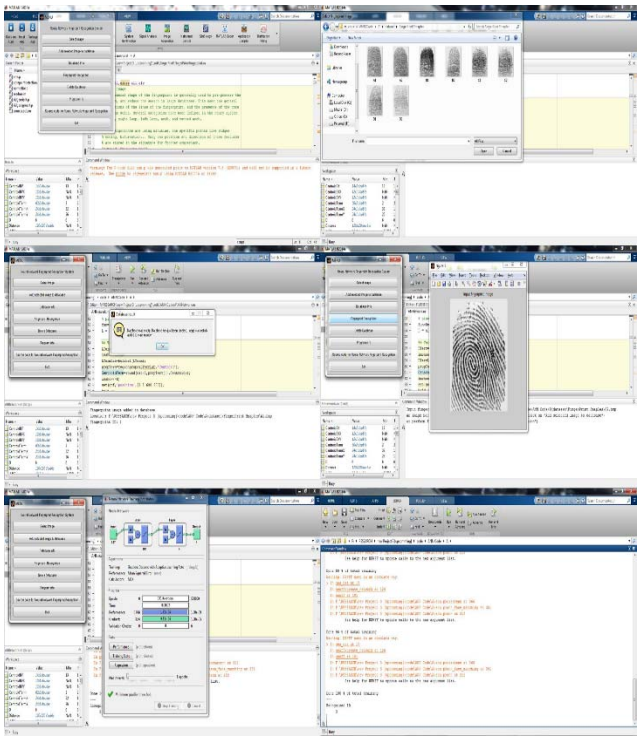


Figure-7: ANN implementation in MATLAB.

Step-1: Executing and administration the codes of ANN to add some samples for the experiment.

Step-2: All samples are in Bitmap (. bmp) image format for our farther steps. Extract and upload system of samples are followed by the database. Here, we provide four different samples for the dataset according to A1, A2; B1, B2; C1, C2; D1, D2.

Samples Name	Dataset Name
A1	1
B2	2
C1	3
D1	4
A2	5

Step-3: Comparing two samples, namely sample 1.bmp with sample 5.bmp, Then the result confirmations that mentioned samples are matched with Sample 1. So the resolution designates the matching of experiment samples.

Step-4: Whenever the samples didn't match, results show the unsuccessful matching results.

V. Experimental Results

In the starting, the proposed method is to implement the AES algorithm. AES has just two operations, encryption, and decryption of the password of each user. After that, the Neural Network toolbox in MATLAB is implemented, where Artificial Neural Network is created for Iris Recognition and biometric Fingerprint. A customized dataset with randomized value is used in case of an experiment and got a great output at the end of the experiment. The experiment is based on time and efficiency issue.

TABLE-1: 1st Layer Encryption, Decryption Analysis

	Iteration	Encryption Cycle (ms)	Decryption Cycle (ms)	Size (KB)
AES Algorithm (256bit)	01	160	110	363
	02	170	140	613
	03	220	180	944

Here, TABLE-1 is representing the result of encryption and decryption cycle of an AES 256 algorithm. Algorithms run for three iterations. For the encryption of AES algorithm gradually takes 160ms, 170ms and finally 220ms for according to 163kb, 313kb and 944kb where, decryption cycle gradually takes time 100ms, 140ms, and 180ms. Where this system provides feedback as a very speedy method.

TABLE-2: 2nd Layer Comparison Analysis

Steps	Iris Recognition (ms)	Finger Recognition (ms)	Size (KB)
Artificial Neural Network Encryption	1134	1465	2654
Artificial Neural Network Decryption	334	255	2654

In the 2nd layer, table-2 is representing the encryption and decryption cycle of the proposed method. Using AI, encryption takes 1134ms for 2654kb for Iris and 1465ms for fingerprint method. On the other hand, the decryption cycle takes 334ms for Iris and 255ms for Fingerprint matching.

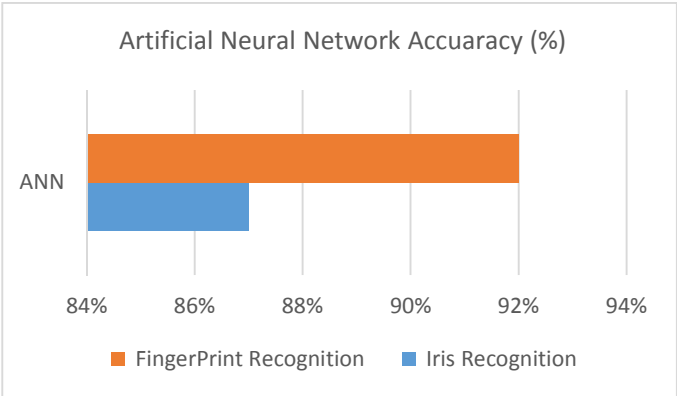


Figure-8: Accuracy Performance of ANN

Here, in the figure-8 shows the accuracy of the recognition is great. Thus the accuracy is showing that 87% accurate for Iris Recognition and for 92% for Fingerprint in biometric technology.

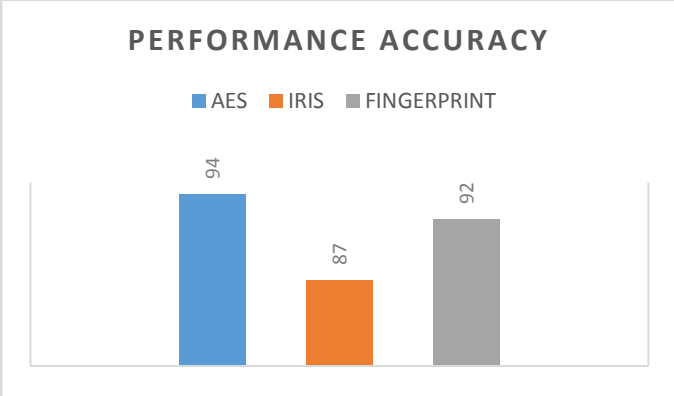


Figure-9: Accuracy Comparison Graph

According to Figure-9; Finally, the performance graph is showing the difference in the final result of the accuracy percentage of the result with the previous method comparison. And the graph surely shows that the proposed method is better working and reliable.

VI. Conclusions

In this paper, the importance of Cloud Computing security has been revealed strongly. Through the proposal method, the first layer AES algorithm provides high performance and accuracy. In the other hand, the second layer; Iris recognition or Fingerprint recognition through Artificial Neural Network will be responsible for the best accuracy of secure access for the user. This proposed procedure can accomplish the secure access of the most uses cloud computing interface for the client users..

Acknowledgment

In this Paper, we would like to acknowledge Mr. Md. Taslim Arefin, Head and Associate professor, Department of ETE, Faculty of Engineering, his project guide for his immense DIU Computer Lab support, genuine concern and invaluable advice that helped me to work on this topic.

I take this opportunity to thank my friends and colleagues for their technical help and moral support.

References

[1] Saini, G. & Singh, P. (2014). Audio Steganography by LSB method and enhanced security with AES. International Journal of Advanced Research in Computer Science & Technology, 2(2), 255-258. Retrieved March 2015.

[2] AES Encryption Algorithm Based on the High-Performance Computing of GPU. Electronic ISBN: 978-1-4244-5727-4, Fei Shao, Zinan Chang, Yi Zhang.

[3] Human Iris recognition for biometric identification. ISBN: 978-1-4244-1550-2. Computer and information technology, 2007. iccit 2007. 10th international conference. Paul Padma Polash, Md. Maruf Monwar.

[4] A Study of Biometric Approach Using Fingerprint Recognition Ravi Subban and Dattatreya P. Mankame. Lecture Notes on Software Engineering, Vol. 1, No. 2, May 2013.

[5] S. Geethalakshmi, P. Subashini, and S. Ramya, "A Study on Detection and Classification of Underwater Mines Using Neural Networks," Int. J. Soft Comput. Eng. IJSCE, vol. 1, no. 5, pp. 150–157, 2006.

[6] B. Howell, S. Wood, and S. Koksal, "Passive sonar Recognition and analysis using hubrid neural networks," IEEE Oceans, vol. 4, pp. 1917–1924, 2003.

[7] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[8] Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders [C]//Proceedings of the 25th international conference on Machine learning. ACM, 2008: 1096-1103.

[9] Bengio Y. Learning deep architectures for AI[J]. Foundations and trends R in Machine Learning, 2009, 2(1): 1-127.

[10] Yang J, Yu K, Gong Y, et al. Linear spatial pyramid matching using sparse coding for image classification[C]//Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009: 1794- 1801.

[11] Collobert R, Weston J. A unified architecture for natural language processing: Deep neural networks with multitask learning[C]//Proceedings of the 25th international conference on Machine learning. ACM, 2008: 160-167.

[12] R. Bodade and S. Talbar, "Fake Iris Detection: A Holistic Approach" International Journal of Computer Applications 19 (2), April 2011.

[13] J. Daugman, "Demodulation by Complex-Valued Wavelets for Stochastic Pattern Recognition", International Journal of Wavelets, Multiresolution and Information processing 1 (1), 2003.

[14] Karthigaikumar, P. & Rasheed, S. (2011). Simulation of image encryption using AES algorithm. Computational Science - New Dimensions and Perspectives, 166. Retrieved March 2015.

[15] Zell, Andreas (1994). "chapter 5.2". Simulation Neuronaler Netze {Simulation of Neural Networks} (in German) (1st ed.). Addison-Wesley. ISBN 3-89319-554-8.

[16] Hinton, G.; Deng, L.; Yu, D.; Dahl, G. E.; Mohamed, A. r; Jaitly, N.; Senior, A.; Vanhoucke, V.; Nguyen, P. (November 2012). "Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups". IEEE Signal Processing Magazine. 29 (6): 82–97. Bibcode:2012ISPM...29...82H. doi:10.1109/msp.2012.2205597. ISSN 1053-5888.

[17] A. K.Jain, S. Pankanti, and R. Bolle, Biometrics: Personal Identification in Networked Society[M]. Boston: Kluwer Academic Publishers, 1999.

[18] R. Thai, "Fingerprint Image Enhancement and Minutiae Extraction", 2003

[19] J. G. Daugman, "Biometric Personal Identification System Based on Iris Analysis", US Patent # 5,291,560, March 1, 1994], [J. Daugman, "How Iris Recognition Works", IEEE Transactions on Circuits and Systems for Video Technology 14 (1), Jan. 2004, 21-30.

[20] M.J. Burge and K.W. Bowyer, Eds., Handbook of Iris Recognition, Springer, 2013. ISBN 978-1-4471-4401-4.

[21] S. E. Baker, A. Hentz, K. W. Bowyer, and P. J. Flynn,"Degradation of Iris Recognition Performance Due to Non-Cosmetic Prescription Contact Lenses," Computer Vision and Image Understanding 114 (9), Sept. 2010,1030-1044.

[22] k.Hughes,KW.Bowyer "Detection of Contact-Lens-Based Iris Biometric Spoofs Using Stereo Imaging" System Sciences (HICSS), 2013 46th Hawaii International Conference on, 18 March. 2013, 13397903.

Non Functional Requirements Classification with Feature Extraction and Machine Learning: An Empirical Study

1st Md. Ariful Haque, 2nd Md. Abdur Rahman, 3rd Md Saeed Siddik
[1,3]Institute of Information Technology, [2]Center for Advanced Research in Sciences
University of Dhaka, Dhaka, Bangladesh
{[1]arifulmit17,[2]mukul.arahman}@gmail.com, [3]saeed.siddik@iit.du.ac.bd

***Abstract*—Non Functional Requirements (NFR) describe a set of quality attributes required for a software such as security, reliability, performance, etc. Extracting and considering NFR from software requirement specification can help developers to deliver quality software which meets users expectations completely. Since, the functional and non-functional requirements are mixed together within the same SRS, it requires a lot of human effort for distinguishing them. This paper proposed automatic NFR classification approach for quality software development by combining machine learning feature extraction and classification techniques. An empirical study with seven machine learning algorithms and four feature selection approaches have been applied to automatically classify NFRs for finding out the best pair. The experiments were measured with statistical analysis including precision, recall, F1-score, and accuracy of the classification results through all the combinations of the techniques and algorithms. It is found that, SGD SVM classifier achieves best result where precision, recall, F1-score, and accuracy reported as 0.66, 0.61, 0.61, and 0.76 respectively. Additionally, TF-IDF (character level) feature extraction technique illustrated higher average score than others.**

***Index Terms*—Non Functional Requirements, Feature Extraction, Classification, Machine Learning, Text Analysis.**

I. Introduction

Software requirements specification is the most important artifacts producing step during the software development life cycle, which is composed of functional and non-functional requirements (NFR). Functional requirements specify software behavior directly listed by stakeholders, where, NFR describe how the software system will provide the means to perform functional tasks. However, the functional and NFRs are mixed together within the same paragraphs, and it becomes hard to detect manually. It is reported that neglecting NFRs leads to project failure or increase production cost [1].

Hence to develop quality software as well as reduce development cost, early NFRs detection is essential because it enables system level constraints to be considered and incorporated into early architectural designs. However, classifying requirements manually is not feasible in terms of required time, budget, and accuracy. To overcome these limitations, various approaches has been presented such as information retrieval based [1], linguistic knowledge based [2], CNN based [3], etc to classify software NFRs into different categories. In 2006, cleland et. al introduced NFRs classifying approach using information retrieval technique [1]. Where, hussain et. al presented a approach to classify text requirements as functional and non-functional using linguistic knowledge [2]. Using convolutional neural network, a method was proposed to classify text based non functional requirements [3]. In 2017, Matsumato et. al proposed a verification method to detect ambiguity, inconsistency, incompleteness, and redundancy of NFRs in software requirement specification [4]. On the other hand, Lu et. al proposed automatic requirements classification approach to classify app user reviews into functional, non-functional and others categories [5].

Although several machine learning methods are existed for automatic NFRs classification task, there is no extensive comparison among machine learning feature extraction, and classification techniques. Where, feature selection is an important task in machine learning to increase classification performance [10]. However, investigation of NFR classification accuracy combining different feature extraction techniques such as Bag of Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), etc. with various classification algorithms named as Multinomial Naive Bayes (MNB), Gaussian Naive Bayes (GNB), Bernoulli Naive Bayes (BNB), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Stochastic Gradient Descent SVM (SGD SVM), and Decision Trees (Dtree) is a significant research direction.

Therefore, this paper proposed an empirical approach to automatic NFR classification technique combining four feature extraction methods with seven machine learning algorithms. The requirements dataset has been pre-processed to remove unnecessary characters which will be converted to feature vector using BoW, TF-IDF (character level), TF-IDF (word level), and TF-IDF (n-gram) techniques. Next, the MNB, GNB, BNB, KNN, SVM, SGD SVM, and DTree algorithms were applied for each type of feature vectors to train the classifier. Finally, based on different NFR classifying results the classification performance of each method has been evaluated. The PROMISE [6] software requirement dataset was used to investigate classification accuracy of considered techniques with several statistical analysis.

The investigated average results reported precision, recall,

and F1-score is higher for TF-IDF (character level). Where, SGD SVM showed best classification result compared to other algorithms. And reports average precision, recall, F1-score, and accuracy as 0.66, 0.61, 0.612, and 0.75 respectively. However, the Dtree achieves worst NFR classification performance for all cases, which reports 0.175, 0.17, 0.14, and 0.51 as precision, recall, F1-score, and accuracy respectively.

The rest of this paper is organized as follows; Section II discussed the existing work related to this thesis. Section III and IV illustrated the proposed method and result analysis respectively. Finally section V concludes this paper with future research direction.

II. Related Works

Software requirements are classified into functional and non-functional based on textual requirements analysis. Various requirements classification research has been investigated in the literature such as information retrieval based [1], linguistic knowledge based [2], multinomial naive bayes based [7], etc.. The approaches have been discussed in this section.

Cleland et. al introduced an approach for detecting and classifying non functional requirements using information retrieval technique in 2006 [1]. First, a set of indicator terms is identified for each NFR category, then additional requirements are classified. Indicator terms which is used to detect requirements were used as the basis of identification of different NFRs. A probabilistic weight for each potential indicator term is calculated from the requirements. This strategy requires less effort for NFR classification compared to semi-automated classification methods. However, this technique suffers for evaluating correctness of candidate NFRs manually.

In 2008, hussain et. al proposed a classification approach to classify text requirements into two categories named as functional and non-functional using linguistic knowledge [2]. The classifier works at sentence level as the characteristics of FR and NFR remain within the scope of sentences. The reported results were cross checked ten times over same data. The investigated results showed that the method improved classification performance significantly. However, more training and testing data could be introduced to improve classification performance.

Ju et. al proposed a method to classify non-functional requirements using convolutional neural network [3].The technique converts text data in low-dimensional vector space using Word2vec to vectorize requirement words in dataset. The average score of precision, recall, and f-measure has been used to evaluate proposed deep learning model. The reported results showed that CNN performs well for text based NFR classification task automatically. However, recurrent neural network could be incorporated to improved classification accuracy.

Matsumato et. al proposed a verification method using requirements frame model to detect ambiguity, inconsistency, incompleteness, and redundancy of NFRs in software requirement specification [4]. The investigation were focused on time-response and usability related requirements. In this experiment, the specific requirements were retrieved using respective keywords. Also multiline statements were converted to single sentence manually. However, such manual conversion is time consuming and error-prone.

Lu et. al proposed automatic requirements classification approach to classify app user reviews [5]. This approach combined BoW, TF-IDF, CHI^2 and AUR-BoW classification techniques with NB, J48, and Bagging algorithms to automatically classify user reviews. The classification technique was focused to four types of NFRs named as reliability, usability, portability, and performance. This research reported that in an automatic classification using unbalanced dataset performed worse when the size of certain types is obviously smaller. However, no solution has been provided for the mentioned unbalanced dataset problem.

Slankas et. al proposed a method called NFR locator and developed a tool to extract relevant NFRs from a wide variety of documents [7]. In this research, a word vector representation was used followed by the use of a multinomial naive bayes and k-nearest neighbor classifiers. After parsing document into an internal representation, classified sentences into specific NFR categories. The k-nearest neighbor classifier with a unique distance metric had outperformed in their experiments the optimal naive Bayes classifier.

Khan et. al investigated different existing methodologies for software requirement classification [8]. This research identifies co-relations between software architecture and NFRs in order to check their applicability in software development life cycle. The presented comparison results revealed that software architecture is important to address NFRs.

Kurtanovi et. al proposed a strategy to handle the imbalanced classes in the dataset and the classifiers were cross-validated using precision, recall, and F1 metrics [9]. The approach was automatically identified NFRs using supervised requirements. Where, four NFRs named as usability, security, operational, and performance were focused. After removing stop words and punctuations, BoW, bigrams, and trigrams techniques were applied for feature selection. The paper stated that part of speech tags are the most informative features. Experiments were conducted using SVM classifier algorithm.

The analysis of existing text classification approaches concluded that different strategies have been proposed for non-functional requirement analysis such as information retrieval based, linguistic knowledge based, CNN based, etc. However, no guideline has been found yet to select the best pair of feature extraction and machine learning for NFR classification by performance investigation of different feature extraction techniques and machine learning algorithms.

III. METHODOLOGY

This paper proposed technique for finding the best pair of feature extraction and machine learning approach on software requirement classification. This process considers NB, KNN, SVM, and D-tree machine learning algorithms and BoW and TF-IDF feature extraction technique for comparison. The software requirement classification performance of each algorithm is measured using different featured extraction methods named

as BoW, TF-IDF (character level), TF-IDF (word level), and N-gram. In this framework, text data has been pre-processed to remove unwanted texts, which then encoded with text encoding techniques. The encoded data has been trained and tested by NB, KNN, SVM, and Dtree algorithms. The whole process of this framework is divided into following four steps and depicted in Figure 1.

- Step 1: Data Preprocessing
- Step 2: Feature Extraction
- Step 3: Train Classifier
- Step 4: Classifying requirements

The details regarding the steps is elaborated in following sections and sub-sections.

A. Step 1: Data Pre-processing

The first step of this method is to cleaning and pre-processing dataset, which is composed of four sub-steps: (i) removing special characters, (ii) case-folding, (iii) removing stop words, and (iv) tokenization. Special characters and symbols are usually non-alphanumeric characters, which add extra noise to the experimental dataset. So, special characters has been removed using regular expression in order to remove noisy data. In English, words appear in upper-case or lower-case having similar meaning, hence, case-folding is applied for unique case consideration. On the other hand, Stop words usually refer to the most common words such as "a", "an", "the", which do not influence the semantics of a software requirement. The most uses natural language processing stop words corpus of nltk has been used in this framework. Finally, tokenization splits longer text into smaller piece or tokens. The requirements has been tokenized into sentences which then splitted into distinct words.

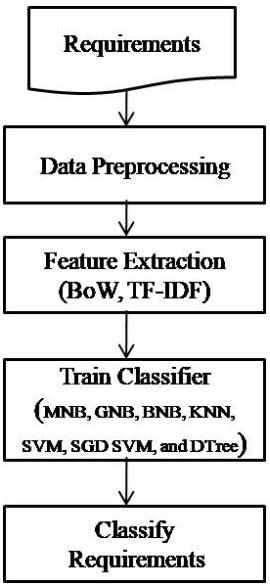


Fig. 1. Proposed Method Overview

B. Step 2: Feature Extraction

Textual feature extraction is a process of transforming list of essential words into a feature set that is usable by machine learning classifier. Textual dataset consists of documents which represents words, sentences or paragraphs of free flowing text and those are inherently unstructured i.e. no neatly formatted data columns. Therefore, extracting significant and meaningful features from text data is essential to building text classification machine learning model. In this step, textual data has been converted from free flowing text into some numeric representations to be understood by machine learning algorithms.

In this framework, a set of feature has been extracted from text data to automatically classify software requirements using machine learning algorithms. Initially this research went through the standards steps for text preprocessing. The feature extraction process from the preprocessed text are described in the following sections.

- **Bag of Words based feature:** Bag of Words (BoW) vector space representation model has been used as a first set of features for training machine learning algorithms. This vector space model represents unstructured text as numeric vectors, where it constructs a word presence feature set from all the words of an instance [5]. In this paper, requirements have been converted into numeric vectors such that each document is represented by one vector (row). As example, the following BoW model based document feature vectors was generated for the sentences of 'System shall be used by realtors', and 'System shall display Events'. The unique words in the corpus are 'The system shall be used by realtors display Events', therefore the vocabulary size will be of 9 words from total 12 words. The vector representation of this corpus is shown in table I.

TABLE I
BOW EXAMPLE

system	shall	be	used	by	realtors	display	Events
1	1	1	1	1	1	0	0
1	1	0	0	0	0	1	1

From the above table, it is clear that each column or dimension in the feature vectors represents a word from the corpus and each row represents one of corpus documents. The value in any cell, represents the number of times that word occurs in the specific documents which is represented by row. Therefore, if a corpus of documents consists of N unique words across all documents, it will have an N-dimensional vector for each of the documents.

- **Term frequency based features:** additionally to the BoW features, tf-idf scores associated with each term present in a given requirement has been used in this classification framework. The tf-idf term weighting technique assigns a higher score to rare words and a lower one to words occuring frequently across all requirements [5]. The tf-idf score calculation process is defined as following Equation (1):

$$tf - idf_{t,d} = tf_{t,d} \times idf_t = tf_{t,d} \times log(\frac{N}{df_t}) \quad (1)$$

where $tf_{t,d}$ is the frequency of word t in requirement document d while, idf_t is the inverse document frequency of terms occurring in various reviews. The three techniques of $tf - idf$ named as character level, word level and n-gram have been used in this paper.

After cleaning the requirements strings, above techniques were applied to extract textual features to train the classifiers.

C. Step 3: Train Classifier

Four classification techniques BoW, TF-IDF (character level), TF-IDF (word level), and TF-IDF (n-gram) has been used in above step to extract textual features, which act as the input of machine learning algorithms of training classifiers. The experimentation has been conducted using seven machine learning algorithms MNB, GNB, BNB, KNN, SVM, SGD SVM, and DTree. The multiple algorithms were applied for requirement classification to compare the performance.

D. Step 4: Classifying requirements

Each of the classifier trained in previous section will output for a given requirement whether is belongs to a category or not. For example, in order to classify requirement according to category PE (performance), the framework will return the list of requirements for which it received a positive answer. Also, the other documents will be classified accordingly.

The combination of four textual feature extraction method and seven machine learning algorithm has been applied in this software requirement classification framework. The textual data has been converted into vector representations to fed as input in machine learning algorithms.

IV. Evaluation

The experimental study, used requirement corpus and evaluated results are demonstrated in this section. Precision, recall, and F1-score have been measured to evaluate how well the model learn to classify software requirement. The labeled data was used for training and unlabeled text strings were used for testing. The experimental results of classification of NFR dataset for different machine learning algorithms are shown in Table II and Table III.

A. Environment Setup

The experimentation has been carried out on intel core i7, 32GB RAM and Ubuntu 16.06. Pandas, numpy, nltk, sklearn and matplotlib packages were used for loading the data preprocessing, and result evaluation.

TABLE III
Precision, Recall and F1-Score Comparison

Methods	Average Score			
	Precision	Recall	F1-Score	Accuracy
MNB	0.22	0.19	0.17	0.53
GNB	0.55	0.54	0.52	0.66
BNB	0.19	0.18	0.16	0.52
KNN	0.62	0.48	0.50	0.69
SVM	0.59	0.46	0.50	0.70
SGD SVM	0.66	0.61	0.61	0.76
Dtree	0.18	0.17	0.15	0.51

B. Datasets

The most popular PROMISE software requirement dataset has been used for experimental analysis [6]. It is a widely used dataset for software requirement classification research. It contains a set of labeled requirements in functional and nonfunctional (11 categories) requirements. The dataset is small in size which consists of 625 requirement sentences. In these sentences 255 are functional and 370 are non-functional requirements (NFRs). The NFR is labeled with eleven categories named as availability, legal, look and feel, maintainability, operational, performance, scalability, security, usability, fault tolerance, and portability. However, the dataset contains majority number of functional requirements which indicates class distribution is imbalanced.

C. Result Analysis

The prepared dataset is trained and tested to evaluate classifying accuracy of proposed model. The experimental results of non functional requirement classification using several feature extraction and machine learning algorithms have shown in Table II. Where, Table III demonstrated the average score of proposed seven machine learning algorithms.

Table II provides an overview of the experiment results obtained through the feature extraction techniques BoW, TF-IDF (character and word level, n-gram) combined with seven machine learning algorithms MNB, GNB, BNB, KNN, SVM, SGD SVM, and DTree, which show the weighted average precision, recall, and F1-score of the classification results.

From Table II performance comparison of different feature extraction methods with respect to different machine learning algorithms is displayed. It is observed that overall performance of TF-IDF (character level) feature extraction is better than other mentioned methods. However, for some cases BoW out performed than other two TF-IDF (word level and n-gram) feature extraction methods. The highest precision, recall, and F1-score is reported by SGD SVM technique for all feature extraction methods. However, SGD SVM best performs with TF-IDF (character level) which showed 0.70, 0.62, and 0.63 for precision, recall, and F1-score respectively. On the other hand, BNB with TF-IDF (N-gram) and MNB with TF-IDF (character level) performs worst, however, in Table III Dtree showed average lowest score as 0.18, 0.17, 0.15, and 0.51 for precision, recall, F1-score, and accuracy respectively.

In case of feature extraction techniques, this experiment showed almost same average score for BoW and TF-IDF (character level) which is shown in Table II. The average score of TF-IDF (character level) method shown as 0.45, 0.40 and 0.40 for precision, recall and F1-score respectively which is higher than TF-IDF (word level, and n-gram) feature extraction methods.

In Table III, the weighted average score has been reported, where, SGD SVM method showed highest score again for all cases. The SGD SVM achieves precision, recall, F1-score, and accuracy as 0.66, 0.61, 0.61, and 0.76 respectively. On the other hand, decision tree technique performs worst among

TABLE II
PRECISION, RECALL AND F1-SCORE COMPARISON

Encoding/ Algorithm	BoW			TF-IDF (Word level)			TF-IDF (n-gram)			TF-IDF (Character level)		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score	Precision	Recall	F1-Score	Precision	Recall	F1-Score
MNB	0.47	0.43	0.40	0.13	0.13	0.11	0.22	0.12	0.1	0.04	0.09	0.06
GNB	0.58	0.54	0.55	0.59	0.51	0.53	0.54	0.68	0.57	0.5	0.41	0.43
BNB	0.15	0.13	0.12	0.15	0.13	0.12	0.04	0.09	0.06	0.4	0.35	0.35
KNN	0.49	0.29	0.28	0.64	0.55	0.58	0.68	0.48	0.52	0.65	0.58	0.59
SVM	0.65	0.58	0.60	0.63	0.53	0.56	0.41	0.22	0.25	0.66	0.52	0.57
SGD SVM	0.60	0.61	0.58	0.64	0.62	0.62	0.71	0.6	0.62	0.70	0.62	0.63
Dtree	0.23	0.20	0.18	0.23	0.21	0.19	0.13	0.1	0.07	0.11	0.18	0.14
Average	0.45	0.40	0.39	0.43	0.38	0.39	0.39	0.33	0.31	0.45	0.40	0.40

the seven algorithms, this method reports precision, recall, F1-score, and accuracy as 0.18, 0.17, 0.15, and 0.51 respectively. The KNN and SVM has shown moderate performance where, precision, recall, F1-score, and accuracy of KNN are 0.62, 0.48, 0.50, and 0.69 respectively. On the other hand, SVM reports 0.59, 0.46, and 0.50, 0.70 as precision, recall, F1-score, and accuracy respectively.

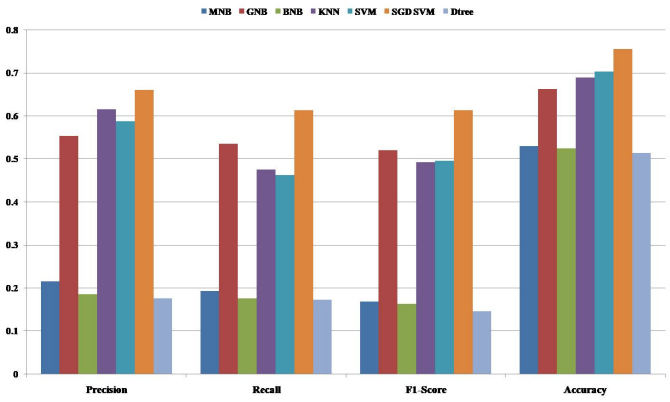


Fig. 2. Average Score Comparison

In Fig 2 the average results have been plotted, where SGD SVM method dominating other algorithms. However, BNB, MNB, and Dtree scored almost similar for textual NFRs classification. All the empirical results concluded that the pair of SGD SVM and TF-IDF (character level) performs best over other machine learning and feature detection pairs for non functional requirements classification. Functional requirement classification is beyond the scope of this research.

V. CONCLUSION

This NFR identifying framework focused on the investigation of various feature extraction and classification techniques classification performance. As feature extractor, BoW, and variation of TF-IDF techniques have been used. After getting, featured vectors from textual data, seven classifier algorithms named as MNB, GNB, BNB, KNN, SVM, SGD SVM, and DTree have been applied to classify requirements into different categories. For experimentation, standard PROMISE NFRs dataset have been used, where, all requirements are categorize into several categories. Among the requirement classifying algorithms, SGD SVM performed better than other algorithms, which reported average precision, recall, F1-score, and accuracy as 0.66, 0.61, 0.612, and 0.75 respectively. Where, decision tree performed worst which reports precision, recall, F1-score, and accuracy as 0.18, 0.17, 0.14, and 0.51 respectively. However, GNB and KNN achieved moderate score which is higher than DTree and lower than SGD SVM methods. On the other hand, TF-IDF (character level) feature extraction technique illustrated higher average score than others. This empirical research recommended TF-IDF (character level) for feature extraction with SGD SVM algorithm to predict best result in NFR classification.

Incorporating other significant machine learning feature extraction techniques such as Chi Squared, and classification algorithms such as bagging, boosting are the future research direction of this paper.

REFERENCES

[1] J. Cleland-Huang, R. Settimi, X. Zou, and P. Solc, "The detection and classification of non-functional requirements with application to early aspects," in *14th IEEE International Requirements Engineering Conference (RE'06)*, pp. 39–48, IEEE, 2006.

[2] I. Hussain, L. Kosseim, and O. Ormandjieva, "Using linguistic knowledge to classify non-functional requirements in srs documents," in *International Conference on Application of Natural Language to Information Systems*, pp. 287–298, Springer, 2008.

[3] R. Ju, and G. Licea, "Towards supporting software engineering using deep learning: A case of software requirements classification," in *2017 5th Int. Conf. in Software Engineering Research and Innovation (CONISOFT)*, pp. 116–120, IEEE, 2017.

[4] Y. Matsumoto, S. Shirai, and A. Ohnishi, A method for verifying non-functional requirements, *Procedia Computer Science*, vol. 112, pp. 157–166, 2017.

[5] M. Lu and P. Liang, "Automatic classification of non-functional requirements from augmented app user reviews," in *Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering*, pp. 344–353, ACM, 2017.

[6] The Promise Repository of Empirical Software Engineering Data. 2015.

[7] J. Slankas and L. Williams, "Automated extraction of non-functional requirements in available documentation," in *2013 1st International Workshop on Natural Language Analysis in Software Engineering (NaturaLiSE)*, pp. 9–16, IEEE, 2013.

[8] F. Khan, S. R. Jan, M. Tahir, S. Khan, and F. Ullah, "Survey: dealing non-functional requirements at architecture level," *VFAST Transactions on Software Engineering*, vol. 9, no. 2, pp. 7–13, 2016.

[9] Z. Kurtanović and W. Maalej, "Automatically classifying functional and non-functional requirements using supervised machine learning," in *2017 IEEE 25th International Requirements Engineering Conference (RE)*, pp. 490–495, IEEE, 2017.

[10] B. Xue, M. Zhang, W. N. Browne, and X. Yao, "A survey on evolutionary computation approaches to feature selection," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 4, pp. 606–626, 2016.

Design and Development of Question Answering System in Bangla Language from Multiple Documents

[1]Samina Tasnia Islam
Computer Science and Engineering Department
Military Institute of Science and Technology
Dhaka, Bangladeh
saminatasnia113@gmail.com

[2]Dr. Md. Nurul Huda
Computer Science and Engineering Department
United International University
Dhaka, Bangladeh
mnh@cse.uiu.ac.bd

***Abstract* – The paper has presented about the experiment of design and development of automatic question answering system for Bangla language. The purpose of proposed answering system is to provide answer based on the keyword, lexical and semantic feature of a question. User gives a question for which answer has to be found from multiple documents. For measurement (time and quantity) related question the system gives specific answer otherwise the system retrieves relevant answers.**

Keywords – Question answering; keyword extraction; stemming; information retrieval; sentence ranking.

I. INTRODUCTION

In Question Answering (QA) system a user finds relevant answer of a question from unstructured document collection. QA is similar with Information Retrieval (IR) as minimum amount of information is retrieved which is enough to fulfill user demands [1].

Nowadays we need a system which will receive a question in natural language from user and find answers form given documents and provide relevant answers quickly to user. This study designed a QA system in NLP for Bangla language that allows a user to give question and provides relevant answers from a single or multiple documents. The originality of this paper is that the proposed system identifies the question type and for measurement (time and quantity) related question it provides relevant specific answer otherwise it retrieves relevant answers.

In this paper Section-II discusses the problems of QA system in Bangla language, section-III is about the implementation of the proposed system, section-IV analyses the evaluation of the system, section-V is the conclusion and finally references are given in section-VI.

II. PROBLEMS OF QUESTION ANSWERING SYSTEM IN BANGLA LANGUAGE

Question answering system in Bangla language faces many problems. Problems of it are following:

1) Automatic QA from unstructured documents is a difficult task as there is a probability that the source text may contain only one answer to any user's question [2].

2) Mapping questions to answers between question strings to answer strings using lexical, syntactic or semantic relationship is another difficult task [2].

3) The greater the answer redundancy in the text document, the more possibility to retrieve an answer in a simple relation to the user's question. Otherwise, we will need to solve the above issues facing NLP systems [2].

4) In Bangla language there is another challenge to identify the keyword or headword from the question as there is no specific rules in which place the *'wh'* word of the question will be appeared.

III. IMPLEMENTATION

The proposed system can be designed by merging the contents of all the documents, removing stop words, stemming of question and text documents, keyword extraction from question, N-grams formation from keywords for approximate matching, retrieving n-best answers and generating specific answers by question type and evaluation of performance and correctness. Details of the steps are following:

A. Removing Stop Words

In the proposed system necessary intelligence about conjunctions, pronouns, verbs and also inexhaustible words have to be provided for removing these stop words from the question as well as from the text documents.

B. Stemming

Intelligence about suffixes is also needed for stemming the words of the given question as well as the text document. Both the training sets of question as well as the documents have to be stemmed to find out the morphological stem of the words for approximate matching and retrieving information from source documents.

C. Keyword Extraction and N-grams formation from Keywords for Approximate Matching:

Keywords or headwords from the question have to be generated. In the proposed system one of the statistical approaches that is the word intermediate distance vector and its mean value was used to extract keyword from sentence [3].

N-grams have to be formed for effective approximate matching. Keywords generated from questions will be used to form n-grams (unigram/bigram/trigram), thus allowing the n-grams to be compared with other sequences for retrieving relevant information from the source text.

Table 1: Keyword Extraction Process. Here the keywords/Headwords are the root of the words.

Question	Keywords
বাংলাদেশে প্রথম কম্পিউটার কবে আসে?	Array ([0] => প্থম [1] => কম্পিউটা [2] => আস)
অ্যাবাকাস কবে আবিষ্কৃত হয়?	Array ([0] => অ্যাবাকাস [1] => আবিষ্কৃত)
জন নেপিয়ার (John Napier) এর অস্থি কি?	Array ([0] => নেপিয়া [1] => (John [2] => Napier) [3] => এ [4] => অস্থি [5] =>)
গটফ্রাইড ভন লিবনিজ কিভাবে যান্ত্রিক ক্যালকুলেটর আবিষ্কার করেন?	Array ([0] => গটফইড [1] => যান্ত্রিক [2] => ক্যালকুলেট [3] => আবিষ্কা [4] => কেন)
রিকোনিং যন্ত্র কি?	Array ([0] => যন্ত্র)
ডিফারেন্স ইন্জিন কি?	Array ([0] => ইন্জিন)

In Table 1, all the questions are set from a document related with computer collected from Wikipedia (Bangla). Here it's shown that keywords are extracted from the question. Here the keywords are stemmed to find out the morphological root of the word.

D. Lexical and Semantic Features:

In Bangla question ***"****wh-word" is* a vital lexical feature. An important role is played by the end marker. If the end marker is "|" then the given question is definition type [4].

In the proposed system, measurement unit is used as semantic feature. According to interrogative (*wh-type*) type and semantic feature answer will be retrieved.

Here, if the question type is time related কবে (kəbɛ)/ কখন (kəkhən) and quantity related কত (kət) then proposed system will give specific answer.

A document related with "Cox's Bazar" which is one of the tourist spot of Bangladesh and another document related with "Computer" have been collected from wikipedia(Bangla) and several questions given below have been set from these documents.

- **Question 1:** "কক্সবাজার থানা প্রথম কবে প্রতিষ্ঠিত হয়? (When Cox's Bazar Thana has been first established?)"
  Answer 1: "১৮৫৪ সাল" ("1854 year")
  Answer 2: "সাল এবং পৌসভা" ("Year and Municipality")
- **Question 2:** "কম্পিউটারে প্রথম বাংলা লেখা সম্ভব হয় কখন? (When Bangla writing was possible in Computer first?)"
  Answer 1: "১৯৮৭ সাল" ("1987 Year")
- **Question 3:** "বাংলাদেশে প্রথম কম্পিউটার আসে কত সালে ? (When Computer has first come in Bangladesh?)"
  Answer 1: "১৯৬৪ সাল" ("1964 Year")
  Answer 2: "১৯৭১ সাল" ("1971 Year")
  Answer 3: "১৯৮১ সাল" ("1981 Year")

In the first question the first answer is correct; the second answer here matches with the keywords of the question. In question 2, here only one answer is retrieved which is correct and in question 3 first answer is correct and other answers matches with the keywords of the question. In these answers every word is the root of the words itself.

E. Ranking the retrieved sentences

Retrieved sentences will be ranked by Textual Entailment Module (TE) [5] and best ranked retrieved information will be considered as answer.

IV. EVALUATION

Evaluation of the proposed system can be carried out by using the following formulas.

$$\text{Precision} = \frac{\text{Relevant Items Retreived}}{\text{Retreived Items}} \dots\dots \text{eq}^{(1)}$$

$$\text{Recall} = \frac{\text{Relevant Items Reteived}}{\text{Relevant Items}} \dots\dots\dots \text{eq}^{(2)}$$

$$\text{F Score} = 2 * \frac{(\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}} \dots\dots\dots \text{eq}^{(3)}$$

For evaluation purpose several documents have been selected from Wikipedia. Approximate 500 questions have been set to test the performance and correctness. At first for every single question relevant items have been retrieved, retrieved items and total no of relevant answer for the question given by user have been identified. Thus for every single question precision and recall have been calculated using $\text{eq}^{(1)}$ and $\text{eq}^{(2)}$ respectively.

Then average precision and average recall have been calculated. With the value of average precision and average recall F Score/ F Measure has been calculated using $\text{eq}^{(3)}$.

After testing approximate 500 question average precision (0.35) average recall (0.65) have been obtained and the F score (0.45) is calculated based on average precision and averave recall. F score/ F measure reveals the system performance. The higher F score indicates that the system is more better.

Table 2: Process for Calculating Precision and Recall

Questions	Relevant Answers Retreived	Retreived Answers	Relevant Answers	Precision	Recall
কম্পিউটার শব্দের উতপত্তি কিভাবে?	1	2	1	0.5	1
যান্ত্রিক ক্যালকুলেটর সর্বপ্রথম কবে আবিষ্কৃত হয়?	1	1	1	1	1
কম্পিউটার শব্দের অর্থ কি?	0	1	1	0	0
বাংলাদেশে প্রথম কম্পিউটার কবে আসে?	1	1	1	1	1
অ্যাবাকাস কবে তৈরি হয়?	0	6	2	0	0
জন নেপিয়ার (John Napier) এর অস্থি কি?	1	2	2	0.5	0.5
গটফ্রাইড ভন লিবনিজ কিভাবে যান্ত্রিক ক্যালকুলেটর আবিষ্কার করেন?	1	2	1	0.5	1
রিকোনিং যন্ত্র কি?	0	9	2	0	0
যান্ত্রিক ক্যালকুলেটর সর্বপ্রথম কবে আবিষ্কৃত হয়?	1	2	1	0.5	1
গণকযন্ত্র কি?	1	1	1	1	1
মাইক্রোপ্রসেসর উদ্ভাবক কোন প্রতিষ্ঠান?	0	5	1	0	0
কম্পিউটারে প্রথম বাংলা লেখা সম্ভব হয় কখন?	1	2	1	0.5	1
বাংলা ওয়ার্ডপ্রসেসিং সফটওয়্যার উদ্ভাবন করে কারা?	1	1	1	1	1
মাইক্রোসফট উইন্ডোজ' এর সঙ্গে ব্যবহারের জন্য ইন্টারফেস 'বিজয়' কবে উদ্ভাবিত হয়?	1	2	1	0.5	1
			Average:	**0.5**	**0.69**

In Table 2 process for calculating precision and recall has been shown. The first question in Table 2, this system has retreived two answers so here retrieved items=2. Among the two answers only one is relavent with the question and in the given document there is only one relevant answer as a result here relevant items retrieved=1 and relevant items=1. So for the first question using $eq^{(1)}$ and $eq^{(2)}$ precision (1/2=0.50) and recall (1/1=1.00) have been calculate. Like this way for every single question precision and recall have been calculated. Using the value of precision and recall average precision and average recall were calculated.

V. Conclusion

This paper has discussed a QA system from multiple documents. This study concludes the following:

- This approach is able to provide relavent answers of a user question for Bangla Language from multiple documents.
- Keywords of a question can be find out in this approach.
- This approach retreives relevant specific answers for time related and quantity related question.
- Precision, recall and F score of the system are 0.35, 0.65 and 0.45 respectively.
- If the document size is large then there is a possibilty that the system may retreive some less relevant information as total no of retrieved information increases.

In future the author of this paper would like to do the following tasks:

- There exists some more answer or information retreival system for some other languages like Chinese, English, Japanese [6][7][[8][9]. So this approach can be compared in future with other existing answer retreival approaches.
- Finding other question type by "*wh*" words and retreiving answers accordingly.
- If any relevant answers starts with pronoun then there is a chance that more relavant answers may exist in other sentences of the text. So finding out more relevant answers if any answers starts with pronoun.

VI. Reference

[1] Davide Buscaldi, et al, "Answering questions with an n-gram based passage retrieval engine," *Journal of Intelligent Information Systems,* vol. 34.2, pp. 113-134, 2010.

[2] Eric Brill, Susan Dumais, and Michele Banko, "An analysis of the AskMSR question-answering system," *Proceedings of the ACL-02 conference on Empirical*

methods in natural language processing, vol. 10, Association for Computational Linguistics, 2002.

[3] Sifatullah Siddiqi and Aditi Sharan, "Keyword extraction from single documents using mean word intermediate distance," *International Journal of Advanced Computer Research,* vol. 6.25, pp. 138, 2016.

[4] Somnath Banerjee and Sivaji Bandyopadhyay, "Bengali question classification: Towards developing qa system," *Proceedings of the 3rd Workshop on South and Southeast Asian Natural Language Processing*, 2012.

[5] Partha Pakray, et al, "A Hybrid Question Answering System based on Information Retrieval," *CLEF (Notebook Papers/Labs/Workshop)*, 2011.

[6] Liang Zhenqiu, "Design of automatic question answering system base on CBR," *Procedia Engineering,* vol. 29, pp. 981-985, 2012.

[7] Eriks Sneiders, "Automated question answering using question templates that cover the conceptual model of the database," *International Conference on Application of Natural Language to Information Systems*, Springer Berlin Heidelberg, 2002.

[8] Yuanzhi ke and Masafumi Hagiwara, "An English neural network that learns texts, finds hidden knowledge, and answers questions," *Journal of Artificial Intelligence and Soft Computing Research,* vol 7.4, pp. 229-242, 2017.

[9] Tetsuya Sakai, et al, "ASKMi: A Japanese question answering system based on semantic role analysis," *Coupling approaches, coupling media and coupling languages for information retrieval,* pp. 215-231, 2004.

[10] L. Hirschman and Robert Gaizauskas, "Natural language question answering: the view from here," *natural language engineering,* vol 7.04, pp. 275-300, 2001.

[11] Andrea Andrenucci and Eriks Sneiders, "Automated Question Answering: Review of the Main Approaches," *ICITA (1)*, 2005.

[12] Md. Zahurul Islam, Md. Nizam Uddin and Mumit Khan, "A light weight stemmer for Bengali and its Use in spelling Checker," *Centre for Research on Bangla Language Processing*, 2007.

Using Boosting Approaches to Detect Spam Reviews

Sifat Ahmed
Department of Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
sifat.austech@outlook.com

Faisal Muhammad Shah
Department of Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
faisal505@hotmail.com

***Abstract*— When it comes down to buying products from online shops, one of the key factor that influences a buyer are the reviews associated with a product. While buying people try to understand the quality and authenticity of the product by reading the previous user feedback. And sellers have started taking advantage of it. Putting fake and spam reviews to deceive the buyers is a common strategy mostly used by newcomers. But these reviews are important when it comes to deciding whether to buy a product or not. We propose a method to detect these fake reviews from Amazon Review Dataset. Rather than using traditional machine learning classifiers we have used boosting algorithms to improve the accuracy of the traditional approach. In this approach, a significant increase in accuracy has been achieved by boosting weak learners. Up to 93% accuracy has been achieved when tried to detect fake reviews where traditional machine learning algorithms achieve an accuracy of up to 89%.**

Keywords—fake review detection, boosting algorithm, spam detection, active learning.

I. INTRODUCTION

For every product sold on the internet and every buyer bought a product, around 60% of people post a review of that. As people are getting more connected to the internet, the style, medium of putting a review has changed. Nowadays people even make video reviews of a product to attract sellers and the industry is still growing where people promote products for compensation. But amongst all, there are people who write honest reviews to help others when it comes to making a decision about buying the product. Some sellers are taking this advantage by posting fake reviews to deceive the buyers.

Detecting these fake reviews are not a simple task when it comes down to handing over the responsibility to the computer. Some fake reviews are so skillfully written that even a human can't detect that let alone computers. And the freedom to write anything with no monitoring has made it more difficult to detect these deceptive reviews. These deceptive reviewers are mainly known as opinion spammers and their activities are known as opinion spamming [5]. In most of the cases, the new sellers are taking the chance of opinion spamming to grow the business quickly. Some websites are also paying review writers to write these fake reviews [6]. As a result, detecting these deceptive reviews are necessary to maintain the trust issue between the buyer and the seller.

There are many researchers who have proposed several techniques to detect review spam. But there are some difficulties which have made this task more complicated. One of them is the unavailability of the real-life enriched labeled dataset. Most of the datasets are based on the pseudo-fake reviews which are different from the real world data. It's a major drawback when it comes down to work with the real-life data depending on the pseudo-review dataset.

In this paper, we first create a labeled dataset from real life data with the help of active learning. This solves the drawback which was discussed previously and enables us to perform an experiment on the real-life data. Traditional machine learning classifiers have been used by many researchers. But it didn't have a significant impact when it comes down to accurately detecting the real-life fake reviews. So we have introduced boosting algorithms such as Extreme Gradient Boosting (XGBoost), Adaptive Boosting (AdaBoost) and the Gradient Boosting Machine (GBM) in fake review detection side by side with the traditional machine algorithms.

Our methodology is different from the previous work for the following aspects-

- **Real life custom labeled dataset:** The dataset used in this experiment were manually prepared and labeled. Because of the lack of gold standard real-life dataset we had to go through this phase. And drastically it has made a significant impact on the accuracy of our model. From a chunk of 100,000 reviews, collected from the authors of [3, 4]. Then from the chunk, fake reviews were collected based on the writers, writing pattern, the time difference between reviews, types of reviews for a product. A keyword-based search was also done to find out keywords in reviews such as 'Fake Review', 'Review exchange for product', 'Honest Review'. These key-word based search added more versatility to the dataset.
- **Supervised boosting approach:** Supervised boosting is a new approach which is not widely applied in spam detection as traditional machine learning algorithms. Amongst all other types of classifiers such as tree-based, boosting, regression, ensemble, and neural network, boosting is on the rise among researchers.

II. RELATED WORKS

Product reviews or consumer opinions are used in business as a guide for new customers. In e-commerce, product reviews written by customers can be beneficial. Positivity from the previous customer can bring new customers into the business. And it's one of the most vital things to generate revenue. Thus sellers always prefer a positive vibe of their product so that more customers get attracted. Throughout the process, some reviews come as opinion spam.

The term fake review was stated by the authors of [7] for the first time. In their research, they categorized it into different portions. The authors of [1] have used active learning to create a hybrid dataset from yelp and ott review dataset. Their proposed model performed very well on it. On the other hand, the authors of [8] and [9] explored time series and distributional analysis. From the crowdsourced fake review data Ott, Choi [9] focused on the fictitious opinions. In their model, they have used an n-grams based classifiers to detect the review spam. On the other hand, the authors of [10] proposed a novel approach of topic modeling using Latent Dirichlet Allocation. But the fixed number of topics was the main drawback when applied to the real-world dataset. Fei, Mukherjee [11] tried a different approach to ranking. They explored the burstiness in the reviews to detect the opinion spammers. The authors of [12] tried a graph relationship between the reviews and proposed a framework named FraudEagle. On an extended work [13], they collected the meta-data of the reviews to detect the review spammers.

Ren and Ji in [14] took the analysis one step further with the introduction of a convolutional neural network (CNN) and recurrent neural network. Their study showed that neural network based models have better generalization ability than discrete models. In [15], the authors proposed a heterogeneous graph model that finds the correlation between the reviews and the reviewers. This approach doesn't use the texts taken from the reviews. Therefore it only identifies the review spammer. In [16] the authors proposed a time series approach to detect the review spam. But the model suffered from a high computational problem during the scoring phase of the reviews.

III. DATASET & ACTIVE LEARNING

As stated before, a dataset with pseudo-reviews is available for analysis. But the problem with the dataset is that the model trained on this dataset didn't perform well on real-life test data. To solve the problem we created a hybrid dataset with the help of active learning. The active learning algorithm can interactively query the user or the information for the new data points to obtain the desired outputs and to reduce the cost of manual labeling of the dataset. A learner chooses a smaller set of examples to learn the concept. The main difference here with the normal supervised learning is that the learner learns from very small data samples than the one required for the normal supervised learning. We manually labeled 200 reviews as truthful reviews and another 200 as fake reviews. Then we fed these data to the learner to label the remaining 1600 data. For the remaining 1600 data, the model labeled 875 data samples as truthful reviews and leftover 725 reviews as fake reviews. So, we ended up with a dataset of 2000 samples where 1075 data samples were labeled as truthful reviews and 925 data samples were labeled as fake. Then this dataset was split into train, test and validation set. The algorithm of active learning is given below. The algorithm is quite same with authors proposed in [1].

Algorithm: Active learning process

```
INPUT:

initialTrain = The initial labeled training set;
initialTest = The initial labeled test set;
procData = unlabeled data;

OUTPUT:
finalTrain = hybrid training set;
//finalTrain    will    consist    intialTrainData    &
newLabeledData

1.    Load procData
2.    for all instances in procData
3.      create a sparse vector using the tf-idf vectorizer
4.       mat[] = sparse vector;
         //store s. Vector in a matrix
5.      feed mat[] to the CLASSIFIER
6.      accuracy = CLASSIFIER.accuracy
7.      if accuracy > 85% then
8.          finalTrain = (finalTrain U procData)
            //assert new labeled data
9.      else
10.       unlabData = DecisionFunction(procData instance)
11.       Manually label unlabData
```

Fig 1: Active Learning algorithm

IV. BOOSTING

The authors of [17] first came up with the concept of converting the weak learners into strong learner with higher accuracy. The work was then followed by the authors of [18] and [19] who made the boosting popular. They utilized it as a functional approximation of the logistic regression. This method of solving a problem is known as gradient boosting. This approach was slightly modified by the authors of [20] who fitted a subsample of the training set of the model without replacing it. This approach was based on the minimization of the loss function of the gradient and stochastic gradient descent in regression.

In [1], the authors introduced a hybrid ensemble approach to detect Amazon fake reviews. In their approach they have used active learning to create a labeled dataset first and then supervised learning with traditional classifiers such as Naïve Bayes, SVM, Decision Tree, Maximum Entropy and majority voting for classification were used. With n-gram features, their proposed model achieved an accuracy of 88%. Our proposed model extends their work to improve the performance with boosting. The extreme gradient boosting (XGBoost) is an end to end tree boosting algorithm which was proposed by the authors of [2]. The authors proposed it

as a novel sparsity-aware algorithm for sparse data and weighted quantile like a sketch for approximate tree learning. But the derivation of XGBoost follows from the existing idea of gradient boosting, originating a second order method from [Freidman at el]. The main factor of XGBoost is the leaf wise tree growth where the gradient boosting uses level wise tree growth. Recently XGBoost has become more popular because of winning kaggle competitions and outperforming classifiers and neural network models in some cases. 17 solutions out of 29 have used XGBoost and won the competition. 8 of the models solely used XGBoost while remaining models were ensembles of XGBoost and Neural Networks. Only some ensemble classifiers have outperformed well-tuned XGBoost. But they are very small in number according to [2]. Thus, this study and exploration of the previous work and kaggle winning solutions give us an overview that XGBoost can be explored with traditional approaches.

Our proposed method focuses more on XGBoost rather than other boosting approaches and classifiers that are commonly used. Different feature selection techniques were used to compare the performance of traditional classifiers and XGBoost. XGBoost performed brilliantly when compared with the previous work. Even with the default values of the hyperparameters described in [2], it beats the accuracy of fake review detection in a big margin. And finally, we ended up with a better result than the previous works.

V. METHODOLOGY

This section elaborates the experimental procedures that have been followed. We divide the whole methodology into 5 steps.

1. Collecting the unlabeled dataset.
2. Manually labeling some data and labeling remaining with Active Learning.
3. Preprocessing the dataset.
4. Feature selection.
5. Traditional and boosting approach to detect the review spam and comparison of results.

In previous researches, authors had their own way of preparing the dataset as gold standard data for amazon fake review detection is yet troublesome to find. That is where active learning played its role. Labeling all the data manually takes a huge effort, time and cost. To minimize these, we manually labeled some data with the help of experts and then using the active learning, labeled remaining dataset. Though it can't be claimed that our dataset is the gold standard but it is moderate enough for our research. From the dataset labeled by active learning, we picked some of the data and checked manually how active learning performed. Throughout the process, we labeled 1600 data with active learning and manually 400 data were labeled before. So we had 2000 labeled data where 1075 were labeled as truthful reviews and 925 reviews were detected as fake reviews. The dataset used here was collected from Amazon raw review data used by the authors of [3, 4]. It contains 142.8 million reviews and metadata spanning from 1996 – July 2016. The details of the dataset can be found at – http://jmcauley.ucsd.edu/data/amazon

After creating the dataset, in the second phase, we preprocess the dataset. The data we collected, needed a lot of preprocessing. The preprocessing part has not been discussed here as for the textual data standard preprocessing methods were followed such as removing stop words, stemming, lemmatizing etc. After preprocessing the dataset, it was split into 70%-10%-20% for training, validation, and testing. A small portion of the training set was used for validation.

In the feature selection phase, TF-IDF (Term frequency-Inverse document frequency) and Chi2 (Chi-Square) were used. TF-IDF increases the weight of uncommon words and decreases the weight of common word and results in creating a vector of features. On the other hand, Chi2 is a statistical approach that tries to find a relationship between every feature variable and the target variable. Features that are independent, gets discarded whereas dependent features get the most importance.

After the feature selection process traditional machine learning classifiers such as Decision Tree, Support Vector Machine, Stochastic Gradient Descent was trained and test data were used to check the accuracy. At the same time XGBoost and AdaBoost, GBM classifiers were trained and validated with training and validation sets and then accuracy was measured on the training set.

To understand the performance we took the accuracy, precision, recall and f1-score measurements. As said above, simulations were run based on different feature selection techniques, supervised approach with and without boosting, with default parameters and tuned hyperparameters. The results are discussed in the experiment section. The algorithm followed throughout the experiment is given below

Algorithm: Experimental Process

```
INPUT:

TrainData = The labeled training set (70%);
ValidationData = The validation dataset (10%)
TestData = Unlabeled dataset (20%)

OUTPUT:
Predictions = prediction from classifiers used.;
//ValidationData is used to validiate the classifier
predictions

1.    Load TrainData
2.    for all instances in TrainData
3.        create different sparse vector using the tf-idf
& chi2

4.       tf_idf_mat[] = sparse vector;
5.       Chi_2_mat[] = sparse vector;
6.       combined_mat[] = combine tf_idf_chi2_matrix
7.     for each feature matrix fed to the CLASSIFIER [DT,
SVM, SGD, XGB, ADB, GBM]
8.       train classifier
9.       accuracy, precision, recall = PREDICTION.metrics
10. RESULT COMPARISON
```

Fig 2: Experimental process Algorithm

VI. EXPERIMENT & RESULTS

In this part, we have discussed the details of the experimental procedure. As said above, using the active learning algorithm and manual labeling we created a dataset of 2000 reviews where 1075 were labeled as truthful reviews and the remaining 925 were labeled as fake reviews. We kept the size small to closely observe the performance. Before feeding this data into the classifiers, we need to select features from the data. In this case, tf-idf and chi2 were used individually. At first, we selected features with tf-idf and fed the data into traditional classifiers and boosting classifiers. In the second phase, we used chi2 as feature selection algorithm and followed the exact same procedure. At the third phase, we combined both to observe how all the classifiers perform. The boosting algorithms used here has several hyperparameters that can be tuned to achieve higher accuracy. Working with huge data can be time-consuming and costly when it comes to finding the best hyperparameter set. Choosing a smaller dataset of 2000 data has given us an advantage here. As for the traditional classifiers we have used

- Decision Tree (DT)
- Support Vector Machine (SVM)
- Stochastic Gradient Descent (SGD)

These traditional classifiers and boosting classifiers were simulated first before using Grid Search to find the best hyperparameter set. The results of the traditional classifiers with different feature selection technique are given below

Table 1: Performance of Traditional Classifiers on Tf-idf features

Classifiers	Feature Selection	DT	SVM	SGD
Accuracy	Tf-idf	0.824	0.833	**0.845**
Precision		0.810	0.845	**0.868**
Recall		0.804	0.860	**0.842**

The table above shows the performance metric for traditional classifiers with tf-idf. Tf-idf is one of the most popular feature selection technique used in natural language processing problems to select features. It can be seen that SGD achieves accuracy up to 84%. After observing performance with tf-idf feature selection performance with traditional classifiers, we then try chi2 features.

Table 2: Performance of Traditional Classifiers on Chi2 features

Classifiers	Feature Selection	DT	SVM	SGD
Accuracy	Chi2	0.846	0.857	**0.882**
Precision		0.829	0.875	**0.895**
Recall		0.839	0.872	**0.888**

Table 2 shows the performance of classifiers on chi2 features. From here we can say that the accuracy, precision, recall have improved. SGD achieved the highest among all up to 88% accuracy. Chi2 features have performed better than tf-idf features when fed into the target classifiers. After observing the accuracy of chi2 features, we combine both tf-idf and chi2 as part of our experiment to observe whether the metrics improve.

Table 3: Performance of Traditional Classifiers on combined tf-idf and chi2 features

Classifiers	Feature Selection	DT	SVM	SGD
Accuracy	Tf-idf + Chi2	0.842	0.868	**0.872**
Precision		0.813	0.845	**0.885**
Recall		0.798	0.879	**0.877**

From table 3, it can be observed that the combination of tf-idf and chi2 features didn't perform well. The overall accuracy was higher when only chi2 features were used. Here now on we have discussed the results of boosting algorithms with their default parameters and best parameter set found with Grid Search. The simulation procedure is the same as the traditional classification approach keeping the training data and feature selection techniques same for comparison purposes.

Table 4: Performance of Boosting Classifiers on Tf-idf features

Classifiers	Feature Selection	XGBoost	AdaBoost	GBM
Accuracy	Tf-idf	**0.932**	0.919	0.922
Precision		**0.947**	0.901	0.917
Recall		**0.911**	0.884	0.898

Boosting classifiers achieved a notable amount of higher accuracy than the traditional machine learning classifiers. These boosting classifiers ran on their default parameter values. XGBoost achieved up to 93% which is slightly higher than others with tf-idf features. In the table below the classifiers have the same configuration of parameters. Only the features were changed to chi2 to observe the performance.

Table 5: Performance of Boosting Classifiers on chi2 features

Classifiers	Feature Selection	XGBoost	AdaBoost	GBM
Accuracy	Chi2	**0.958**	0.942	0.952
Precision		**0.951**	0.911	0.939

Recall		**0.898**	0.887	0.906

As the traditional classifiers had slight improvement of accuracy in chi2 features from tf-idf features, boosting classifiers have also performed in the same way. Using chi2 features, evaluation metrics have slightly improved results. So combining tf-idf and chi2 features to observe the change in results.

Table 6: Performance of Boosting Classifiers on combined tf-idf and chi2 features

Classifiers	Feature Selection	XGBoost	AdaBoost	GBM
Accuracy	Tf-idf + Chi2	**0.937**	0.902	0.918
Precision		**0.932**	0.895	0.927
Recall		**0.909**	0.889	0.894

The combination of features causes the accuracy to be slightly reduced than only chi2 features. The accuracy achieved here is 93%. It can be said from the observations that all the classifiers performed well with chi2 features. The below figure demonstrates the overall highest accuracy achieved in the different experimental setup. It can be seen that XGB achieves the highest accuracy with Chi2 as a feature selection technique.

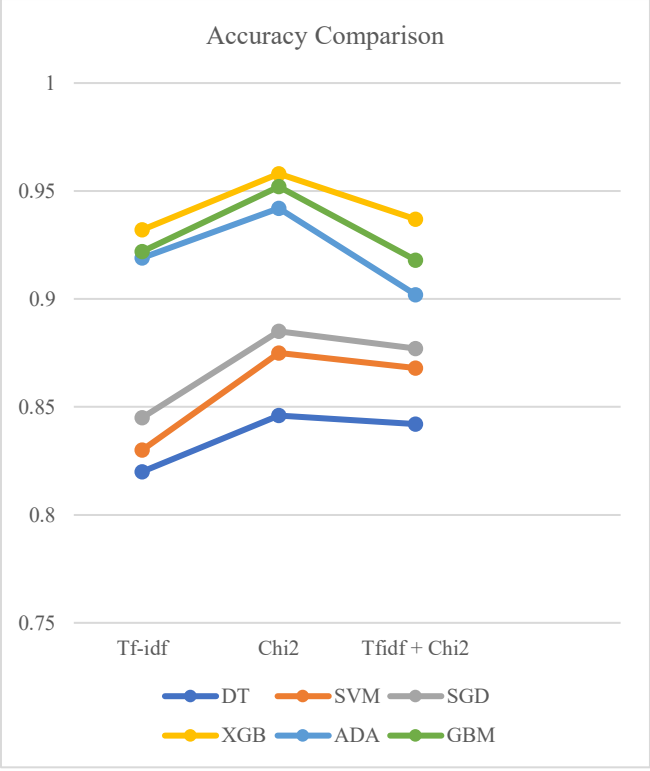


Fig 3: Accuracy Comparison

VII. CONCLUSION

As discussed earlier, we mentioned a lack of gold standard public dataset for analysis for which we had to introduce active learning. Manually labeling the data may increase the performance of the model. Using Extreme Gradient Boosting on fake review detection is still new and on the rise. There are many branches available to explore. As said, in our proposed method and experiments we didn't do deep hyperparameter tuning. Hyperparameter tuning is costly and time-consuming. Finding the best set of parameters can be very tricky. But with default parameters, it can be seen that XGBoost performed better achieving accuracy up to 95%. Finally, we hope to improve the dataset more and look forward to taking other aspects of the boosting algorithms to observe the change in results.

REFERENCES

[1] M.N. Istiaq Ahsan, Abdullah All Kafi, and Tamzid Nahian. Faisal Muhammad Shah, "An Ensemble approach to detect Review Spam using hybrid Machine Learning Technique." 2016 19th International Conference on Computer and Information Technology (ICCIT).

[2] Chen, Tianqi & Guestrin, Carlos. (2016). XGBoost: A Scalable Tree Boosting System. 785-794. 10.1145/2939672.2939785.

[3] Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering R. He, J. McAuley WWW, 2016

[4] Image-based recommendations on styles and substitutes. J. McAuley, C. Targett, J. Shi, A. van den Hengel SIGIR, 2015

[5] Algur, S., Hiremath, E., Patil, A. and Shivashankar, S., "Spam Detection of Customer Reviews from Web Pages." In Proceedings of the 2nd International Conference on IT and Business Intelligence.2010.

[6] Streitfeld, David. "Buy reviews on Yelp, get black mark." New York Times. Available:http://www.nytimes.com/2012/10/18/technology/yelp-tries-to-halt-deceptive-reviews. html. (2012)

[7] Jindal, Nitin, and Bing Liu. "Opinion spam and analysis." In Proceedings of the 2008 International Conference on Web Search and Data Mining, pp. 219-230. ACM, 2008.

[8] Xie, Sihong, et al. "Review spam detection via temporal pattern discovery."Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2012.

[9] Feng, S., Xing, L., Gogar, A., and Choi, Y. "Distributional Footprints of Deceptive Product Reviews". ICWSM. 2012

[10] Kyungyup Daniel Lee, Kyungah Han and Sung-Hyon Myaeng. Capturing Word Choice Patterns with LDA for Fake Review Detection in Sentiment Analysis. WIMS 2016. Available at: https://dl.acm.org/citation.cfm?id=2912868

[11] A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, "Fake Review Detection: Classification and Analysis of Real and Pseudo Reviews," UIC-CS-03-2013. Tech. Rep., 2013.

[12] Jindal N, Liu B (2007) Review spam detection. In: Proceedings of the 16th international conference on World Wide Web (pp. 1189–1190). ACM, Lyon, France.

[13] Pennebaker, J.W. et al., The Development and Psychometric Properties of LIWC2007 The University of Texas at Austin. ,

pp.1–22.

[14] Ren Y, Ji D. Neural networks for deceptive opinion spam detection: An empirical study. Information Sciences. 2017;385–386:213–24.

[15] Wang G, Xie S, Liu B, Yu PS. Review Graph Based Online Store Review Spammer Detection. Proceedings of the 2011 IEEE 11th International Conference on Data Mining; Vancouver, Canada. 2118325: IEEE Computer Society; 2011. p. 1242–7.

[16] Heydari A, Tavakoli M, Salim N. Detection of fake opinions using time series. Expert Systems with Applications. 2016;58:83–92.

[17] Schapire RE. The Strength of Weak Learnability. Maching Learning. 1990;5(2):197–227.

[18] Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). The annals of statistics. 2000;28(2):337–407.

[19] Friedman JH. Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics. 2001;29(5):1189–232.

[20] Friedman JH. Stochastic gradient boosting. Computational Statistics & Data Analysis. 2002;38(4):367–78.

Performance Analysis of Different Decision Tree Based Methods for Identifying Drop Out Students

Nafisa Tasnim
Dept. of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
nafisa.ruetcse16@gmail.com

Mahit Kumar Paul
Dept. of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
mahit.cse@gmail.com

A. H. M. Sarowar Sattar
Dept. of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
sarowar@gmail.com

***Abstract*—The main concern of an educational institution is to enhance the performance of students. But dropping out of students is a large obstacle that creates bad impact in this purpose. For this reason, most of the educational institutions are conscious about reducing the rate of drop out students. To reduce the drop out students, firstly it is needed to identify the reasons behind the drop out students. In this paper to address the issue, different Decision Tree (DT) based methods as well as the boosted approaches have been used. Firstly, the features of the dataset have been sorted according to their importance based on Information Gain (IG). After the features being sorted, the dataset is partitioned into three categories: the all features, two-third of the features and one-third of the features. Thereafter the standard DTs and the boosted DTs are being experimented over the partitioned dataset to identify drop out students.**

***Index Terms*—Important features, information gain, decision tree, boosted decision tree, drop out.**

I. Introduction

Education facilitates our learning, knowledge and skill. Our mind and personality is completely changed because of education. It helps us to gain positive attitude. For every nation education is a key factor in its progress. For this purpose, many educational institutes give importance for completing graduation on time. But it is an enormous challenge for educational department to improve students graduation rate. The causes for not graduating on time is diverse; including economic problems, lack of motivation, and unexpected life changes can delay students graduation or cause them to drop out [1]. It is a lengthy process if we identify the students by observing students behavior. It needs huge man power. In this case, data mining techniques help us to identify drop out students.

Data mining is the process for discovering interesting patterns from huge amount of data [2]. It has both predictive and descriptive nature. Usually supervised learning approach has predictive nature. On the other hand, unsupervised learning has descriptive nature [3]. In supervised learning method we are given input instances with their class labels for training purposes. By analyzing the given data, we can predict the outcome for the test data. To identify drop out students, we have used some decision tree based method like J48, Random Tree (RT), Random Forest (RF), Logistic Model Tree (LMT), Decision Stump (DS) and Hoeffding Tree (HT).

For identifying drop out students, we have selected most important features by using information gain. The feature with highest gain value is considered as the most important feature. According to this gain value, we have sorted the features. After sorting the features, we have considered three categories of the dataset: all features, two-third of the features and one-third of the features. Then we have implemented the tree based methods to identify drop out students.

The leavings of the paper is systematized as follows: in section II, the related tasks by many researchers have been discussed. Important features are selected in section III. Different tree based methods have been described in section IV. In section V, a brief introduction to the datasets has been given. The evaluation criteria for performance analysis has been discussed in section VI. The performance of the tree based approaches are analyzed in section VII. And in section VIII, the conclusion of our work has been provided.

II. Related Work

Nowadays identification of drop out students is a major task. Because it strains the nation backward. Besides this, it attenuates the resource of an educational institutions. For this, researchers have tried to identify drop out students using several classification algorithms. In a recent work [4], the authors have implemented several classification algorithms (Naive Bayes, Neural Network, Support Vector Machine, Decision Tree and Random Forest) and tried to compare among them. They also have generated rules using Decision Tree. The authors of the paper [5] have reported on a large-scale study. They have used several classification algorithms such as Decision Tree, Decision Table, Logistic Regression, and Naive Bayes for identifying students of not meeting acceptable levels. In [6], the authors have tried to find out who, when and why are at risk of not graduating on time. Decision Tree, Logistic Regression, Stepwise Regression, and Cox Regression model have been used for their way of task. They have suggested predictors for prioritizing risk students. In another work [7], the authors have proposed a new threshold based method to identify drop out students and compared their method with some machine learning methods such as Logistic Regression, Naive Bayes and Support Vector Machine. In a work [8], the authors have compared different tree based

methods and have shown that each tree performs better for each condition. They have concluded that choice of parameter is important for knowledge discovery. The authors of paper [9] have compared REPTree and Decision Tree algorithms. They have shown that how pruning algorithm performs better. They have also focused on the influence of REPTree on the change of accuracy and complexity. In [10], the authors have compared among Random Forest, Logistic Model Tree (LMT) and Classification and Regression Tree (CART). They have shown that the performance of all the three methods are reasonably good. Among them Random Forest outperms than CART and LMT for prediction.

III. Feature Importance

To identify the drop out students, the mostly used features should be considered. In our way to work, Information Gain (IG) has been used to extract the most important features. It is based on Claude Shannon Entropy [11]. The feature with highest information gain is referred as the most important feature. The information gain can be calculated as eq. 1 [12]

$$IG(t) = -\Sigma_{i=1}^{|c|} P(c_i) log P(c_i) + P(t)P(c_i|t) log P(c_i|t) + P(1-t)P(c_i|(1-t)) log P(c_i|(1-t)) \quad (1)$$

where, c_i represents the i-th category, $P(c_i)$ is the probability of i-th category, P(t) is the probability appears in the documents, P(1-t) is the probability that doesn't appear in the documents, $P(c_i|t)$ is the conditional probability given that t is appeared, $P(c_i|(1-t))$ is the conditional probability given that t is not appeared.

Using the eq. 1, the most important features can be found out. Using these most important features and ignoring the less important features the dimensionality of the datasets can be reduced. This is to address as high dimensional data are the most important concerns in classification problem because of space and time complexity [13].

IV. Methodologies

A. J48

J48 is an modified version of ID3 and an implementation of C4.5 algorithm in WEKA. Depth-First strategy is used for growing this tree [8]. The algorithm considers all possible tests that can split the dataset. Here information gain is used as splitting criteria. The dataset selects an attribute with highest information gain. This algorithm can handle both discrete and continuous attributes. J48 algorithm also can handle missing values. After the tree is fully constructed, this algorithm performs pruning. Because of outliers pruning is a significant step to the result.

B. Random Tree

Random Tree (RT) is an ensemble learning algorithm that generates different learners. It is a supervised classifier. From different learner it employs a bagging idea to generate a tree. Every node is split using the best split among all variables in a standard tree. Random Tree is an ensemble of tree predictors called forest. The steps of the classification are as follows: the random trees classifier gets the input feature vector, classifies it with every tree in the forest, and outputs the class label that received the majority of "votes" [8]. Random Tree is a combination of two machine learning algorithms: single model tree and random forest. Single model trees are merged with random forest for creating Random Tree. In a single model tree every single leaf holds a linear model. Random Forests are used to enhance the performance of single model tree.

C. Random Forest

Random Forest (RF) is a supervised learning method. It is an ensemble learning method for classification, regression and other tasks. In the Random Forest algorithm, the more the number of trees in the forest, it gives higher accuracy. To create trees in the forest, we need not follow the information gain or gini index for splitting criteria. Trees are formed by randomly chosen features. Prediction is made by combining (majority vote for classification or averaging for regression) the predictions of the ensemble [8]. Random Forest classifier can handle missing values. As the number of trees in the forest becomes large, the generalization error for forests converges to a limit. The generalization error depends on the strength of the individual trees.

D. Logistic Model Tree

Logistic Model Tree (LMT) is a combination of two complementary classification methods: logistic regression and tree induction [14]. To generate an LMT, firstly it needs to build a standard classification tree, using C4.5 method. Afterwards, a logistic regression model is built at every node because every node is a 'candidate leaf' during pruning. In the logistic variant information gain is used as splitting criteria. Splits are either binary or multi way. The LogitBoost algorithm is used to form an LR model in every node. The tree is pruned using a CART algorithm. LMTs perform competitively with other classifiers while being easy to interpret. But the main drawback is the time needed to build them.

E. AdaBoost.M1

AdaBoost can be used to boost the performance of a weak learner. Freund and Schapire proposed a simple and natural extension of AdaBoost for k>2 classes, which is called AdaBoost.M1 [15]. It is an ensemble technique of classification. In this technique, learners are learned with weak learners fitting simple models to the data. After that analyze the data for errors. Boosting maintains a weight for each instance. The higher the weight, the more the classifier influence the learner [16]. When an input is misclassified by an assumption, its weight is increased so that next assumption is more likely to classify it correctly. It focuses to convert weak learners into strong one. By combining the whole set at the end, it converts weak learners into better performing model. The main disadvantage of AdaBoost.M1 is that it is unable to handle

weak hypothesis with error greater than 1/2. The algorithm of AdaBoost.M1 [17] is as follows:

Algorithm 1. AdaBoost.M1

Input: sequence of m examples $((x_1, y_1), ..., (x_m, y_m))$ with labels $y_i \epsilon Y = 1, ..., k$
weak learning algorithm **WeakLearn**

Output: the final hypothesis:
$h_{fin}(x) = arg\ max_{y \epsilon Y}\ \Sigma_{t:h_t(x)=y}\ log\frac{1}{\beta_t}$

Initialization :

1: $D_1(i) = 1/m$ for all i
2: **for** $i = 1$ to T **do**
3: Call WeakLearn, providing it with the distribution D_t
4: Get back a hypothesis $h_t = X \longrightarrow Y$
5: Calculate the error of $h_t : \epsilon_t = \Sigma_{i:h_t(x_i) \neq y_i}\ D_t(i)$. If $\epsilon_t > 1/2$, then set $T = t - 1$
6: **end for**
7: Set $\beta_t = \epsilon_t/(1 - \epsilon_t)$
8: Update D_t :

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} \beta_t, & \text{if } h_t(x_i) = y_i \\ 1, & \text{otherwise.} \end{cases}$$

F. Decision Stump

Decision Stump (DS) is a prediction based machine learning algorithm. The algorithm is based on one-level decision tree [18]. Decision stump is a kind of decision tree in which only one rule is used. One internal node (Root node) is connected with only the terminal node. This algorithm predicts using only one input feature. So, it is also called 1-rules method.

G. Hoeffding Tree

Hoeffding tree (HT) induction uses information gain as splitting criteria [19]. It is an incremental decision tree induction method capable of learning from huge amount of data. In Hoeffding tree, Hoeffding bound is used because it holds the true distribution generating values, and depends only on the range of values, number of observations and desired confidence.

V. Dataset Description

The dataset "Student Performance Analysis" has been taken from the renowned UCI Machine Learning Repository [4]. This dataset contains 33 features such as: final grade, second period grade, first period grade, number of school absences, number of past class failures, going out with friends, wants to take higher education and so on. The dataset has 395 instances. The dataset is used to assess the performance of different tree based methods (J48, Boosted J48, Random Tree, Boosted Random Tree, Random Forest, Boosted Random Forest, LMT, Boosted LMT, Decision Stump, Boosted Decision Stump, Hoeffding Tree and Boosted Hoeffding Tree) for identifying drop out students. We have used WEKA for executing the methods.

VI. Evaluation Metrics

To assess the performance of our work, different metrics such as accuracy, precision [20], recall [20], F1-score [20] and ROC Area [21], PRC Area [22], MCC [22], Mean Absolute Error [22], Root Mean Squared Error [22] have been used.

A. Precision, Recall and F1 measure

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1measure = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

where,TP = True Positive, means that the number of students who can continue study are correctly classified, FP = False Positive, means that number of drop out students are wrongly classified as the student who can continue studying. FN = False Negative, means that the number of students who can continue study are wrongly classified.

B. ROC Area

In a ROC curve, the true positive rate is plotted in function of the false positive rate. The accuracy of the test depends on how well the test separates the group. If the ROC Area is 1, then it represents perfect test and if the area is 0.5, then it is worthless.

C. PRC Area

In PRC curve, precision is plotted in function of recall. PRC is more useful when someone is only interested how the classifier is behaving in only one class. It is more informative than ROC when evaluating binary classifications on imbalanced data.

D. MCC

Matthews Correlation Coefficient (MCC) is used as a measure of the quality of binary classification. It considers true and false positives and negatives for measuring the correlation coefficient. It returns a value between -1 and +1. A +1 represents a perfect prediction, 0 represents no better than random prediction and -1 represents total disagreement between prediction and observation.

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{5}$$

where, TN = True Negative, means that number of drop out students are correctly classified.

E. Mean Absolute Error

Mean absolute error (MAE) is used to measure how close the predictions are to the actual outcomes. The mean absolute error is an average of the absolute errors. The mean absolute error is calculated by

$$MAE = \frac{1}{n}\Sigma_{i=1}^{n}|f_i - y_i| = \frac{1}{n}\Sigma_{i=1}^{n}|e_i| \tag{6}$$

Where, f_i is the predicted value and y_i is the actual value.

F. Root Mean Squared Error

Root mean squared error (RMSE) represents the standard deviation of the differences between the predicted value and the actual value. It can be calculated as

$$RMSE = \sqrt{\frac{\Sigma_{t=1}^{n}(f_i - y_i)^2}{n}} \quad (7)$$

VII. Performance Analysis

At first, we have calculated information gain (IG) of the features. Feature with highest information gain is considered as most important feature. After calculating the information gain, features have been sorted according to their gain value. The feature with highest information gain is considered first and the rest of the features are considered in the same way. After sorting the features according to the IG value the dataset is divided into three different categories. We have used all features, two-third of the features with highest information gain and one-third of the features with highest information gain for this purpose. Then these three categories of dataset have been considered to analyze the performance of different tree based algorithm. The overall flow of our work is depicted in Fig 1.

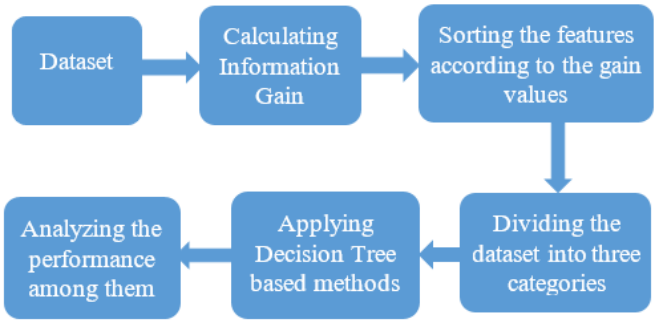


Fig. 1. Flow of the Work

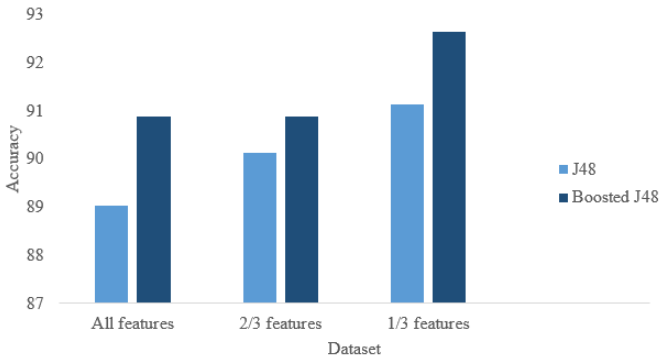


Fig. 2. Accuracy for J48 and Boosted J48

Accuracy for J48 and Boosted J48 are shown in Fig. 2. From this figure, it can be seen that Boosted J48 performs always better than J48. The performance for one third of the features is better than two third of the features and the performance for two third features is better than all features. Table I represents the comparative performance between J48 and Boosted J48.

TABLE I
Comparative analysis between J48 and Boosted J48

Evaluation Metrics	All features		2/3 features		1/3 features	
	J48	***Boosted J48***	***J48***	***Boosted J48***	***J48***	***Boosted J48***
Precision	0.895	0.909	0.901	0.908	0.911	0.926
Recall	0.896	0.909	0.901	0.909	0.911	0.927
F1 measure	0.896	0.909	0.901	0.908	0.911	0.926
ROC Area	0.878	0.964	0.876	0.957	0.926	0.972
PRC Area	0.867	0.960	0.869	0.951	0.912	0.973
MCC	0.727	0.762	0.741	0.759	0.767	0.808
MAE	0.115	0.091	0.113	0.089	0.096	0.074
RMSE	0.296	0.289	0.290	0.289	0.268	0.262
Time (sec)	0.02	0.11	0.02	0.06	0.01	0.05

From all the assessment criteria, it can be seen that Boosted J48 outperforms J48. Using one third features that means using most important features performance is better than other two cases.

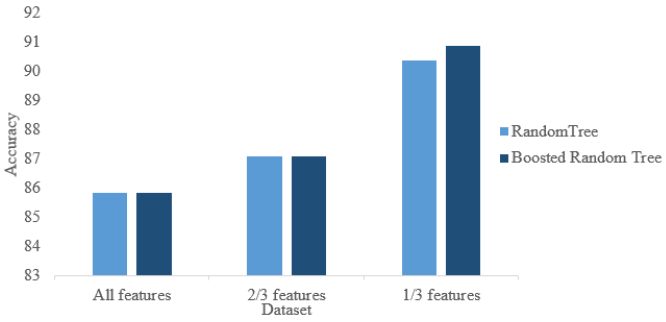


Fig. 3. Accuracy for Random Tree and Boosted Random Tree

TABLE II
Comparative analysis between Random Tree and Boosted Random Tree

Evaluation Metrics	All features		2/3 features		1/3 features	
	RT	***Boosted RT***	***RT***	***Boosted RT***	***RT***	***Boosted RT***
Precision	0.855	0.854	0.868	0.873	0.903	0.908
Recall	0.858	0.858	0.871	0.876	0.904	0.909
F1 measure	0.856	0.855	0.868	0.874	0.903	0.908
ROC Area	0.808	0.794	0.812	0.836	0.876	0.874
PRC Area	0.884	0.803	0.818	0.836	0.871	0.870
MCC	0.621	0.617	0.653	0.668	0.746	0.759
MAE	0.141	0.147	0.131	0.124	0.096	0.092
RMSE	0.374	0.378	0.359	0.347	0.305	0.302
Time (sec)	0.02	0.01	0.01	0	0	0

Fig. 3 shows that both Random Tree and Boosted Random Tree have equal accuracy for all features and two-third features. It can be also seen that, using most important features the result is outstanding than using all features. From Table II, it can be observed that Boosted Random Tree performs better than Random Tree.

Bar graph in Fig. 4 shows the accuracy comparison between Random Forest and Boosted Random Forest. Though

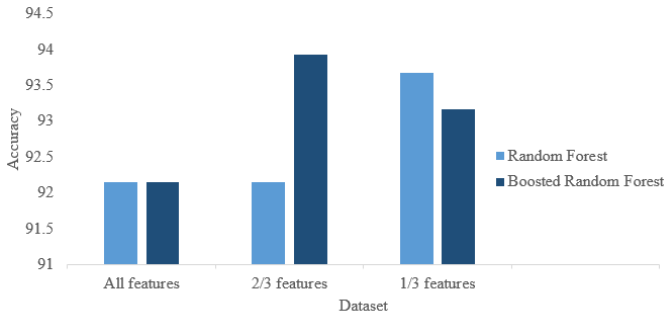


Fig. 4. Accuracy for Random Forest and Boosted Random Forest

TABLE III
COMPARATIVE ANALYSIS BETWEEN RANDOM FOREST AND BOOSTED RANDOM FOREST

Evaluation Metrics	All features		2/3 features		1/3 features	
	RF	*Boosted RF*	*RF*	*Boosted RF*	*RF*	*Boosted RF*
Precision	0.920	0.928	0.921	0.939	0.936	0.932
Recall	0.922	0.929	0.922	0.939	0.937	0.932
F1 measure	0.920	0.928	0.921	0.939	0.936	0.932
ROC Area	0.971	0.974	0.974	0.975	0.976	0.977
PRC Area	0.972	0.974	0.975	0.976	0.977	0.978
MCC	0.791	0.811	0.792	0.840	0.833	0.822
MAE	0.166	0.162	0.145	0.144	0.116	0.116
RMSE	0.255	0.249	0.244	0.242	0.235	0.234
Time (sec)	0.09	0.09	0.06	0.08	0.06	0.08

performance of Boosted Random Forest is lower for one third features, it outperforms Random forest in a larger scale for two third features. Using different evaluation metrics we can also see the same results from Table III.

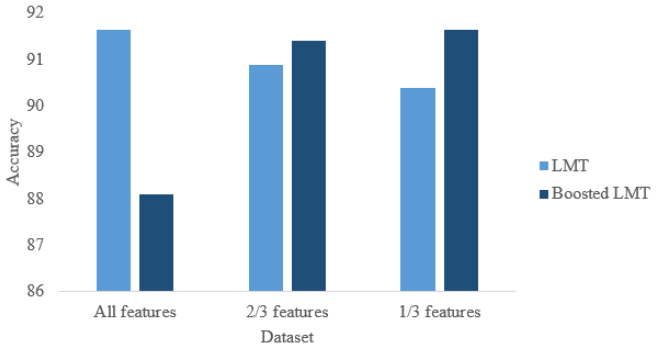


Fig. 5. Accuracy for LMT and Boosted LMT

The overall evaluation summary of both LMT and Boosted LMT is shown in Fig. 5 and Table IV. From these, it can be seen that Boosted LMT beats LMT for two-third features and one-third features. In case of Boosted decision tree performance using most important features exceeds the other cases.

From Fig. 6 and Table V, it can be seen that Boosted Decision Stump identifies drop out students more accurately than standard Decision Stump algorithm. Though the accuracy for standard Hoeffding Tree is better than Boosted Accuracy,

TABLE IV
COMPARATIVE ANALYSIS BETWEEN LMT AND BOOSTED LMT

Evaluation Metrics	All features		2/3 features		1/3 features	
	LMT	*Boosted LMT*	*LMT*	*Boosted LMT*	*LMT*	*Boosted LMT*
Precision	0.915	0.881	0.907	0.913	0.903	0.917
Recall	0.916	0.881	0.909	0.914	0.904	0.916
F1 measure	0.915	0.881	0.908	0.914	0.903	0.917
ROC Area	0.977	0.946	0.972	0.961	0.964	0.963
PRC Area	0.979	0.948	0.971	0.962	0.969	0.957
MCC	0.776	0.690	0.757	0.774	0.746	0.783
MAE	0.139	0.121	0.139	0.09	0.133	0.079
RMSE	0.239	0.341	0.247	0.291	0.254	0.274
Time (sec)	0.31	14.35	0.27	9.39	0.19	3.11

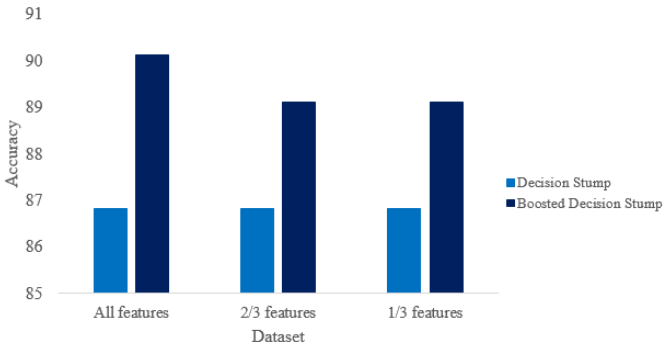


Fig. 6. Accuracy for Decision Stump and Boosted Decision Stump

TABLE V
COMPARATIVE ANALYSIS BETWEEN DECISION STUMP AND BOOSTED DECISION STUMP

Evaluation Metrics	All features		2/3 features		1/3 features	
	DS	*Boosted DS*	*DS*	*Boosted DS*	*DS*	*Boosted DS*
Precision	0.897	0.899	0.897	0.889	0.897	0.889
Recall	0.868	0.901	0.868	0.891	0.868	0.891
F1 measure	0.874	0.9	0.874	0.889	0.874	0.889
ROC Area	0.886	0.96	0.886	0.960	0.886	0.960
PRC Area	0.865	0.959	0.865	0.961	0.865	0.961
MCC	0.712	0.736	0.712	0.706	0.712	0.706
MAE	0.179	0.121	0.179	0.122	0.179	0.122
RMSE	0.311	0.277	0.311	0.279	0.311	0.279
Time (sec)	0.01	0.03	0.01	0.02	0.01	0.01

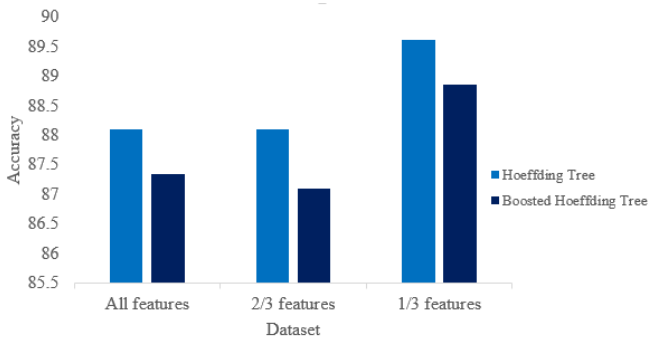


Fig. 7. Accuracy for Hoeffding Tree and Boosted Hoeffding Tree

TABLE VI
COMPARATIVE ANALYSIS BETWEEN HOEFFDING TREE AND BOOSTED HOEFFDING TREE

Evaluation Metrics	All features		2/3 features		1/3 features	
	HT	***Boosted HT***	***HT***	***Boosted HT***	***HT***	***Boosted HT***
Precision	0.883	0.872	0.883	0.868	0.897	0.888
Recall	0.881	0.873	0.881	0.871	0.896	0.889
F1 measure	0.882	0.873	0.882	0.869	0.897	0.888
ROC Area	0.931	0.927	0.940	0.914	0.943	0.941
PRC Area	0.916	0.925	0.924	0.917	0.927	0.939
MCC	0.695	0.666	0.695	0.656	0.732	0.707
MAE	0.141	0.140	0.132	0.145	0.125	0.127
RMSE	0.312	0.315	0.295	0.318	0.285	0.290
Time (sec)	0.02	0.16	0.01	0.08	0.01	0.07

from Fig. 7 and Table VI it can be found that it performs better when we use most important features.

From the above it can be perceived that, Boosted trees outperform than standard trees. Because from a number of weak classifiers Boosted tree creates a strong classifier. Boosted algorithm builds a model from training data. After that it creates a second model that attempts to correct the error from first model. The method continues until the training set is predicted perfectly.

The performance is also higher for the most important features than using whole features. From all the tables, it can be seen that with the number of features are reducing the time taken for model building is also reducing. So, when only important features will be used, it will save time for building a model.

VIII. CONCLUSION

In this paper, different tree based methods are compared with their boosted approaches for identifying drop out students. For identifying drop out students important features are calculated. Features with highest gain value are considered as the most important features. To assess the performance of the methods, several evaluation criteria are used. J48, Random Tree, Random Forest, LMT, Decision Stump, Hoeffding Tree and their corresponding Boosted Trees are used for this purpose. Finally, it is observed that for most of the cases performance is higher for one third of the features, i.e., the most important features and Boosted DTs outperforms than the standard DTs. In future, a comparison with the recent tree based method such as XGBoost can be performed.

REFERENCES

[1] A. Bowers, and R. Sprott, "Why tenth graders fail to finish high school: a dropout typology latent class analysis," *Journal of Education for Students Placed at Risk (JESPAR)* , vol. 17, no. 3, pp. 129-148, 2012.

[2] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "Knowledge discovery and data mining: Towards a unifying framework," in the 2nd *International Conference on Knowledge Discovery and Data Mining (KDD)*, Aug. 1996, pp. 82-88.

[3] C. J. Carmona, P. Gonzales, M. J. Jesus, and F. Herrera, "NMEEF-SD: Non-dominated Multi-objective Evolutionary algorithm for Extracting Fuzzy Rules in Subgroup Discovery," in *IEEE international conference on fuzzy systems*, pp. 1706-1711, 2010.

[4] P. Cortez and A. Silva, "Using data mining to predict secondary school student performance," in Proceedings of 5th *FUture BUsiness TEChnology Conference (FUBUTEC 2008)*, pp. 5-12, Porto, Portugal, EUROSIS, ISBN 978-9077381-39-7, April 2008.

[5] A. Tamhane, S. Ikbal, B. Sengupta, M. Duggirala and J. Appleton, "Predicting Student Risks Through Longitudinal Analysis," in Proceedings of the 20th *ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1544-1552, New York, New York, USA August 24 - 27, 2014.

[6] E. Aguaiar, H. Lakkaraju, N. Bhanpuri, D. Miller, B. Yuhas and K. L. Addison, "Who, When, and Why: A Machine Learning Approach to Prioritizing Students at Risk of not Graduating High School on Time," Proceedings of the Fifth *International Conference on Learning Analytics And Knowledge*, pp. 93-102, Poughkeepsie, New York March 16 - 20, 2015.

[7] N. Tasnim, M. K. Paul, and A. H. M. S. Sattar, "Identification of Drop Out Students in Educational Data Mining," in 2nd *International Conference on Electrical, Computer and Communication Engineering (ECCE)*, pp. 1-5, February 09-11, 2019.

[8] Y. Zhao, and Y. Zhang, "Comparison of Decision Tree Methods for finding active objects," *National Astronomical Observation*, vol. 41, pp. 1955-1959, 2008.

[9] W. N. H. W. Mohamed, M. N. M. Salleh, and A. H. Omar, "A Comparative Study of Reduced Error Pruning Method in Decision Tree Algorithms," In *IEEE International Conference on Control System, Computing and Engineering*, 23 - 25 Nov. 2012.

[10] W. Chen, X. Xie, J. Wang, B. Pradhan, H. Hong, D. T. Bui, Z. Duan, and J. Ma, "A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility," Catena 151, pp. 147-160, 2017.

[11] J. Han, M. Kamber, and J. Pei, "Data mining concepts and techniques," 3rd ed., vol. 2. by *Elsevier*, 2012, pp.336–340.

[12] H. Ugus, "A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm," in *Knowledge-Based Systems*, vol. 24, issue 7, pp. 1024-1032, October 2011.

[13] A. K. Uysal, "A novel probabilistic feature selection method for text classification," in *Knowledge-Based Systems*, vol. 36, pp. 226-235, December 2012.

[14] M. Samner, E. Frank, and M. Hall, "Speeding up Logistic Model Tree Induction," In *European Conference on Principles of Data Mining and Knowledge Discovery. Springer*, Berlin, Heidelberg.

[15] E. A. Cortes, M. G. Martinez, and N. G. Rubio, "Multiclass Corporate Failure Prediction by Adaboost.M1," Int *Adv Econ Res*, (2007) 13: 301.

[16] J. R. Quinlal, "Bagging, Boosting, and C4.5", In *AAAI/IAAI*, Vol. 1, pp. 725-730. 1996.

[17] Y. Freund, and R. E. Schapire, "Experiments with a new boosting algorithm," in *icml*, 1996 Jul 3 (Vol. 96, pp. 148-156).

[18] W. Iba, and P. Langley, "Induction of One-Level Decision Trees," in *Machine Learning Proceedings*, pp. 233-240, 1992, Morgan Kaufmann.

[19] G. Holmes, R. Kirkby, and B. Pfahringer, "Stress-Testing Hoeffding Trees," in *European conference on principles of data mining and knowledge discovery*, pp. 495-502, Springer, Berlin, Heidelberg, 2005.

[20] T. Fawcett, "An introduction to ROC analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861-874, 2006.

[21] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern recognition*, vol. 30, no. 7, pp. 1145-1159, 1997.

[22] K. H. Walse, R. V. Dharaskar, and V. M. Thakare, "A study of human activity recognition using AdaBoost classifiers on WISDM dataset," *The Institute of Integrative Omics and Applied Biotechnology Journal*, 2016 Jan 1;7(2):68-76.

Student's Academic Performance Evaluation Method Using Fuzzy Logic System

Abdul Aziz, Md. Asaf-uddowla Golap and M. M. A. Hashem

Department of Computer Science and Engineering
Khulna Univeresity of Engineering & Technology, Khulna-9203, Bangladesh
abdulaziz@cse.kuet.ac.bd, asaf.golap@kuet.ac.bd, hashem@cse.kuet.ac.bd

Abstract— **For evaluating student's performance, the existing methods are not so much fair and satisfiable. The ranges of marks used for grading here are sometimes blessing and sometimes not. A method is proposed to evaluate performance based on Fuzzy Inference System (FIS). This system considers the student's continuous assessment which includes the attendances, total times he/she is spent in class and marks obtained in class test. This method increases the fairness and excellency of a student's assessment than existing methods. In existing FIS for grading the continuous assessment of student is not fuzzified where only the final exam papers mark is fuzzified. But, the proposed system fuzzified both the continuous assessment and exam paper's mark which actually increase more precise justification of student performances.**

Keywords— Fuzzy Inference System, Expert System, Fuzzification, Defuzzification, Continuous Assessment.

I. INTRODUCTION

Student performance evaluation is one of the most important tasks in an educational institute. Proper evaluation method ensures fair evaluation what is highly desired and student grade should comply with performances. Student's performance depends on various factor such as attendance, class test, final exam etc. Considering, these factors final grade is calculated. In many universities only, number of class simply is considered for evaluating attendance. However, both number of class as well as time spent on class should be considered in the evaluation process, as class size is large in most of the universities in developing countries like Bangladesh and student sometimes do not attend the class of full time. For this purpose, a specialized hardware is used which scores number of class and time spent on the class.

Examinations are conducted in most of the universities, college and institutes of a subject in a subjective manner. In these examinations' students ask to write either short or detail answer which is indicated in the exam paper and depending on allotted time of that specific question. Usually, to solve 100 marks exam paper maximum 3 to 4 hours is given as a whole but time is not specified for each question, overall time management depends on students. Finally, performance is computed from their examination, class test, attendance etc. and each of these sections has a predefined weight.

Student educational performance is measured based on the results of examinations and usually expressed numerically or by later grades [1]. Conventional method compares student performance with predefined measurement criteria. Evaluation and measurement are vital parts of the educational process. Various methods are using for evaluating student performance.

A. Related Work

In 1965, fuzzy set theory was introduced by Lotfi A. Zadeh [2]. In recent years, it has been used extensively in a number of research fields including educational system to remove vagueness that conventional methods unable to solve properly. Echauz and Vachtsevanos in [3] proposed a system using fuzzy set theory for converting conventional result into later-grades. Law in [4] developed a fuzzy educational grading system structure model and proposed an algorithm to aggregate various test scores to generate a final single score of a student. It also discussed the method to build the membership functions of several linguistic values with different weights. Bai and Chen in [5] proposed a scheme to generate membership function of fuzzy rules automatically. Then this membership function used in fuzzy logic to obtain the marks of a student. It devised an unbiased and smooth scheme for evaluating students answer scripts. Then Biswas offered a method in [6] to evaluate student performance using fuzzy logic and fuzzy inference system. It took too much time to predict performance and there was chance to assign same grade for to different fuzzy marks. Later, Chen and Lee in [7] present a method which remove the limitation of the previous scheme by Biswas. Their method is faster enough and do not need any complicate matching operations.

Wilson et. al in [8] proposed a method using fuzzy set theory and genetic algorithm. Ramjeet and Vijendran in [9] consider student performance evaluation problem as a clustering problem where clusters (classes) are generated from student talent level and no. of cluster should be less than or equal to the predefined capacity. In [10] Arya et. al designed a dynamic model mainly for engineering students which take final exam marks as well as lab and project performance into account to devise final grade applying fuzzy model. Shilpa in [11] used fuzzy logic to measure a student performance considering vagueness of question paper along with accuracy rate, relative importance and complexity. Reference [12] represents a method by James using fuzzy expert system to evaluate student writing samples with high level of accuracy comparable to the best teacher graders. Ibrahim in [13] proposed an easy, transparent, fair, automatic method taking the difficulty, the importance and

the complexity of questions in account applying fuzzy inference system to provide student learning achievements. Wang and Chen in [14] presented a method for evaluating students' answer scripts using fuzzy numbers associated with degrees of confidence of the evaluator. Bardul proposed a method in [15] to evaluate the student's performance as an individual and as a group. This focus of this method is to evaluate student's answer sheet and group performance assessment of a student.

B. Basic Fuzzy Logic

Fuzzy logic was introduced by Prof. Lotif Zadeh to handle imprecision, vagueness and uncertainty of data. On the basis of working procedure, a fuzzy controller is similar to a traditional system. It takes inputs then carry out calculations and finally provide an output value. Fig. 1 illustrates the basic structure of a fuzzy system. Fuzzification is a process of converting crisp input values into fuzzy values using fuzzy membership functions. Defuzzification is the reverse process of fuzzification which converts this latter output into crisp values again.

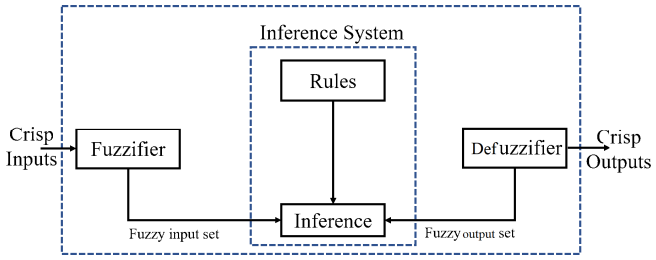


Fig. 1 Basic Fuzzy System

C. Motivation

Existing grading system is strict. This grading system has some predefined range of marks defined by University Grants Commission (UGC) of Bangladesh and a grade is awarded (e.g. A+, A, A-, B+, B, B-, C+, C, D, F etc.) equivalent to an integer interval such as 80-100 (A+), 75-(<80) (A), 70-(<75) (A-), 65-(<70)(B+), 60-(<65) (B), 55-(<60) (B-), 50-(<55) (C+), 45-(<50) (C), 40-(<45) (D), <40 (F) etc. The mark is simply obtained from different sections as a mean (class test, attendance, final exam etc). In this method same grade is awarded for a range of marks, for example if a student gets 80 and another student gets 100 percent of mark both of them will get the same grade A+, it is not fair. Therefore, an evaluation system is demanded which will provide fairness, transparency, objectivity, logical reasoning, and easy computer implementation.

Unlike existing grading system in fuzzy logic system every single mark is considered for calculating final grade as every mark has a degree of membership value in membership function. Although performance assessment process using fuzzy logic is intricate, it is flexible, whereas traditional grading system stick to fixed mathematical calculation.

D. Contribution

This paper proposes a fair method to evaluate student performance using fuzzy logic and fuzzy inference system. We summarize our contributions in this paper as follows:

- We consider number of classes as well as total time spent on the class. In credit system there is a direct relation between class time and credit. So, we have considered the total time spent on class for assessment.
- Continuous assessment (attendance, time spent on class and Class Test) and Final examination both are vital for measuring student performance. Our system considers both of these for assessing student's performance.
- Comparison between existing method and our proposed method has been done.

E. Organisation

This research paper forms as: section II which discusses the overall methodology how does the system work. Section III contains the performance analysis of students and results are compared with traditional method. In section IV, the conclusion of the research work is added.

II. METHODOLOGY

1 Proposed System for Evaluating Students Performance

In Bangladeshi Universities, for evaluating student's performance in each subject the mark is divided into two parts. So, our system consists of two parts. First part is student's continuous assessment which includes class test, attendance and total time spent in class. Second part is semester final examination's where exam paper is separated into two parts named Part-A and Part-B. The mark distribution for Part-A and Part-B is equal. The overall architecture of the proposed system is shown in Fig. 2.

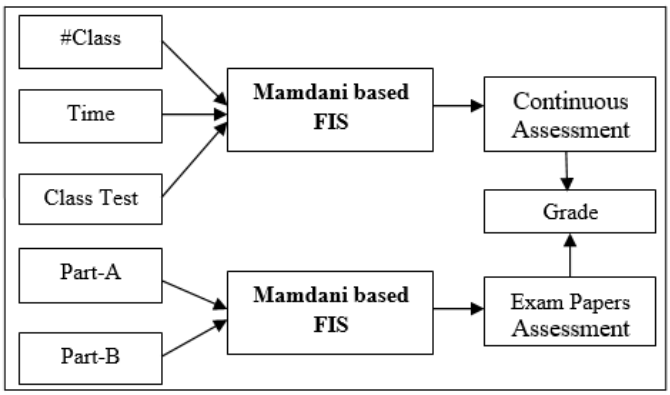


Fig. 2 Proposed System Architecture

1.1 Students Continuous Assessment System

In existing Student's Continuous Assessment System, the total no. of class and Class Test (CT) is considered which is not very much fair. But in our proposed system the total amount of time a student spent on class room is also considered which makes the assessment process more fair than existing system.

1.1.1 Fuzzification of Students Continuous Assessment

In this part, we have three linguistic input variables called Total number of class (#Class) and Total time spent in class (Time) and Class_Test. The linguistic values and interval for the variables are shown in the TABLE I, TABLE II AND TABLE III.

TABLE I: LINGUISTIC VALUES AND INTERVALS FOR INPUT VARIABLE "#CLASS"

Linguistic Values	Intervals
Few (F)	(0, 0, 0.20, 0.30)
Average (A)	(0.20, 0.30, 0.45, 0.55)
High (H)	(0.45, 0.55, 0.70, 0.80)
Very High (VH)	(0.70, 0.80, 1, 1)

TABLE II: LINGUISTIC VALUES AND INTERVALS FOR INPUT VARIABLE "TIME"

Linguistic Values	Intervals
Low (L)	(0,0, 0.20, 0.40)
Average (AV)	(0.20, 0.40, 0.60, 0.80)
Large (LG)	(0.60, 0.80, 1, 1)

TABLE III: LINGUISTIC VALUES AND INTERVALS FOR INPUT VARIABLE "CLASS_TEST"

Linguistic Values	Intervals
Very Low (VL)	(0,0,0.15,0.25)
Low (L)	(0.15, 0.25, 0.35, 0.45)
Average (A)	(0.35, 0.45, 0.55, 0.65)
High (H)	(0.55, 0.65, 0.75, 0.85)
Very High (VH)	(0.75, 0.85,1,1)

The trapezoidal membership function is used for all the input variables "#Class", "Time" and "Class_Test" as shown in Fig. 3, Fig. 4 and Fig. 5 respectively. The output variable for this part is assessment whose linguistic value is as same as input variable Class_Test shown in TABLE III and Fig. 5.

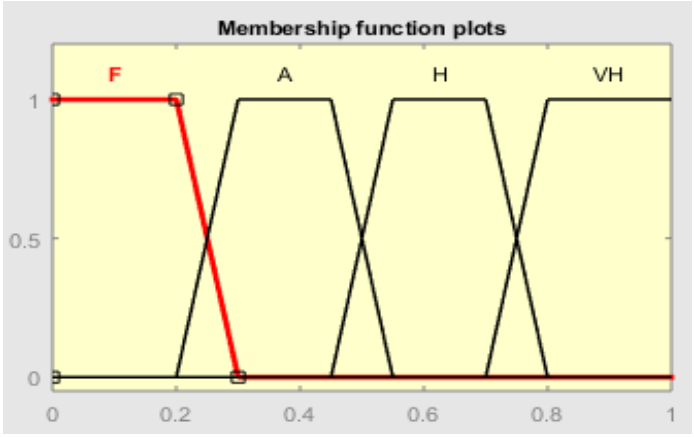


Fig. 3 Membership Function for #Class

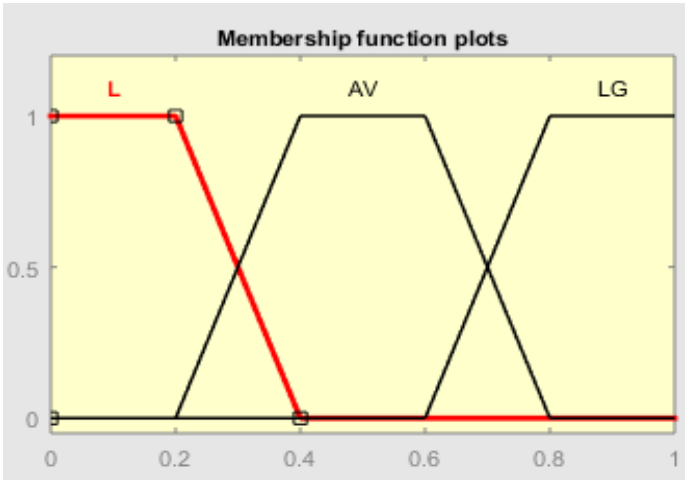


Fig. 4 Membership Function for Time



Fig. 5 Membership Function for Class_Test

1.1.2 Rules Evaluation for Assessment

The Fuzzy inference system uses the "IF Then Else" rules. For each of the input variables value the rules generated as follows.

1. If (#Class is F) and (Time is L) and (Class_Test is VL) then (Assessment is VL)
2. If (#Class is F) and (Time is L) and (Class_Test is LW) then (Assessment is VL)
3. If (#Class is F) and (Time is L) and (Class_Test is AG) then (Assessment is VL)
4. If (#Class is F) and (Time is L) and (Class_Test is HH) then (Assessment is L)
5. If (#Class is F) and (Time is L) and (Class_Test is VHH) then (Assessment is L)

……..

58. If (#Class is VH) and (Time is LG) and (Class_Test is AG) then (Assessment is H)
59. If (#Class is VH) and (Time is LG) and (Class_Test is HH) then (Assessment is VH)
60. If (#Class is VH) and (Time is LG) and (Class_Test is VHH) then (Assessment is VH)

The degree of membership value is generated by the trapezoidal membership function for each of the linguistic value of an input variable. In Fuzzy rules there are more than one value exists. So, a single value is need to choose. There are several ways to choose it. One is MIN rule and another one is MAX rule. The MIN rules select the minimum values among the degree of membership values which are ANDed by each other's. The MAX rules select the maximum values among the degree of membership values which are ORed by each other's.

1.1.3 Determining the Assessment Score

There are many others ways to calculate the final assessment score which is called defuzzification. In this paper we have used center of gravitation (COG) method or center of area calculating method.

1.2 Students Exam Papers Assessment System

1.2.1 Fuzzification of Student's Exam Papers

The answer script or the Exam paper is divided into two parts as Part-A and Part-B. So, in this part we have two linguistic input variables Part-A and Part-B. The linguistic value and interval for both of these variables are shown in TABLE IV.

TABLE IV: INPUT VARIABLE AND IT'S INTERVAL FOR "PART-A" AND "PART-B"

Linguistic Value	Interval
Very Low (VL)	(0,0, 0.02, 0.18)
Low (L)	(0.02,0.18,0.22,0.38)
Average(A)	(0.22,0.38,0.42,0.58)
High (H)	(0.42,0.58,0.62,0.78)
Very High (VH)	(0.62,0.78,0.82,0.98)
Very Very High (VHH)	(0.82,0.98,1,1)

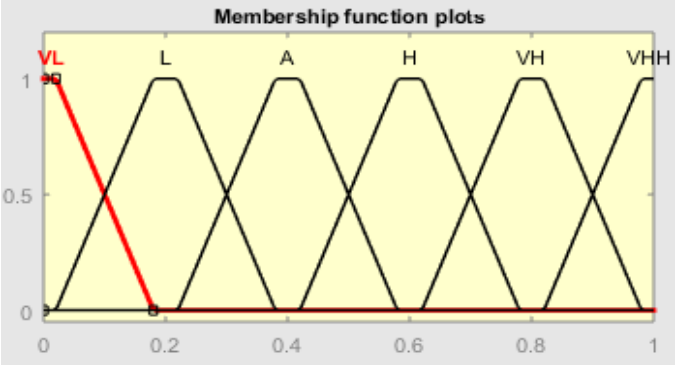


Fig. 6 Membership Function for Part-A and Part-B

Assessment is the output variable in this part whose linguistic variable and interval is same as input variable Part-A and Part-B as shown in TABLE IV and Fig. 6.

1.2.2 Rules for Performance of Two Variable

The same "IF Then Else" rules are used to measure the performance of student.

1. If (Part-A is VL) and (Part-B is VL) then (Performance is VL)
2. If (Part-A is VL) and (Part-B is L) then (Assessment is VL)
3. If (Part-A is VL) and (Part-B is A) then (Assessment is L)
4. If (Part-A is VL) and (Part-B is H) then (Assessment is L)
5. If (Part-A is VL) and (Part-B is VH) then (Assessment is A)

.....

31. If (Part-A is VHH) and (Part-B is VL) then (Assessment is A)
32. If (Part-A is VHH) and (Part-B is L) then (Assessment is H)
33. If (Part-A is VHH) and (Part-B is A) then (Assessment is VH)
34. If (Part-A is VHH) and (Part-B is H) then (Assessment is VH)
35. If (Part-A is VHH) and (Part-B is VH) then (Assessment is VHH)
36. If (Part-A is VHH) and (Part-B is VHH) then (Assessment is VHH)

The MIN and MAX rules are also used in this part for selecting only one value from several values.

1.2.3 Determining the Assessment Result

Same method (COG) is used for determining exam paper Assessment for a student which is used for determining the assessment score.

III. PERFORMANCE ANALYSIS

A. Dataset

Actually, our system is divided into two parts for evaluating student's performance for considering the student's continuous assessment related data and final exam paper data. In continuous assessment data, there are three fields named number of class, time spent in class and class test. For exam paper, also considered two fields named marks of exam papers Part-A and Part-B. The benchmark dataset is collected from Khulna University of Engineering & Technology (KUET), Khulna-9203, Bangladesh.

B. Result Analysis

For evaluating overall performance, we have used Mamdani type fuzzy inference engine which we have described briefly in methodology Section II. The performance of 15 out of 60 students in 5 individual subjects based on their benchmark dataset collected from Computer Science and Engineering Department, KUET is shown in TABLE V. Total number of class is considered 39 for each course and length of every class lecture is 50 minutes. So, total 50x39 = 1950 minutes (Time) a student can be spent in class. The final grade of an individual subject is calculated considering 30% marks in continuous assessment and 70% marks in exam papers evaluation. Then, the total marks obtained from continuous assessment and exam papers assessment is added to get the grade.

In Fig. 7 shows the assessment score is 0.9 for a student whose attendance is 94.9%, time spent in class is 96.8%. and class test is 96.67%. Similarly, Fig. 8 shows the Assessment

value is 0.8 for a student whose exam paper marks are 78.1% and 67.6% in Part-A and Part-B respectively.

TABLE V: ASSESSMENT OF STUDENT BY PROPOSED METHOD

SI No.	# Class	Time Spent on Class (Minute)	Class Test	Continuous Assessment	Part-A Mark	Part-B Mark	Exam Paper Assessment	Grade
1	37	1790	58	0.900	82	71	0.800	A+
2	28	660	60	0.888	52	67	0.636	A-
3	26	1000	40	0.700	59	54	0.566	B
4	25	1200	28	0.326	56	55	0.533	C
5	32	1400	10	0.338	49	76	0.667	B-
6	20	952	35	0.494	73	69	0.695	B
7	29	1400	18	0.444	58	57	0.557	C+
8	34	1400	55	0.900	62	63	0.600	B+
9	16	672	44	0.500	64	55	0.600	B-
10	15	500	33	0.300	48	57	0.547	C
11	12	444	23	0.176	73	66	0.695	C+
12	36	1303	39	0.813	69	73	0.695	A-
13	08	350	29	0.300	42	69	0.600	C+
14	35	1700	51	0.900	21	44	0.215	D
15	20	1144	10	0.151	32	76	0.505	F

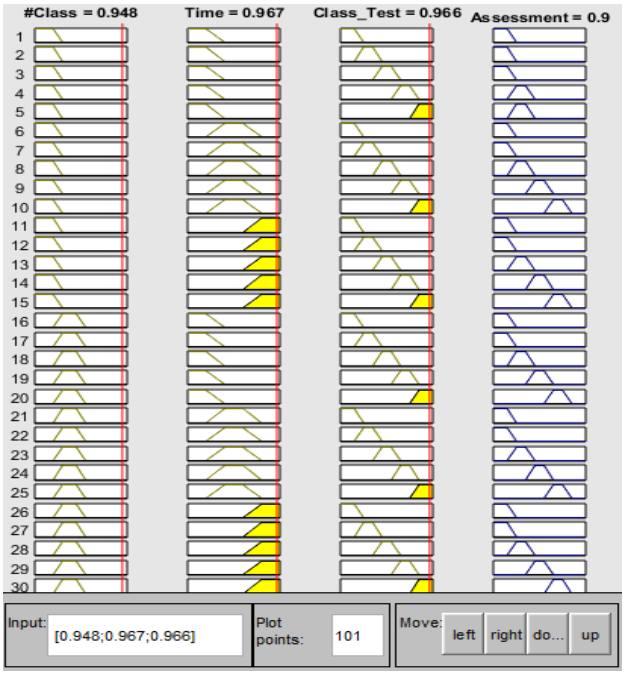


Fig. 7 Assessment Value for Student 1 Who was Presented 37 Class, spent 1890 Minutes in Class and Got 58 in Class Test.

C. *Comparison with Existing Method*

The comparing study between the existing system and the proposed fuzzy assessment system for the same 15 students out of 60 student's datasets for subject 1 is shown in TABLE VI. It shows, in the existing system the assessment score is always 0.00 for the students whose attendance is less than 60%.

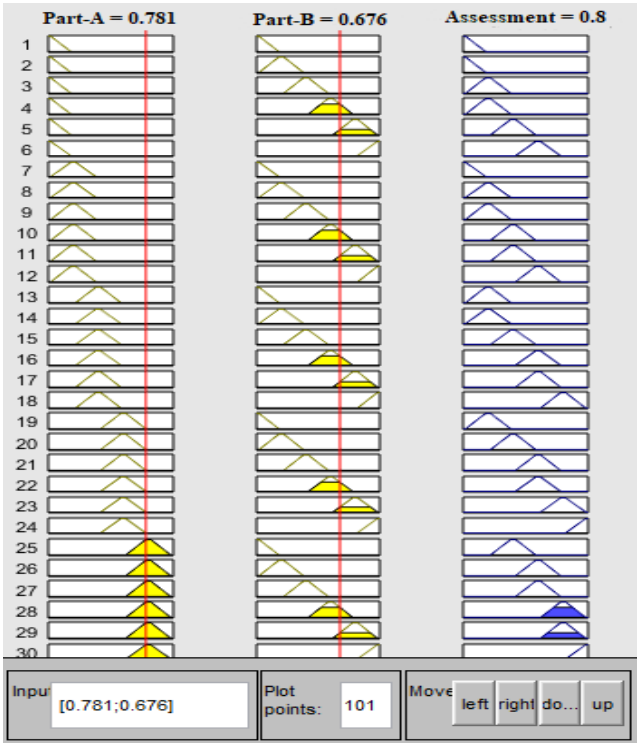


Fig. 8 Assessment Value for Exam papers Part-A and Part-B Based on Marks 82 and 71 Respectively.

TABLE VI: COMPARISON OF ASSESSMENT BETWEEN EXISTING METHOD AND PROPOSED METHOD

SI No.	Existing Method Continuous Assessment	Proposed Method Continuous Assessment	Existing Method Exam Paper Assessment	Proposed Method Exam Paper Assessment	Existing Method Grade	Proposed Method Grade
1	1.00	0.900	0.729	0.800	A+	A+
2	0.60	0.888	0.567	0.636	B+	A-
3	0.50	0.700	0.538	0.566	B-	B
4	0.40	0.326	0.529	0.533	C+	C
5	0.80	0.338	0.595	0.667	C+	B-
6	0.00	0.494	0.676	0.695	B-	B
7	0.60	0.444	0.548	0.557	C+	C+
8	0.90	0.900	0.595	0.600	B+	B+
9	0.00	0.500	0.567	0.600	C+	B-
10	0.00	0.300	0.500	0.547	C	C
11	0.00	0.176	0.662	0.695	C+	C+
12	1.00	0.813	0.676	0.695	A-	A-
13	0.00	0.300	0.529	0.600	C	C+
14	0.90	0.900	0.310	0.215	C	D
15	0.00	0.151	0.514	0.505	F	F

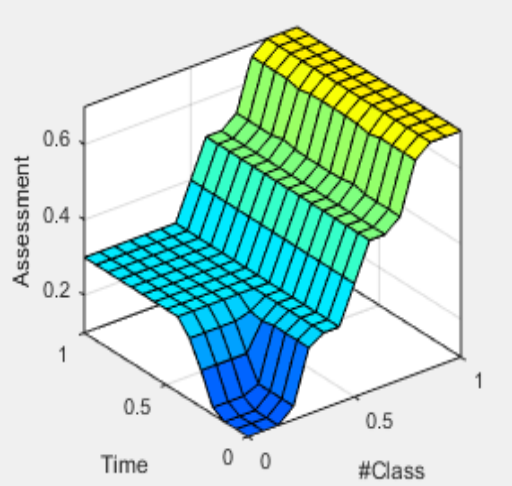


Fig. 9 Surface Viewer of Continuous Assessment.

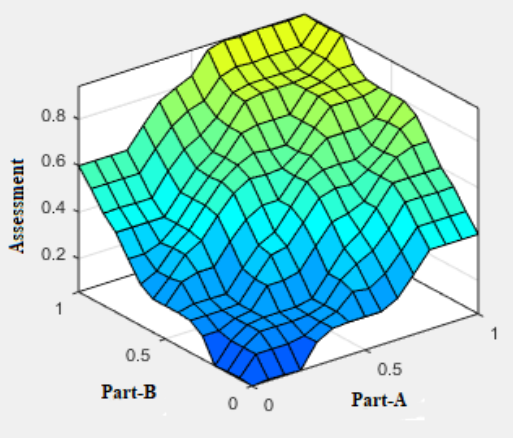


Fig. 10 Surface Viewer of Exam Papers Assessment.

TABLE VII: RESULT CHANGES BY EXISTING METHOD AND PROPOSED METHOD

Subject	# Total Students	Result Upgraded than Classical Method	Result Degraded than Classical Method	Total Changed
1	60	9	5	14
2	60	13	3	16
3	60	10	11	21
4	60	13	6	19
5	60	3	5	8

TABLE VII shows the changes in result for proposed method against existing method. Approximately, in average 16 students result is changed out of 60 in the proposed method but number of upgradation and degradation in results are varied subject to subject.

IV. CONCLUSION

A fair and unbiased performance evaluation system is needed for promoting education and to achieve learning outcome. Proposed method is very helpful in evaluating subjective types of answer paper. Our system considers many aspects of student's participation, not only grades got on exam marks in a formative way. Proposed system provides a smarter, fairer, transparent way to evaluate student performance. Proposed fuzzy logic centric evaluation method has greater flexibility with associated reliability and fairness while classical method uses a fixed mathematical rule. The method considers the aspect such as time spent on class making it more flexible and fairer than the conventional method which is clearly presented in the comparison TABLE VI. The system is implemented using Fuzzy Logic Toolbox by MATLAB.

REFERENCES

[1] Gokhan Gokmena, Tahir Çetin Akincib, Mehmet Tektaú, Nevzat Onatc, Gokhan Kocyigita, Necla Tektaú, "Evaluation of student performance in laboratory applications using fuzzy logic", Procedia Social and Behavioral Sciences 2 (2010) 902–909, WCES-2010

[2] L.A. Zadeh, "Fuzzy sets", Information and Control, (8), pp. 338-353, 1965.

[3] Echauz, J. R., & Vachtsevanos, G. J., "Fuzzy grading system. IEEE Transactionson Education", 38(2), 158–165, 1995.

[4] Chiu-Keung Law, "Using fuzzy numbers in educational grading system", Fuzzy Sets and Systems, 83(3), 311–323, 1996.

[5] S.M. Bai and S.M. Chen "A new method for students' learning achievement evaluation using fuzzy membership functions", Proceeding of the 11th Conference of Artificial Intelligence and Applications, Kaohsiung, Taiwan, Republic of China, pp. 177-184, 2006

[6] R. Biswas, "An application of fuzzy sets in students' evaluation," Fuzzy Sets System, vol. 74, 1995.

[7] S. M. Chen and C. H. Lee, "New methods for students' evaluating using fuzzy set," Proceedings of International Conference on Artificial Intelligence, 1996.

[8] Wilson, E., Karr, C. L., & Freeman, L. M., "Flexible, adaptive, automatic fuzzy-based grade assigning system", In Proceedings of the North American fuzzy information processing society conference (pp. 334-338) 1998.

[9] Ramjeet Singh Yadav and Vijendra Pratap Singh, "Modeling Academic Performance Evaluation using Fuzzy C-Means Clustering Techniques", International Journal of Computer Applications (0975 – 8887) Volume 60– No.8, December 2012.

[10] Arya A Surya, Merin k kurian and Surekha Mariam Varghese, "Overall Performance Evaluation of Engineering Students Using Fuzzy Logic", International Journal on Cybernetics & Informatics (IJCI) Vol. 5, No. 2, April 2016.

[11] Shilpa N. Ingoley and Dr. J.W. Bakal, "Students' Performance Evaluation Using Fuzzy Logic", Nirma University International Conference On Engineering, Nuicone-2012, 06-08 December, 2012.

[12] James R. Nolan, "An expert fuzzy classification system for supporting the grading of student writing samples", Expert Systems With Applications, Vol. 15, 1998.

[13] Ibrahim Saleh, Seong-in Kim "A fuzzy system forevaluating students' learning achievement" (Expert Systems with Applications 36 (2009) 6236–6243)

[14] Hui-Yu Wang and Shyi-Ming Chen, "Evaluating Students' Answer scripts Using Fuzzy Numbers Associated With Degrees of Confidence", IEEE Transactions on Fuzzy Systems, Vol. 16, No. 2, April 2008.

[15] Bardul and Mohamad, "A Fuzzy Approach on Students' Answer Scripts Criteria Based Evaluation"

Using Social Networks to Detect Malicious Bangla Text Content

Tanvirul Islam
Department of Computer Science and Engineearing
Daffodil International University
Dhaka,Bangladesh
tanvirul15-6117@diu.edu.bd

Nadim Ahmed
Department of Computer Science and Engineearing
Daffodil International University
Dhaka,Bangladesh
nadim15-5869@diu.edu.bd

Subhenur Latif
Department of Computer Science and Engineearing
Daffodil International University
Dhaka,Bangladesh
snlatif@daffodilvarsity.edu.bd

***Abstract* - Digital technology has accelerated social networking and brought revolutionary changes. Facebook, YouTube and Twitter are the most popular platforms for social communication, marketing and information distribution. Unhappily, these networks have been invaded by spammers who misuse these online social networking platforms with misinformation, fake news, rumors, malicious links, unsolicited messages, comments etc. Therefore, spam detection in social networks has become a novel framework for distribution of information, sentiments, and news. Our research reflects the current experimental study on spam detection from different textual data. Extensive researches have been done on spam detection field from English texts or other languages. But detection of spam from malicious Bangla text content still needs a lot of attention. In this experimental research, we have used Multinomial Naïve Bayes (MNB) classifier, a supervised machine learning algorithm with feature extraction to detect spam from Bangla text at the sentence level. Our proposed system detects spam on the basis of the polarity of each sentence associated with it. Finally, our experiment shows that the model has an accuracy of 82.44% in detecting spam Bangla text content.**

Keywords— Bangla Text., Multinomial Naïve Bayes classifier, Spam Detection, Social media.

I. INTRODUCTION

Social networking sites are basically represented by Facebook, Twitter, YouTube, and many others. At present, people spend most of their time on social media instead of search engines. Celebrities, public figures, business icons create their social pages for interacting with their users and fans. Unfortunately, malicious behaviors on social media bring troubles to the users. The initial purpose of social networking sites was to facilitate the connection and message sharing. Therefore, people are heavily dependent on online interaction for communication.

The increasing number of users in social networks, along with the trust they inherently have in their virtual profile, makes a propitious environment for the spammers. In fact, reports clearly indicate that the volume of spam over the social network is dramatically increasing year by year [1]. It represents a challenging problem for traditional filtering methods since such messages or links are usually fairly short and normally rife with slangs, idioms, symbols and acronyms that make tokenization a difficult task. Even no strong authentication mechanism is provided by the social networking sites to find out the spammers. There is a crucial need in the society and industry for saving the image and maintaining a healthy environment in social media. Improved accuracy and consistency in text-mining techniques can help us overcome the current problems.

However, spam analysis of Bangla texts is still a new area and there is a scope for improvement. There are 250-300 million Bangla speakers worldwide [16]. With the development of information technology, these native Bangla speakers are now able to use Bangla text in the social networking sites and everyday huge amount of Bangla texts are generated online. As a result, it will be easier to check the polarity i.e. how much positive or negative the sentence is. After analyzing the text pattern, the sentence can be categorized according to the polarity it belongs to.

Most of the researches on Bangla text were performed using news corpus and blogs that were extracted by scrapping the websites [5] [6]. Another source of data is social media where the opinionated texts are shorter in length but are informal, full of grammatical and spelling errors, in mixed languages and characters.

In our approach, we have mainly researched on how we can identify spam from every new generated Bangla sentences. Our textual data were collected from popular Facebook groups and pages e.g. ‘’Sakib Al Hasan”, “Hero Alom”, “Prothom Alo” etc. and YouTube videos of the viral YouTube channel “Hero Alom OFFICIAL”. The spammers were also attracted by these groups and video posts for their unethical interactions. They put their malicious links or spam texts which degrades normal user’s experience. So, we have collected open source data from these platforms and applied Multinomial Naïve Bayes classifier algorithm for spam detection from Bangla texts.

II. RELATED WORK

The increasing popularity and the vast amount of information residing on social network attracts researchers to study issue faced by the social network user community. A large proportion of work has been done on various problems e.g. spam detection and filtering, community detection, and information diffusion. Spam emails are classified [2] using four different classification techniques (Naive Bayes, SMO, J48, and random forest). The evaluation showed that random forest has the highest weighted precision, the highest weighted recall, the highest accuracy, and the highest weighted F-measure of 95.50%. A semi-supervised spam detection framework named S^3D was proposed in [3]. S^3D employs four lightweight detectors to detect spam tweets in

real time. In [4], authors proposed a spam detection approach for Twitter based on sentiments using five classifiers i.e. Naive Bayes, Random Forest, Support Vector Machine (SVM), J48 and Random Forest. Multinomial Naive Bayes has been used in [5] based on features selection method (Mutual Information) for sentiment analysis on both Bangla and English dataset. The experimental result shows that the accuracy of English dataset differs from the Bangla dataset of the proposed approach. A lexicon based backtracking approach was proposed to detect emotion from Bengali text [6]. Two basic emotion, 'happiness' and 'sadness' are considered to detect emotion from Bengali text and the experimental result shows that the proposed system can produce an accuracy of 77.16%. Houshmand Shirani-Mehr [7] applied several machine learning algorithm on an SMS spam dataset from UCI Machine learning repository where the evaluation showed that 10-fold cross-validation showed the best accuracy in their work. A new supervised machine learning approach has been introduced in [8] to detect spam accounts in social networks through their behavioral characteristics. They collected messages crawling from Sina Weibo and then selected three categories of features extracted from message contents, social interactions, and user profile properties and applied the ELM based spam accounts detection algorithm. To improve the performance of detecting spam SMS, a new content-based feature was implemented in [9]. The effectiveness of the proposed features is validated using several classification methods and the results demonstrate that the proposed system can improve the performance of SMS spam detection. Moreover, authors in [10] proposed a scalable and online social media spam detection system for social network security. For dealing with the scalability and real-time detection challenges, they implemented GAD clustering algorithm for large scale clustering and integrated it with the designed active learning algorithm. Another approach is presented in [11] who explored the four supervised learning Methods, namely Decision Tree, KNN, Naïve Bays, and Support Vector Machine (SVM) for categorization of Bangla web documents. An experiment has been adopted in [12] which emulates incremental training of personalized spam filters. Several classification algorithms were applied in [13] e.g. Logistic Regression, Boosted Tree and SVM to classify sentiment (Negative or Positive) of Bengali newspaper headline and Support Vector Machine (SVM) performs better having higher accuracy. In [14], a text processing approach for semantic analysis and context detection is proposed. They evaluated their approach with a public, real and non-encoded dataset along with several established machine learning methods which can enhance instant messaging and SMS spam filtering.

III. RESEARCH METHODOLOGY

It is mandatory to prepare a convenient dataset for training and testing before applying a classification algorithm. Pre-processing and features extraction of Bangla documents are also prerequisite for the construction of a model for efficacious classification. Figure 1 illustrates the overall system of the classification process.

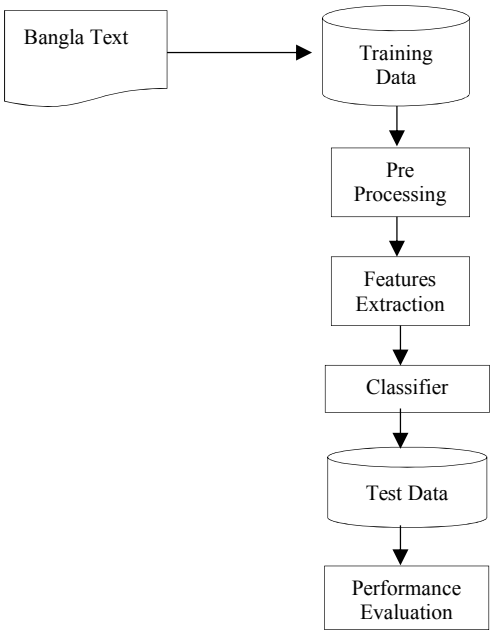


Figure 1. Proposed system of the classification process

A. *Proposed Dataset*

As no dataset is available for research purpose in terms of Bangla language, we have used some comment extractor tools e.g. "socialfy", "ytcomments" etc. to extract data from social platforms e.g. Facebook and YouTube. These social networking platforms have higher interactions of the users and vast amount of contents are posted every single day. We have accessed only public data from random users. Our proposed dataset consists of a wide variety of 1965 instances including 646 spam and 1319 ham. The percentage in dataset is given in the following table.

TABLE I. THE DISTRIBUTION OF HAM AND SPAM IN DATASET

Instances	**Percentage in Dataset**
Spam(1)	32.87
Ham(0)	67.13

B. *Pre-Processing*

A proper representation of words within the text document is important to achieve satisfactory classification performance. To train our model, it requires tagged data. We have formatted our dataset into two columns i.e. One is "Text" that contains actual text data and the other is "Status" that contains value 0 or 1. Spam text was labeled as "1", and non-spam was labeled as "0". It's recommended to apply a classification algorithm on cleaned corpus instead of noisy corpus. The noisy corpus includes insignificant things within the text e.g. misspelling, numerical values, emoticon, stop words e.g. ',', '!', '?', '.', '.', '~', '||', 'ı' etc. They degrade the processing which provides a low accuracy rate. Elimination of these entities from corpus will enhance the accuracy. For example – emoticon like ":-P", ":-D" etc. are meaningful while sentiment analysis but may not be meaningful while

other classification. After elimination of all the punctuation marks, numerical values, and emoticons, we have a clean dataset to fit into a classification algorithm.

C. Feature Extraction

After the pre-processing phase, we have to extract features from these words prior to applying the algorithm. Different statistical approaches can be used to extract features from this text corpus e.g. Count Vectorizer, TF-IDF (term frequency–inverse document frequency) vectorizer etc. Count vectorizer only counts the frequency of each word. We used TF-IDF vectorizer to extract the features from the document because it provides a way to score the importance of word based on how frequently they appear across multiple documents.

- If a word appears frequently within a document, give that word higher score.
- If a word appears frequently across multiple documents, that means it's not unique identifiers, give that lower score.

That is how, a common word e.g. "আমি", "তুমি", "ও", "এবং" that frequently appears across many documents, will be scaled down and word that appears frequently within a single document, will be scaled up. This will lead towards a better classification performance. The weight for a term 'i' in terms of TF-IDF is given by

$$W_i = \frac{(\mathrm{TF_i} \times \log(\mathrm{N}/\mathrm{n_i}))}{\sqrt{\sum_{i=1}^{n} (\mathrm{TF_i} \times \log(\mathrm{N}/\mathrm{n_i}))^2}}$$

Where N= total number of documents and n_i= document frequency of term 'i'.

D. Training

In this section, the processed data are split into a testing-set and a training-set. The 80/20 split is chosen. This research implements the algorithm called Naive Bayes Classifier which is widely used for spam detection. When dealing with text, it's very common to treat each unique word as a feature and since the typical person's vocabulary consist of numerous thousands of words, it produces a huge number of features. The simplicity of the algorithm and the capability of handling a large number of features makes Naïve Bayes Classifier a strong performer for classifying texts. It becomes a huge challenge regarding which variant of Naïve Bayes should be applied. It has been shown experimentally in [15] that Multinomial Naïve Bayes performs better than the Bernoulli Naïve Bayes and Gaussian naive Bayes in case of text classification. Multinomial naive Bayes treats features as event probabilities. Multinomial Naïve Bayes classifier are applied on the proposed dataset.

IV. EXPERIMENTAL RESULTS

After training, using training dataset, it's time to test our dataset using test set that is unknown to our model. From results, Multinomial Naive Bayes yields an overall accuracy of 82.44%. The confusion matrix is shown as follows in the following table.

TABLE II. CONFUSION MATRIX

	Predicted Ham	Predicted Spam
Actual Ham	269	11
Actual Spam	58	55

Test results after performing the tests have been pictured in the following figures.

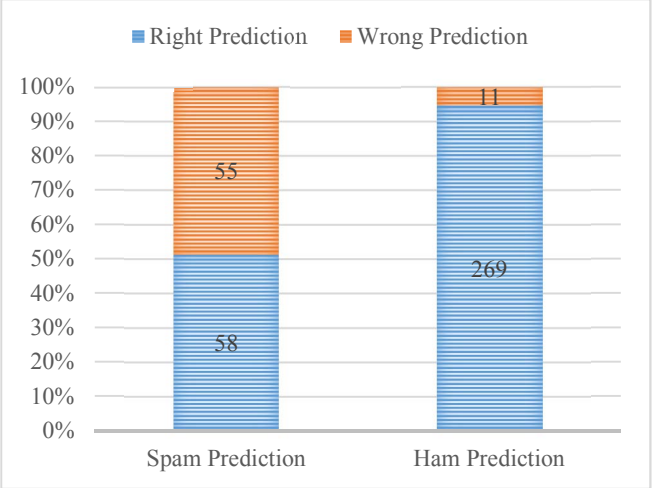


Figure 1. Confusion Matrix in Bar Chart

This chart shows the performance of the classifier in case of both spam and ham prediction. The accuracy of the ham prediction is higher than the spam prediction as the corpus has more ham compared to spam. So, the overall accuracy is satisfactory.

Error, Precision, Recall, F1 Score, AUC (Area Under Curve) are shown in the following table.

TABLE III. PERFORMANCE MEASURE

Precision	0.825
Recall	0.824
F-Score	0.808
Error	17.56
AUC (Area Under Curve)	0.72

In the testing phase, precision: 0.825 determines that 82.5% of positive prediction is actually correct. The recall: 0.824 indicates that 82.4% of actual positive was predicted correctly. F-score; the weighted average of precision and recall is 0.808.

The ROC is shown in the following figure.

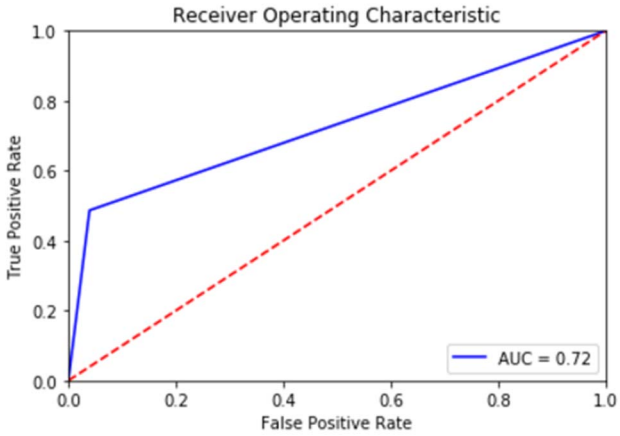


Figure 4. The ROC (Receiver Operating Characteristics) Curve.

V. CONCLUSION AND FUTURE WORK

During the last few decades, text classification has received an incredible attention from people because it helps to classify spam data and threats. Therefore, a lot of work has been done in this domain to find the finest classifier for text classification. Our text classification method helps us to detect exact expression that majority people think about. Among different approaches, we used Multinomial Naïve Bayes classification algorithm to extract semantic information from textual data for detecting spam. After extensive pre-processing, Multinomial Naïve Bayes was applied and it turned out to be 82.44 % successful in classifying malicious Bangla text content at the sentence level.

Besides, we also acknowledge that our corpus doesn't have sufficient lexicons and the mechanisms of feature extraction may be low in future. As every day, new data are generated through social media, a collection of sentences having new patterns is necessary to enrich the corpus. Further research on using combined-classifiers may increase the accuracy rate in future. Moreover, a stemmer may be added to reduce the size of the corpus and improve model performance.

REFERENCES

[1] *Go.nexgate.com*, 2019. [Online]. Available: https://go.nexgate.com/nexgate-social-media-spam-research-report. [Accessed: 26- Mar- 2019].

[2] S. Saha, S. DasGupta and S. Das, "Spam Mail Detection Using Data Mining: A Comparative Analysis", *Smart Intelligent Computing and Applications*, pp. 571-580, 2018. Available: 10.1007/978-981-13-1921-1_56 [Accessed 11 April 2019]

[3] S. Sedhai and A. Sun, "Semi-Supervised Spam Detection in Twitter Stream," in *IEEE Transactions on Computational Social Systems*, vol. 5, no. 1, pp. 169-175, March 2018. doi: 10.1109/TCSS.2017.2773581

[4] Nasira Perveen, Malik M. Saad Missen, Qaisar Rasool, Nadeem Akhtar, "Sentiment Based Twitter Spam Detection" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 7, 2016.

[5] A. Paul and P. Shill, "Sentiment mining from Bangla data using mutual information", *2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)*, 2016. Available: 10.1109/icecte.2016.7879569 [Accessed 11 April 2019]s

[6] Tapasy Rabeya, Sanjida Ferdous, Himel Suhita Ali, Narayan Ranjan Chakraborty, "A Survey on Emotion Detection A lexicon based backtracking approach for detecting emotion from Bengali text", *20th International Conference of Computer and Information Technology (ICCIT)*, 2017.

[7] Houshmand Shirani-Mehr, "SMS spam detection using machine learning approach", pp. 1-4, 2013.

[8] Liu, C., Wang, G.: Analysis and detection of spam accounts in social networks. In: 2016 2nd IEEE International Conference on Computer and Communications (ICCC). IEEE (2016)

[9] A. Karami and L. Zhou, "Improving static SMS spam detection by using new content-based features," vol. 51, no. 1, pp. 1–9, 2014

[10] X. Jin, C.X. Lin, J. Luo, and J. Han. Socialspamguard: A data mining-based spam detection system for social media networks. PVLDB, pages 1458–1461, 2011.

[11] A. Mandal and R. Sen, "Supervised Learning Methods for Bangla Web Document Categorization", *International Journal of Artificial Intelligence & Applications*, vol. 5, no. 5, pp. 93-105, 2014. Available: 10.5121/ijaia.2014.5508 [Accessed 12 April 2019].

[12] Jin, Z. (2008). Spam message self-adaptive filtering system based on Naive Bayes and support vector Machine. Journal of Computer Applications, 28(3), 714-718. doi:10.3724/sp.j.1087.2008.00714

[13] M. Hossain, I. Jui, and A. Suzana, "Sentiment analysis for Bengali newspaper headlines", Hdl.handle.net, 2019. [Online]. Available: http://hdl.handle.net/10361/8246. [Accessed: 01- Mar- 2019].

[14] T. A. Almeida, T. P. Silva, I. Santos, and J. M. G. Hidalgo, "Text normalization and semantic indexing to enhance Instant Messaging and SMS spam filtering," Knowledge-Based Systems, vol. 108, pp. 25–32, 2016.

[15] A. McCallum and K. Nigam, "A Comparison of Event Models for Naive Bayes Text Classification," in *Learning for Text Categorization: Papers from the 1998 AAAI Workshop*, 1998, pp. 41--48.

[16] "Bengali language", *En.wikipedia.org*, 2019. [Online]. Available: https://en.wikipedia.org/wiki/Bengali_language. [Accessed: 14- Apr- 2019].

Classification of Bengali Questions Towards a Factoid Question Answering System

Syeda Tamanna Alam Monisha
Computer Science and Engineering
Shahjalal University of
Science and Technology
Sylhet-3114, Bangladesh
alammonisha@gmail.com

Sourav Sarker
Computer Science and Engineering
Shahjalal University of
Science and Technology
Sylhet-3114, Bangladesh
sourav39@student.sust.edu

Md Mahadi Hasan Nahid
Computer Science and Engineering
Shahjalal University of
Science and Technology
Sylhet-3114, Bangladesh
nahid-cse@sust.edu

***Abstract*— This paper demonstrates question classification for Bengali language questions which is the first step towards developing a closed domain based Bengali factoid question answering system. Question classification is an influential step towards a question answering system on which the performance of the whole system depends. We have used machine learning algorithms Stochastic Gradient Descent(SGD), Decision Tree(DT), Support Vector Machine(SVM) and Naive Bayes(NB) for the question classification process where Support Vector Machine(SVM) with linear kernel performs the best providing the accuracy 90.6% over five coarse-grained categories.**

***Index Terms*— Question Classification, Bengali Question Classification, Question Answering System, Factoid Question Answering System, Answer Type.**

I. INTRODUCTION

Bengali is the eighth most spoken language in the world. In Bangladesh Bengali is spoken as the primary language and is the second most commonly spoken language in India. About 242 million native and 261 million people in total all over the world speak in Bengali. But still compared to English the work in computational linguistic with Bengali language is yet very low. There is no automated question answering system in Bengali which will reply to our queries by processing the Bengali Language. Towards developing a question answering system question classification is an important first step [1].

Question classification is the process of classifying questions to predefined classes where the classes reflect the answer type of the respective questions. For example for the question "Where was Rabindranath Tagore born?", the task of question classification is to add the label "Location" to this question as the expected answer is of "Location" type.

A question answering system provides the most pertinent answer to a question asked in natural language which is to be found from a large set of documents. As the number of documents from which the answer is to be extracted is huge, it is difficult to find the most relevant one. But question classification makes the task easier as knowing the expected answer type of a question reduces the search space by a considerable amount [2]. Therefore a well performing question classifier algorithm makes the path easy to an efficient question answering system.

II. RELATED WORKS

A lot of work have been done and are also being done on question classification for different language questions. Besides there are also works on question features and question type taxonomy which is the set of classes to which the questions are to be classified. In QA system there are basically two different ways for the question classification process which are rule-based approach and machine learning approach. There are also researches where hybrid approaches have been used which integrates both the rule-based approach and the machine learning based approach [3][4].

The early approaches to question classification were mostly rule-based where many handcrafted rules were used for classifying the questions to their expected answer type. MULDER, a QA system determines the type of the question just by looking into the question's interrogative pronoun [5]. Author's used the object of the verb in the question to determine the type. 276 handcrafted rules were written by Authors Hermjakob et al., 2001 for question classification where there were 122 classes for the questions to be classified in [6].

In the machine learning approach features are extracted from the questions to classify them. Author's Zhang et. al., 2003 experimented question classification with five machine learning algorithms: Nearest Neighbors (NN), Naïve Bayes (NB), Decision Tree (DT), Sparse Network of Winnows (SNoW), and Support Vector Machines (SVM) using two kinds of features: bag-of-words and bag-of-ngrams where SVM outperforms the other four methods giving the highest accuracy 87.4% for both the features [7]. Another machine learning based approach was made by author's Hacioglu et. al., 2003 using Support Vector Machines and Error Correcting Codes which provides with highest accuracy 82% [8]. A question classification system based on Sparse

Network of Winnows (SNoW) architecture was developed by author's Li et. al., 2002 [9]. A collection of semantic and syntactic features were used for the system and it shows an accuracy of 91.0% on general coarse-grained question types and 84.20% on more specific fine-grained types.

Although the task of question classification is difficult in Bengali, there have been some works done. Author's Banerjee et. al., 2012 have studied suitable lexical, syntactic and semantic features to classify Bengali questions and used machine learning classifiers Naïve Bayes (NB), Kernel Naïve Bayes (KNB), Rule Induction and Decision Tree where the highest accuracy 87.63% was provided by Decision Tree [1]. Bengali question classification using Support Vector Machine(SVM) with hybrid feature extraction method was done by a team which provides an accuracy of 89.14% [10]. Word/Phrase based answer type classification for Bengali question answering system has been done by Islam et. al., 2016 [11]. Authors used stochastic gradient descent (SGD) classifier method and bi-gram and stop words for feature extraction. The system based on stochastic gradient descent (SGD) classifier provides average precision 0.95562 for coarse and 0.87646 for finer classes.

III. CORPUS CONSTRUCTION & ANALYSIS

Dataset is an important issue towards developing a question answering system. As we are working on a closed domain based Bengali factoid question answering system on the domain Shahjalal University of Science & Technology (SUST) we could not get any prepared dataset for our work. So we had to manually prepare our questions. This data collection task was one of the challenging tasks as we had no resource available. Table I shows some sample question from our corpus.

TABLE I
Sample factoid questions on SUST

Sample Question
সাস্টে সব মিলিয়ে সিট সংখ্যা কত?
শাহজালাল বিজ্ঞান ও প্রযুক্তি বিশ্ববিদ্যালয়ে অনুষদ কয়টি?
সাস্টের বর্তমান উপাচার্যের নাম কি?
শাবিপ্রবিতে একাডেমিক কার্যক্রম বছরে কয়টি সেমিস্টারে হয়?
সাস্টে কয়টি বিভাগ রয়েছে?

We used different sources for our data collection part. We used various sources for building question dataset such as crowd sourcing, social media, manual generation etc. But among them, we collected most of the data from crowd sourcing. We collected questions from the students of Shahjalal University of Science & Technology as well as from the official website of SUST where we got the frequently asked questions by the candidates. We also prepared questions from different documents, articles and news on SUST. Table III shows the statistics of our question database.

TABLE II
Sources of Question Database

Type	Source	Amount of Data
Crowd Sourcing	2nd Year students, Dept of CSE,SUST	11300
Social Media Data Crawling	Facebook Group: SUST Admission Aid, SUST Admission Help etc.	1055
Manual Generation	Authors	3000

TABLE III
Statistics of Question Database

Attribute	Description
Question type	Factoid
Total number of questions	15355
Avg. words carried by a question	6
Avg. characters carried by a word	5

IV. METHODOLOGY

A. Answer Type Taxonomy

Because we are working on the domain Shahjalal University of Science & Technology we defined five coarse-grained classes related to SUST for the question classification process. The questions are classified in these five categories. Table IV shows the category details and table V shows some sample tagged question in our corpus.

TABLE IV
Question Categories

Class Name	Short form	Description
Administration	ADS	Questions that require administrative information as answer are of ADS type
Admission	ADM	Questions that require admission related information as answer are of ADM type
Academic	ACD	Questions that require academic information as answer are of ACD type
Campus	CAM	Questions that require campus related information are of CAM type
Miscellaneous	MISC	Questions that require information other than the above four types as answer are of MISC type

TABLE V
Sample Tagged Question

Question	Label
সাস্টে মোট কতজনকে ভর্তি পরীক্ষার সুযোগ দেওয়া হয়	ADM
সিএসই বিভাগ হতে স্কলারশিপ পাওয়া যায় কি	ACD
সাস্ট মানে কি	MISC
প্রতিবর্ষে কি আলাদা পরীক্ষা ফি দিতে হয়	ADS
সাস্টের আয়তন কত	CAM

Our question dataset contains 15355 Bengali questions in total related to SUST. Among them most of them are of administration type which is they require administrative information. Figure 1 reflects the percentage of questions in each of the five categories that we defined for our system.

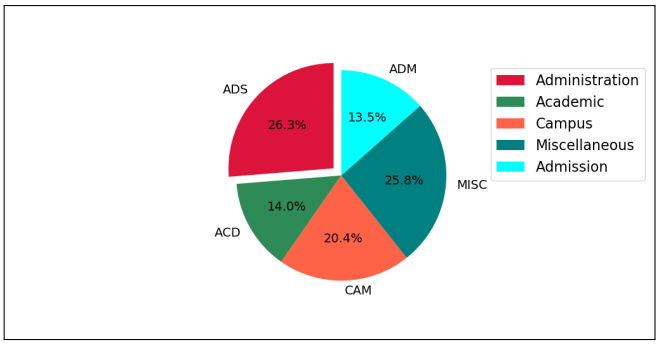


Fig. 1. Percentage of Questions in each category

B. Feature Extraction

The selection of features is an important part for the question classification process and all classification tasks.There are many features that are used for detecting the answer type of a question. Some of the features that are used frequently are lexical features, named entities, parts-of-speech tags, wh-words, semantic features etc. We used the following features for our question classification process:

1) Word level n-gram: A question Q can be written as the following expression.

$$Q_k = W_1, W_1, W_3, ..., W_n$$

where Q_k is any question and $W_1, W_1, W_3, ..., W_n$ represents consecutive words of that question. This successive words or n-gram is important for question classification process as questions of same class is likely to share word n-grams.

For our classification process for feature extraction we experimented with both bi-gram and tri-gram method where tri-gram provides better results than bi-gram.

2) Similar words: Another feature that we have used are words that carry the same meaning. We have mapped a word to words that carry the same meaning. This mapping makes it easy to determine the questions of the same meaning.

TABLE VI
Sample Similar Words

Similar Words
লাইব্রেরী, গ্রন্থাগার, পাঠাগার
ভিসি , উপাচার্য , ভাইস-চান্সেলর
⋮
শাবি, শাবিপ্রবি, শাহজালাল বিশ্ববিদ্যালয়, শাহজালাল বিজ্ঞান ও প্রযুক্তি বিশ্ববিদ্যালয় , সাস্ট

C. Classification Algorithms

We have done the question classification using machine learning based approaches. We used four machine learning algorithms Stochastic Gradient Descent(SGD), Decision Tree(DT), Support Vector Machine(SVM) and Naive Bayes(NB) where Support Vector Machine(SVM) outperforms the other three classifiers. The Stochastic Gradient Descent(SGD) classifier also shows better performance over our dataset.

D. Performance Metric

In general to evaluate the performance of a question classifier the accuracy measure is used. Accuracy is the number of correct predictions made by a model over all kinds of predictions made. Some other measures are precision, recall, f1 score which are also used for performance evaluation. We used the accuracy measure to evaluate our experiments. The accuracy is measured by the formula

$$Accuracy = \frac{TP + TN}{P + N}$$

where,

P = total number of positive samples

N = total number of negative samples

TP (True Positive) = number of cases when the actual class of the sample is true and the predicted is also true

TN (True Negative) = number of cases when the actual class of the data point is false and the predicted is also false

V. EXPERIMENTS & RESULT ANALYSIS

For the training purpose of the question classification stage we had to manually classify the questions of our dataset to predefined classes. We have used 75% of the question dataset for the training purpose and 25% for testing.

In our experiment we have used both bi-gram and tri-gram along with the feature similarity or similar words for feature extraction. We used the classifier algorithms Stochastic Gradient Descent(SGD), Decision Tree(DT), Support Vector Machine(SVM) and Naive Bayes(NB) for the question classification process where in both the cases Support Vector Machine(SVM) with linear kernel performs the best. As Support Vector Machine methodology works by pairing with kernel trick, so using fine tune of the kernel we probably have a better result. Again, we have used only five classes and that is one of the key reasons of showing better accuracy according to this work [7]. Though our classes are different but we have worked on a single domain which also holds a significance in accuracy.

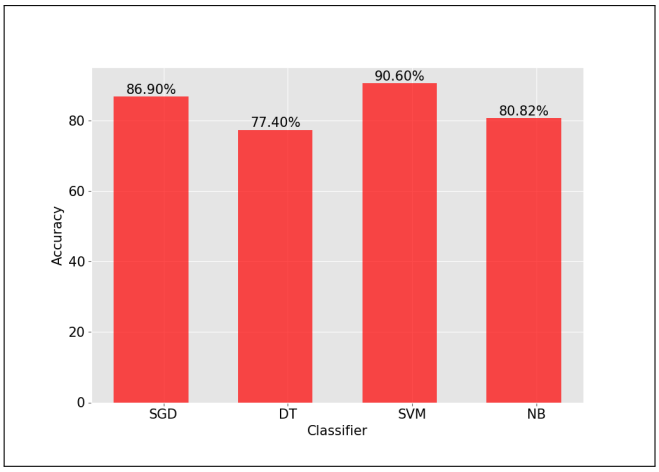


Fig. 2. Best performance for different classifier

We have experimented Support Vector Machine(SVM) with linear, rbf, sigmoid and polynomial kernels where linear kernel provides the best performance.

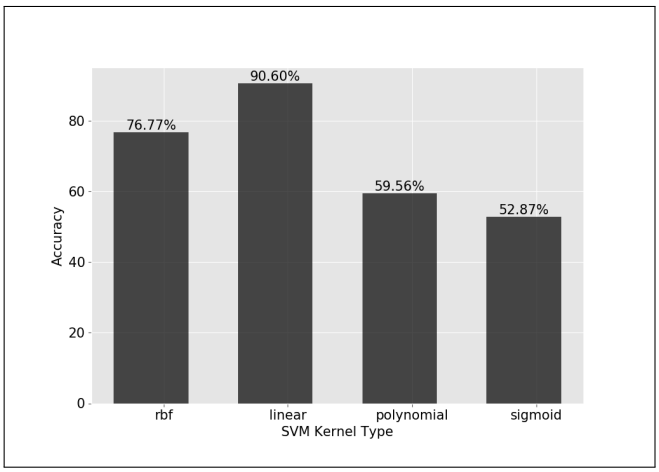


Fig. 3. Best performance for different SVM kernels

We have experimented with both bi-gram and tri-gram for feature extraction and for all the classifiers that we used for question classification process tri-gram gives better accuracy than bi-gram. The following graph shows the comparison between bi-gram and tri-gram for the classifiers.

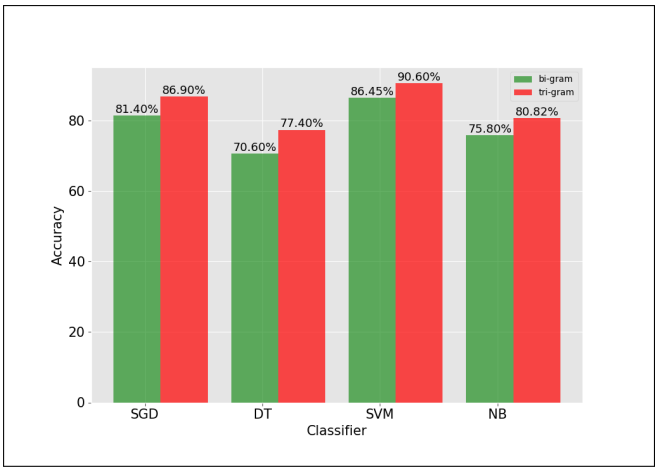


Fig. 4. Performance comparison for different classifier using bi-gram and tri-gram

VI. CONCLUSION

In this paper we have presented an approach for question classification towards building a factoid question answering system for Bengali language on a closed domain. We have built a corpus consisting of 15355 questions on the domain Shahjalal University of Science & Technology and used machine learning classifiers for the classification process over our dataset. As for Bengali language no standard dataset is yet available this can be helpful for the researchers in this field. Natural language processing(NLP) tools like named entity recognizer(NER) can help to enhance the performance of the classification process as the question classification phase plays a vital role in the performance of answer extraction in question answering system.

So far we have done the work for factoid questions only. We have future plans to work for complex and descriptive questions also and also want to implement more techniques to enhance the performance of the classification process.

References

[1] S. Banerjee and S. Bandyopadhyay, "Bengali question classification: Towards developing qa system," in *Proceedings of the 3rd Workshop on South and Southeast Asian Natural Language Processing*, 2012, pp. 25–40.

[2] E. Haihong, Y. Hu, M. Song, Z. Ou, and X. Wang, "Research and implementation of question classification model in q&a system," in *International Conference on Algorithms and Architectures for Parallel Processing*. Springer, 2017, pp. 372–384.

[3] S. K. Ray, S. Singh, and B. P. Joshi, "A semantic approach for question classification using wordnet and wikipedia," *Pattern Recognition Letters*, vol. 31, no. 13, pp. 1935–1943, 2010.

[4] Z. Huang, M. Thint, and Z. Qin, "Question classification using head words and their hypernyms," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2008, pp. 927–936.

[5] C. Kwok, O. Etzioni, and D. S. Weld, "Scaling question answering to the web," *ACM Transactions on Information Systems (TOIS)*, vol. 19, no. 3, pp. 242–262, 2001.

[6] U. Hermjakob, "Parsing and question classification for question answering," in *Proceedings of the ACL 2001 workshop on open-domain question answering*, 2001.

[7] D. Zhang and W. S. Lee, "Question classification using support vector machines," in *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*. ACM, 2003, pp. 26–32.

[8] K. Hacioglu and W. Ward, "Question classification with support vector machines and error correcting codes," in *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology: companion volume of the Proceedings of HLT-NAACL 2003–short papers-Volume 2*. Association for Computational Linguistics, 2003, pp. 28–30.

[9] X. Li and D. Roth, "Learning question classifiers," in *Proceedings of the 19th international conference on Computational linguistics-Volume 1*. Association for Computational Linguistics, 2002, pp. 1–7.

[10] S. M. H. Nirob, M. K. Nayeem, and M. S. Islam, "Question classification using support vector machine with hybrid feature extraction method," in *2017 20th International Conference of Computer and Information Technology (ICCIT)*. IEEE, 2017, pp. 1–6.

[11] M. A. Islam, M. F. Kabir, K. Abdullah-Al-Mamun, and M. N. Huda, "Word/phrase based answer type classification for bengali question answering system," in *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)*. IEEE, 2016, pp. 445–448.

Evidential SVM for binary classification

Md. Eusha Kadir
and Pritom Saha Akash
Institute of Information Technology
University of Dhaka
{bsse0708, bsse0604}@iit.du.ac.bd

Amin Ahsan Ali
Department of
Computer Science & Engineering
Independent University, Bangladesh
aminali@iub.edu.bd

Mohammad Shoyaib
and Zerina Begum
Institute of Information Technology
University of Dhaka
{shoyaib, zerin}@du.ac.bd

***Abstract*—Support Vector Machine (SVM) is one of the most popular supervised learning methods for its better performances over diversified applications. SVM constructs a maximum-margin hyperplane and predicts the class of a new incoming data point based on that hyperplane. However, the hyperplane may not give a reliable decision in some cases specially when the training data is imprecise and noisy. To improve this situation, we propose an evidence based binary SVM classifier (EbSVM) where we first identify the sources of information (SoI) and propose a method to generate mass value from these SoIs. Finally, we combine these mass values using DS theory of evidence. An experiment over six benchmark datasets illustrates that EbSVM significantly performs better than the existing state-of-the-art methods.**

***Index Terms*—SVM, Belief, Evidential, Mass value**

I. Introduction

Machine learning is an area of artificial intelligence that mainly focuses on data analysis and interpretation of data to enable learning, reasoning and decision making. In machine learning, supervised learning approaches are commonly used for classification and categorization of data. A significant amount of research on supervised learning has been conducted to obtain an accurate classification. Many supervised learning algorithms such as K-Nearest Neighbors (KNN) [1], Decision Tree (DT) [2], Support Vector Machine (SVM) [3] and Naive Bayes (NB) [4] have been developed in this regard.

One of the simplest but most successful method is NB classifier. But for small amount data, it often assumes equi-probable priors and may fail in many cases. On the other hand, classifiers like KNN, DT and SVM do not assume any priors. Performances of all of these classifiers degrade when samples are located in highly overlapping area or the samples are noisy. To deal with these situations, various theories, such as fuzzy set theory [5], [6], imprecise probability theory [7], Dempster-Shafer (DS) theory of evidence [8], possibility theory [9], are used to build learning algorithms. Among them, DS theory of evidence is very popular. Different aspects of uncertainty are clearly distinguished by DS theory of evidence. As it is difficult to classify uncertain and imprecise data correctly to specific classes, DS theory of evidence allows one to assign mass values for meta classes for datapoints. Decisions are made using these mass values. DS theory of evidence is briefly described in section II-B.

SVM is considered to be one of the most successful and theoretically sound supervised learning methods. A brief descriptions on SVM is given in section II-A. A binary SVM classifies new datapoints based on which side of a hyperplane the datapoint is located. SVM finds the hyperplane that reduces the classification error, while having a maximum margin between the classes. However, SVM does not assign class-wise probability of that datapoint. Providing probabilistic outputs is necessary when they are used as a part of an overall decision making process. In [10], Platt presents a method (PSVM) which estimates probabilistic outputs for SVM. To estimate the probabilistic outputs, all the training points are projected on the kernel space and the perpendicular distance for each training point from the hyperplane is measured. Then a logistic regression is performed on these distances to extract probabilities. The class probability of a new input pattern x_i is thus calculated using (1).

$$\Pr(y = 1|x_i) = \frac{1}{1 + \exp(Af(x_i) + B)} \tag{1}$$

where, A and B are estimated by minimizing the negative log-likelihood function of the training data, $f(x)$ is the perpendicular distance of x from the maximum margin hyperplane, and $y \in \{1, -1\}$. Based on the distance from the hyperplane, a single probability is obtained in this case. As an extension to this classical probabilistic calibration method, an evidential calibration of binary SVM (EcSVM) is proposed [11]. In EcSVM, instead of estimating class membership probabilities, a mass function is estimated. This method better represents the uncertainty using the calibration procedure when less training data are available.

However, the aforementioned calibration methods do not take into account the density of the datapoints on either side of the hyperplane and the closeness of these datapoints to the test datapoint. For example, in Fig. 1, there are two test points named x_p and x_q, and situated in the same side of the separating hyperplane. So, SVM classifies both points as Class-1. Again as the distances to the hyperplane for both points are same, PSVM will assign the same probability to these two points. However, from Fig. 1 we observe that, x_p is more closer to the data points of Class-2 than the points of Class-1, whereas x_q is more closer to the data points of Class-1. Therefore, intuitively x_q should obtain higher probability (or mass value in the evidential case) for belonging to Class-1 than x_p. We need to consider this case and find a way so that

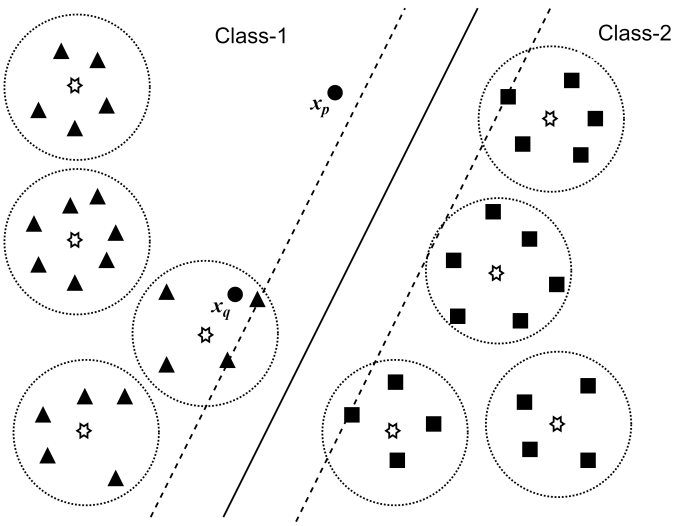


Fig. 1. Overview of plotting data points of binary classes where ▲ represents points belonging to Class-1, ■ represents points belonging to Class-2, • represents the test points, dotted lines represent the marginal hyperplane, the solid line represents the separating hyperplane of SVM.

x_q will be more probable for Class-1 than x_p which satisfies the intuition.

In recent years, many approaches using DS theory of evidence have been developed to estimate mass values [12]–[15]. In this paper, we propose a mechanism of generating mass values based on the different source of information (SoIs) provided by the SVM which also satisfies our above mentioned intuitions. These multiple evidences from these SoIs may provide more reliable estimates of the mass value. Our main contributions in this paper are as follows:

- Firstly, we propose a method to find multiple SoIs from the information provided by SVM.
- Secondly, we propose the mechanism of combining the evidences provided by SoIs to generate the mass value.
- Thirdly, we propose evidential SVM for binary classification and examine its performance on benchmark datasets.

The rest of the paper is organized as follows. In section II, we give an overview of the methods that are required to understand our proposed method. Our proposed method is discussed in section III. In section IV, we present experiments that we performed to validate the performance of our proposed method. Finally, we conclude and discuss our future work in section V.

II. Preliminaries

To understand our approach, the knowledge about SVM, DS theory of evidence, and K-means clustering is important and these methods are briefly discussed here.

A. Support Vector Machine

Let, consider a binary classification problem with features $X = \{x_1, x_2, ..., x_n\}$ and labels $\theta = \{\theta_1, \theta_2, ..., \theta_n\}$. A training set of n data points can be derived as $\{(x_i, \theta_i)\}$ where $x_i \in \mathbb{R}^l$ (l-dimensional feature vectors) and $\theta_i \in \{1, -1\}$ (for $i = 1, 2, ..., n$). SVM finds the maximum-margin *hyperplane* that separates the group of points into two classes [3]. The hyperplane can be described using (2).

$$\sum_{i=1}^{n} \alpha_i \theta_i \kappa(x, x_i) + b = 0 \tag{2}$$

where, α_i ($0 \leq \alpha_i \leq C$ for $i = 1, 2, ..., n$ and C represents the trade-off between margin and misclassification error) are positive constants, $\kappa(x, x_i)$ is a symmetric positive definite kernel function and b is bias.

For linear SVM, $\kappa(x, x_i) = x_i^T x$ is used. So, the weight vector, w, can be expressed as $w = \sum_{i=1}^{n} \alpha_i \theta_i x_i$ and the hyperplane for (2) can be represented by (3).

$$w^T x + b = 0 \tag{3}$$

A decision function, $f(x) = \text{sgn}(w^T x + b)$, is a scalar function of an input feature vector, x. A new incoming feature vector will be classified according to the sign of the decision function that can be described using (4).

$$f(x) = \begin{cases} -1 & \text{if } w^T x + b < 0 \\ 1 & \text{otherwise.} \end{cases} \tag{4}$$

B. DS Theory of evidence

Let, Θ is a finite set, which is called the *frame of discernment* [8]. It is assumed that, Θ is a set of mutually exclusive and collectively exhaustive hypotheses. A number in the range [0, 1] is used to indicate the belief in a hypothesis based on a piece of evidence. This belief is the degree to which the evidence supports the hypothesis. The power set of Θ is 2^Θ. A basic belief assignment (BBA) is a function $m : 2^\Theta \rightarrow [0, 1]$ which satisfies the following properties:

$$m(\emptyset) = 0 \tag{5}$$

$$\sum_{A \subseteq \Theta} m(A) = 1 \tag{6}$$

where, $m(A)$ is the quantity, which can be called mass value, represents the belief reflecting how strongly A is supported. The mass value assigned to Θ is called the degree of ignorance and the subsets A of Θ with non-zero mass values are called the focal elements.

Several pieces of evidence can be characterized by different BBAs. These different BBAs can be combined using Dempster's rule of combination. The Dempster's rule of combination from two BBAs $m_1(.)$ and $m_2(.)$ can be expressed using (7).

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - K} \quad A \neq \emptyset \tag{7}$$

where, $A, B, C \in 2^{\Theta}$ and $K = \sum_{B \cap C = \emptyset} m_1(B) m_2(C)$ which is the degree of conflict.

The lower and upper probabilities of A usually interpret as belief $Bel(A)$ and the plausibility $Pl(A)$ [8]. These are two common evidential measures which are mathematically defined in (8) and (9).

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{8}$$

$$Pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \tag{9}$$

C. K-means clustering

K-means clustering is a process that puts n training data points in an l-dimensional space into K clusters [16], [17]. Each of the clusters can be parameterized by its mean (mc). The algorithm consists of two steps namely *assignment step* and *mean update step*. At first, the means of K clusters are initialized with random values. In the *assignment step*, each training data point is assigned to the nearest cluster mean. In the *mean update step*, the cluster means are adjusted by the data points that the clusters contain. The two steps will be continued until all the cluster means remain unchanged after *mean update step*.

III. Evidential SVM for binary classification

EbSVM consists of three steps. First, SoIs are identified and mass values are generated from the SoIs. Then, classification decisions are made using the Bel function. These steps are described in the following subsections and presented in algorithms 1 and 2.

A. Source of Information Generation

SVM gives us support vectors (sv), a weight vector (w) and a bias (b). This weight vector and the bias are used to represent the separating hyperplane. Moreover, we can find two sets of support vectors (sv_+, sv_-) from sv on the basis of two marginal hyperplanes. Data points of a particular class are found in one side of the hyperplane. However, misclassified data points may remain on the other side. To obtain SoIs, we first eliminate these misclassified points. This process is defined in function $ClassifyPoints$ (line 10-25) of Algorithm 1. Applying this function, we get two sets of training data (X_{train+}, X_{train-}).

Another source of information is the clusters of data on either side of the hyperplane. It is usually observed that the density of the data for each class on the either side of the separating hyperplane are not equal. Moreover, there may be several highly dense areas on the either side of the maximum margin hyperplane. We plan to use the distance to high density clusters as SoIs. To capture these dense areas, we apply K-means clustering on X_{train+} and X_{train-} individually. Each cluster from each of the training sets are represented by S_+ and S_- in line 7 of Algorithm 1. In the following discussion, we use the notation S_{ji} for the i-th cluster of the j-th class (θ_j). Note that, for binary classification $\theta_j \in \{+, -\}$.

Algorithm 1: Source of Information Generation

input : Training Data (X_{train}), Training Data Labels (Y_{train}), Regularization Parameter (C), number of clusters (q)

output: List of Positive Support Vectors (sv_+, sv_-), List of cluster means (mc_+, mc_-), Mean distance of Clusters (mcd_+, mcd_-)

1 $svmModel \leftarrow$ Build SVM model with X_{train}, Y_{train}, C
2 $sv \leftarrow$ Get list of support vectors from $svmModel$
3 $w \leftarrow$ Get weight vector from $svmModel$
4 $b \leftarrow$ Get bias or intercept from $svmModel$
5 $(sv_+, sv_-) \leftarrow ClassifyPoints(sv, \emptyset, w, b)$
6 $(X_{train+}, X_{train-}) \leftarrow ClassifyPoints(X_{train}, Y_{train}, w, b)$
7 $(S_+, S_-) \leftarrow$ Apply K-means clustering over X_{train+}, X_{train-} where k=q
8 $(mc_+, mc_-) \leftarrow$ Get cluster means from S_+, S_-
9 $(mcd_+, mcd_-) \leftarrow$ Calculate average distance from cluster mean (mc_+, mc_-) to the points belonging respective cluster using (10)
10 **function** `ClassifyPoints`(X, Y, w, b) **:**
11 Initialize List pts_+ and pts_-
12 **forall** $i \in 1, ...X.length()$ **do**
13 $x = X[i]$
14 $y_{pt} \leftarrow$ Calculate decision using (4)
15 **if** ($Y \neq \emptyset$ *and* $Y_i \neq y_{pt}$) **then**
16 **continue**
17 **end**
18 **if** ($y_{pt} \neq +1$) **then**
19 Insert x to pts_-
20 **else**
21 Insert x to pts_+
22 **end**
23 **end**
24 **return** pts_+, pts_-
25 **end**

We can find two important properties such as cluster mean (mc_{ji}) and density (mcd_{ji}) for each cluster i in class θ_j. Calculation of mean is described in section II. Cluster density is defined by (10). It is easy to see that a smaller value of mcd indicates a denser cluster.

$$mcd_{ji} = \frac{\sum_{x_m \in S_{ji}} |x_m - mc_{ji}|}{\sum_{x_m \in S_{ji}} 1} \tag{10}$$

The process of using these SoIs to generate mass value is described in the next subsection.

B. Mass value Generation

Algorithm 2 presents the mass value generation process. To generate mass value we consider two intuitions. First, a data

Algorithm 2: Mass value Calculation & Decision Making

input : Test data (x_p), number of nearest clusters (k)
output: Predicted class label (Y)

1 $t_+ \leftarrow min$ (all distances from sv_+ to x_p)
2 $t_- \leftarrow min$ (all distances from sv_- to x_p)
3 $smc_+, smcd_+ \leftarrow$ Sort cluster means (mc_+) and avg. distance from cluster mean to the points belonging respective cluster (mcd_+) on the basis of distance from cluster mean and x_p
4 $smc_-, smcd_- \leftarrow$ Sort cluster means (mc_-) and avg. distance from cluster mean to the points belonging respective cluster (mcd_-) on the basis of distance from cluster mean and x_p
5 Initialize a list m of mass values
6 **forall** $i \in \{1, ...k\}$ **do**
7 $d \leftarrow$ Calculate distance from x_p to smc_{+i}
8 $m_{2i-1} \leftarrow$ Generate mass values from using (11), (12), (13), (14) where $t_j = t_+$ and $mcd_{ji} = smcd_{+i}$
9 $d \leftarrow$ Calculate distance from x_p to smc_{-i}
10 $m_{2i} \leftarrow$ Generate mass values from using (11), (12), (13), (14) where $t_j = t_-$ and $mcd_{ji} = smcd_{-i}$
11 **end**
12 $M \leftarrow$ Combine mass values from m using (7)
13 Calculate belief of singleton classes using (8)
14 $Y \leftarrow$ Calculate decision using (15)

point will be more probable to class θ_i than class θ_j when the distances from the k nearest cluster means of class θ_i are less than the distances from the k nearest cluster means of class θ_j. Second, a data point will be more likely to be assigned in class θ_i than class θ_j when the distance from nearest support vector of class θ_i is less than that of class θ_j.

The algorithm 2 proceeds as follows: as new test datapoint x_p arrives, we first find the k nearest clusters ($k \leq q$) from either side of the separating hyperplane. The k nearest clusters of both sides are considered as SoIs. The clusters we have generated in the previous subsection provide us mc_{ji} and mcd_{ji} which will help to generate mass value satisfying the first intuition. Now, suppose that, t_i is the distance of the test data point from the nearest support vector of class θ_i. According to the second intuition, if t_i is less than t_j, then the data point will be more likely to be assigned in class θ_i. Now in line 8 and 10 of algorithm 2, the mass values assigned by cluster S_{ji} computed using (11), (12) and (13) are added to a list m.

$$m(\{\theta_j\}) = \alpha * \Psi_{ji}(d, t_j) \tag{11}$$

$$m(A) = 0 \quad \forall A \in 2^{\Theta} \setminus \{\Theta, \{\theta_j\}\} \tag{12}$$

$$m(\Theta) = 1 - m(\{\theta_j\}) \tag{13}$$

where, $0 < \alpha < 1$ and d represents the distance between the cluster mean mc_{ji}) and the given data point (x_p). In our method, we use function (14) that satisfies the two aforementioned intuitions.

$$\Psi_{ji}(d, t_j) = \exp\left(\frac{-d^2 * t_j}{mcd_{ji}}\right) \tag{14}$$

According to (14), a larger mass value is assigned to θ_j, if x_p is close to both the center of cluster i and nearest sv of class θ_j.

C. Decision Making

For *Decision Making* process, we will combine the mass values from all SoIs using DS theory of evidence. For decision making, we can use belief, plausibility or combined mass value. We assign the given data point to class θ_j when θ_j has the maximum belief value among the singleton classes. The belief and combined mass value are identical for singleton classes. The decision function can be derived in (15).

$$Y = arg \max_{\theta_j} Bel(\theta_j) \tag{15}$$

where $\theta_j \in \Theta$ and θ_j is singleton class.

D. Illustrative Example

Consider the training dataset with 30 datapoints with two features as given in Table I. Fig. 2 presents the dataset and the SoIs to find mass values.

TABLE I
SAMPLE TRAINING DATA FOR EXAMPLE

Training Point	Label	Training Point	Label
(-1.328, 0.035)	-1	(0.408, 1.508)	+1
(-1.051, 0.792)	-1	(1.103, 0.245)	+1
(-1.282, 0.329)	-1	(1.450, 0.792)	+1
(-1.143, 0.456)	-1	(1.566, 0.456)	+1
(-1.259, 0.666)	-1	(1.219, 0.330)	+1
(-1.375, -1.017)	-1	(0.755, -0.386)	+1
(-1.282, -1.311)	-1	(0.871, -1.017)	+1
(-1.236, -1.143)	-1	(1.010, -0.470)	+1
(-1.143, -1.437)	-1	(1.172, -0.470)	+1
(-1.027, -1.227)	-1	(1.265, -0.638)	+1
(-0.194, -0.062)	-1	(0.177, 1.886)	+1
(0.015, -0.933)	-1	(0.640, 1.718)	+1
(-0.448, -0.848)	-1	(0.871, 1.297)	+1
(-0.356, -1.059)	-1	(0.756, 1.129)	+1
(-0.448 -1.101)	-1	(0.293, 1.507)	+1

After applying SVM model (linear kernel with $C = 10$) over the training data, we find a maximum margin hyperplane and list of support vectors sv_+, sv_-.

$sv_+ = \{(0.177, 1.886), (0.871, -1.017)\}$
$sv_- = \{(-0.194, -0.062), (0.015, -0.933)\}$

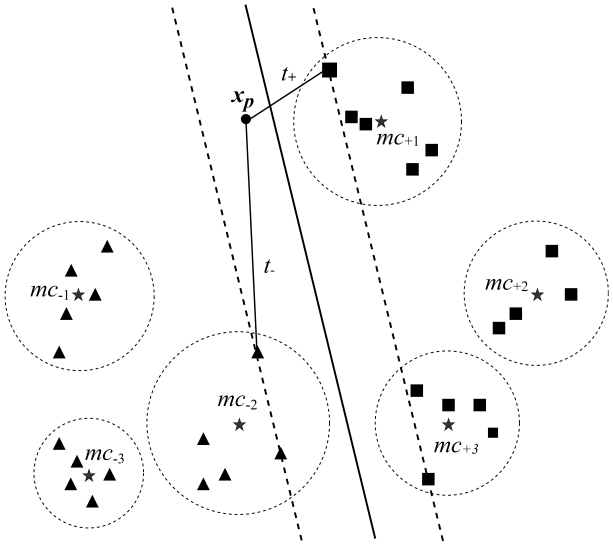


Fig. 2. An illustrative example of plotting data points of binary class problem where ▲ represents point belonging to Class {-1}, ■ represents point belonging to Class {+1}, ● represents test point, solid line represents the separating hyperplane of SVM and marginal hyperplane.

We create 3 clusters in the each side of the hyperplane. mc_+, mc_- represents the cluster means and mcd_+, mcd_- represents the average distance of all the points belonging that cluster from the cluster mean.

$mc_+ = \{(0.524, 1.508), (1.335, 0.456), (1.015, -0.596)\}$

$mc_- = \{(-1.21, 0.46), (-0.286, -0.8), (-1.21, -1.23)\}$

$mcd_+ = \{0.325199, 0.267903, 0.272036\}$

$mcd_- = \{0.247776, 0.370238, 0.173745\}$

When a new test point $x_p = (-0.323, 1.5)$ has come, we find one nearest cluster from each side of the hyperplane and generate mass values from these sources. The mass values from the two sources and combination of the mass values are shown in Table II.

TABLE II
BASIC BELIEF ASSIGNMENT FROM DIFFERENT SOURCE OF INFORMATION

Source	m({+1})	m({-1})	m({+1, -1})
1	0.23572	0.0	0.76428
2	0.0	0.0000006	0.999994
Combined	0.23572	0.000005	0.7642755

After we have calculated the belief of each singleton class, we get the following values.

$Bel(\{+1\}) = 0.23572$

$Bel(\{-1\}) = 0.000005$

When we apply decision function described in (15), we find a decision that is $Y = \{+1\}$. The test point will belong to class {+1}.

TABLE III
PROPERTIES OF DATASETS USED IN OUR EXPERIMENTS

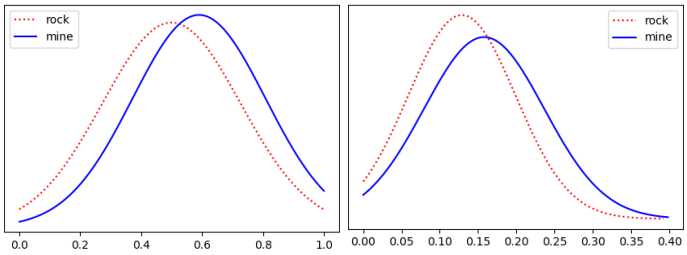
Dataset	Number of instances	Number of features
Ionosphere	351	34
Madelon	2600	500
Musk	473	165
Parkinsons	195	23
Pima	786	8
Sonar	208	60

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The dataset description, implementation details and a comparison of EbSVM with the state-of-the-art methods along with discussion are given in the following subsections.

A. Dataset Description

Six benchmark datasets are collected from UCI machine learning repository [18]. The characteristics of the datasets are illustrated in Table III. The features of these datasets are real and most of the cases, they are imprecise. For example, if we look at the Sonar dataset, most of the cases, the data distribution of the features are overlapping which are shown in Fig. 3. Features having these characteristics are difficult to classify.



(a) (b)

Fig. 3. Data distribution of Rock and Mine class with respect to (a) feature 20 and (b) feature 8 in $Sonar$ dataset

B. Implementation Details

The performances of EbSVM are compared to that of SVM, EcSVM, PSVM. In our experiment, we standarize the features of the chosen datasets using z standarization (mean is 0 and variance is 1). For all methods in the experiment, we use linear kernel and $C = 1.0$ as generalization parameter. The value of k, q, α are set in this experiment at 3, 15, 0.95 for EbSVM respectively. The average of 10-fold stratified cross validation are used as a performance metric and are reported for all the methods. The standard accuracy measure is used in this experiment. We have also performed paired t-test (at 5% level of significance) to show whether a method shows significantly better compared to other methods in terms of accuracies.

C. Result and Discussion

TABLE IV
COMPARISON OF EBSVM AND THREE OTHER METHODS

Dataset	SVM	EcSVM	PSVM	EbSVM
Ionosphere	86.35	86.61	87.48	**89.18**
Madelon	55.88*	55.80*	55.62*	**62.38**
Musk	87.77	**88.38**	87.13	**88.38**
Parkinsons	83.73*	86.26	84.75	**87.84**
Pima	77.21	**77.35**	77.21	76.43
Sonar	72.68*	73.10*	73.59*	**83.28**
Win/Tie/Loss	5/0/1	4/1/1	5/0/1	-
Sig. Win/Loss	3/0	2/0	2/0	-

* represents that EbSVM wins significantly against that method

In Table IV, accuracies of SVM, EcSVM, PSVM and EbSVM over the selected datasets are shown. We bold face the accuracies of the method which performs better than that of other methods. Besides this, accuracies are marked with asterisk(*) when EbSVM performs significantly better than that method. The second last line of Table IV summarizes Win/Tie/Loss of different methods. Here, Win means EbSVM performs better than that of other method, Tie means EbSVM performs as same as that method and Loss means EbSVM performs worse than that method.

From Table IV, it is clearly observed that, EbSVM outperforms the other state-of-the-art methods. For example, EbSVM wins five times and losses onoy lne time while comparing with SVM. It means that EbSVM obtains better accuracies for five datasets than SVM and worse accuracy for one dataset. In comparison with EcSVM, the average accuracy is same for Musk dataset. From last line of Table IV, we observe that EbSVM performs significantly better for three datasets than SVM, and EbSVM doesn't perform worse at all.

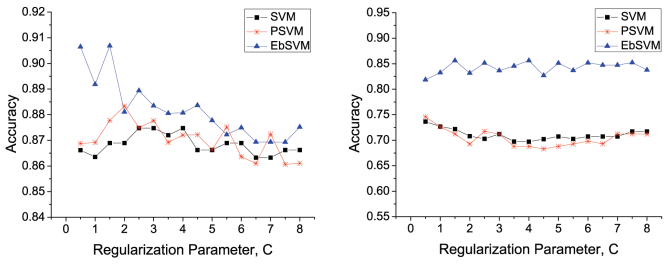


Fig. 4. Influence of regularization parameter, C on accuracy. (a) on $Ionosphere$ dataset; (b) on $Sonar$ dataset

It is easily understandable that, regularization parameter, C can play an influence over accuracy of these classifiers. The influences of choosing this parameter is shown in Fig. 4 for two different datasets namely $Ionosphere$ and $Sonar$. From Fig. 4, it is observable that, EbSVM reasonably performs better compared to SVM and PSVM for different values of C.

V. CONCLUSION

In this work, we present an approach of calculating mass values using SVM. These mass values are used to generate final decision. The solution is intuitive and performs better than the existing state-of-the-art methods over six datasets. However, one important limitation is that the mass value generation process proposed in this paper is defined in ad-hoc manner. A future research direction is to justify this mass value generation process. An extension of EbSVM to multi-class problem can also be considered in future.

ACKNOWLEDGMENT

This research is supported by the University Grants Commission, Bangladesh under the Dhaka University Teachers Research Grant No Reg/Admin-3/61070.

REFERENCES

[1] Evelyn Fix and J. L. Hodges. Discriminatory analysis. nonparametric discrimination: Consistency properties. *International Statistical Review*, 57(3):238–247, 1989.
[2] Leo Breiman. *Classification and regression trees*. Routledge, 2017.
[3] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.
[4] Pedro Domingos and Michael Pazzani. On the optimality of the simple bayesian classifier under zero-one loss. *Machine Learning*, 29(2):103–130, Nov 1997.
[5] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338 – 353, 1965.
[6] George J. Klir and Bo Yuan. *Fuzzy sets and fuzzy logic - theory and applications*. Prentice Hall, 1995.
[7] Peter Walley. Towards a unified theory of imprecise probability. *Int. J. Approx. Reasoning*, 24(2-3):125–148, 2000.
[8] Glenn Shafer. *A mathematical theory of evidence*, volume 42. Princeton university press, 1976.
[9] Didier Dubois and Henri Prade. Ranking fuzzy numbers in the setting of possibility theory. *Inf. Sci.*, 30(3):183–224, 1983.
[10] John C. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in large margin classifiers*, pages 61–74. MIT Press, 1999.
[11] Philippe Xu, Franck Davoine, Hongbin Zha, and Thierry Denoeux. Evidential calibration of binary SVM classifiers. *Int. J. Approx. Reasoning*, 72:55–70, 2016.
[12] T. Denoeux. A k-nearest neighbor classification rule based on dempster-shafer theory. *IEEE Transactions on Systems, Man, and Cybernetics*, 25(5):804–813, May 1995.
[13] Simon Petit-Renaud and Thierry Denoeux. Nonparametric regression analysis of uncertain and imprecise data using belief functions. *Int. J. Approx. Reasoning*, 35(1):1–28, 2004.
[14] Hicham Laanaya, Arnaud Martin, Driss Aboutajdine, and Ali Khenchaf. Support vector regression of membership functions and belief functions - application for pattern recognition. *Inf. Fusion*, 11(4):338–350, October 2010.
[15] Philippe Xu, Franck Davoine, and Thierry Denœux. Evidential logistic regression for binary svm classifier calibration. In *International Conference on Belief Functions*, pages 49–57. Springer, 2014.
[16] John A Hartigan and Manchek A Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.
[17] John A Hartigan. Clustering algorithms. 1975.
[18] Uci machine learning repository. http://mlr.cs.umass.edu/ml/datasets.html. Last Accessed on 15-08-2018.

Sentiment Analysis of Bengali Texts on Online Restaurant Reviews Using Multinomial Naïve Bayes

Omar Sharif[1], Mohammed Moshiul Hoque[2] and Eftekhar Hossain[3]
Department of Computer Science and Engineering[1,2]
Department of Electronics and Telecommunication Engineering[3]
Chittagong University of Engineering and Technology[1,2,3]
Chittagong-4349, Bangladesh
omar.sharif1303@gmail.com[1], mmoshiulh@gmail.com[2], eftekhar.13ete@gmail.com[3]

***Abstract*—Recently, determining the customer impression is considered one of the prominent factors on the success of the restaurant businesses. Due to the rapid growth of digital contents related to restaurant or foods in the web, people are more inclined on reviews before going to any restaurant so the significance of customer review is inevitable. In order to selects a restaurant customer needs to check thousands of feedback's to understand the restaurant quality or services. Therefore, classification of a significant amount of reviews into a sentimental category is required to attain meaningful insights so that the customer can choose restaurants based on their preferences. This classification can be done by sentiment analysis. This paper proposes a system that can classify customer reviews into positive and negative classes based on their sentimental feedback. We have tested the proposed system with 1000 restaurant reviews text written in Bengali. The experimental result shows that the proposed the system can classify restaurant reviews with 80.48% accuracy using multinomial Naïve Bayes.**

***Keywords*—Bangla language processing, Sentiment analysis, Customer reviews, Machine learning.**

I. Introduction

The amount of people has increased significantly who are taking their foods or meals in restaurants due to rise their financial capability, miscellaneous food habits and improved lifestyle. Therefore, restaurant business have achieved remarkable popularity in recent years. Due to the rapid growth of Internet contents, customers has affinity to check the reviews of a restaurant before visit it. Thus, reviewing a restaurant from text comments become a common scenario.

The key objective of this paper is to build a supervised learning model to categorize the customer's feedback's in terms of positive and negative sentiments which are written in Bengali sentences. Any restaurant with plenty of positive reviews can attain a symbol of faith, ensure quality of foods and services among the customers. Conversely, it is very difficult to attract new customers as well as retain existing customers by a restaurant without positive reviews or feedback. Usually, more positive reviews conveys trustworthiness to customers. Thus, people appreciate the experience of the consumer and consent to others and the review on a restaurant is the only way to comprehend others opinion on the restaurant. Opinions collected from users experiences regarding specific criteria of a restaurant like food quality, environment and service quality, straightforwardly make an impact on the future customer decisions. Likewise, negative reviews often yield disinclination on the customer choices. Therefore, restaurants needs to take into account the customer's feedback for future improvements and attracts much attention of customers. This might totally unusual to check every customer reviews one by one as it requires ample amount of money and innumerable manpower to lead such a surveys. Taking consideration the rapid growth of visitors and users requirements, this imposes a concealment for a system that can understand the attitude of a reviewer comments written in various social media communities or blogs. In that case, a term sentiment analysis come uppermost that aim is to uncover how customer sense about a certain foods.

The sentiment analysis is used to attain a grasp of sentimental note behind the person's reviews. It is also known as opinion mining and it determines whether the sentiment of a person comment is positive, neutral or negative based on their writing. Many researches in sentiment analysis have been done in many languages like Chinese [1][2], Vietnamese [3], Arabic [4], Thai [5], Myanmar [6] and Bengali [7]. Research on Bengali text based opinion mining of restaurant reviews is in rudimentary stage still now. This paper presents an automated system for sentiment analysis of Bengali text for restaurant reviews to classify an opinion into two classes: positive or negative sentiment. Three machine learning algorithms such as decision tree, random forest and multinomial naïve Bayes classifier are implemented to classify the sentiment from Bengali review text.

II. Literature Review

The sentiment analysis of customer reviews has become one of the interested research area. Some of the research work by noted scholars on restaurant reviews on different languages are studied and discussed below. Aye et.al developed a lexicon-based sentiment analysis model in Myanmar text. On their work, the customer reviews on food and restaurant quality was the prime insight for the sentiment analysis. The sentiment classification was done by using a sentiment dictionary which contains a list of sentiment words. Their system successfully distinguished data for the reviews on Myanmar Language at

96% overall accuracy over 500 customers reviews on food and restaurant domain [6].

Kaviya et.al developed a sentiment analysis model where the rating system of a restaurant was done depending upon the user reviews. Their system splits user comments to find sentiment keywords and associates the comment with a sentiment rank. The reviews were categorized into positive and negative class and an overall rating was assigned to the restaurant. The Yelp dataset was taken for the testing of their developed system.

Perera et.al adopted an aspect-based opinion mining for restaurant reviews. The aspect-based mining finds a sentence aspect and the opinion of the users comments on these aspects to make a positive and negative review. SentiWordNet a lexical dictionary was used to determine the positive and negative polarity value for each opinion word. Dependency parser was used for the opinion word extraction. The evaluation of their developed system was done on both manually and systematically. From the comparison of the two approaches they found that their developed system provided a satisfactory accuracy of 70% [8].

Doan et.al developed a sentiment analysis model for customers reviews where a variant of online random forests was taken to perform the analysis. Their proposed model was an ensemble of tree-based classifiers and a portion of the restaurant reviews from the Yelp dataset was taken for the sentiment analysis. For the experiment, they set up an ensemble learning model for 100 trees using Online Random Forests (ORF) and incremental random forest (proposed approach) (iRF). Their model has longer run time than iNB because incremental learning was adopted in their design that induces trees instance by instance [9].

Sarkar et.al developed a Bangla sentiment analysis model where Bengali tweet dataset released for SAIL contest 2015 was used for classification. They used Multinomial Naïve Bayes and SVM classifiers for the classification of sentiment polarity expressed in tweets. N-gram and SentiWordnet features were used as their feature set.Their devleloped model showed that SVM classifier trained with unigram and SentiWordNet features performs best on Bengali tweet dataset [10]. Gan et.al designed a model for restaurant rating based on sentiment analysis. In their work, in addition to food, ambience and service, the overall rating of the restaurant also focuses on the pricing and special contexts but there was no emphasizing done on the scores [11]. Lak et.al developed a model to compare sentiment analysis results with star ratings in three different domains [12]. Xing Fang et.al developed a system for sentiment analysis of product reviews which used the sentiment polarity categorization process. They selected naïve bayesian, support vector machine and random forest classification methods for categorization. For the sentiment polarity classification a sentiment score formula was taken to generate the feature vector for a sentiment and based on this score they performed the sentiment polarity categorization [13]. Lada Banic et. al designed a hotel review mining system using natural language processing and machine learning. Number of negative and positive terms or phrases are used for the generation of cumulative information on the level of final evaluation. A grade from 1 to 5 in which 1 and 5 is referred as bad and excellent respectively was assigned in final evaluation[14]. Singla et.al done sentiment analysis on unstructured data of Mobile phone reviews. They used decision tree, naïve Bayes, SVM, and cross validation for the purpose of classification. Among these classification models they concluded that svm provides highest accuracy in opinion mining of product reviews [15].

III. Dataset Accumulation

Success of any machine learning system heavily depends on the dataset used to train that system. But unfortunately, it is really difficult to build a corpus which contains a large amount of Bangla text reviews on restaurant because there is lack of resources in Bengali language as well as researchers do not publish their dataset. For the purpose of research, Bengali benchmark dataset is not yet published. So an English benchmark dataset have taken [12] and translated into Bengali to build the corpus. Fig.1 shows sample positive and negative restaurant reviews.

	Sample Reviews	Sentiment
English	Preferably hot and spicy chicken fry. Excellent environment and hospitality. I'll give it a rating in the top restaurant in Dhaka.	Positive
Corresponding Bengali	গরম এবং মসলাযুক্ত চিকেন ফ্রাইটা পছন্দনীয়। চমৎকার পরিবেশ এবং আতিথেয়তা। আমি এটিকে ঢাকায় শীর্ষস্থানীয় রেস্তোরাঁর মধ্যে রেটিং দিব।	
English	The grill was cold, the bread was dry and strong, the stock was old ... very bad food ... frustrating ...	Negative
Corresponding Bengali	গ্রিলটা ঠান্ডা ছিল, রুটিও শুকনো এবং শক্ত ছিল, স্টকটা পুরাতন ছিল ... খুবই খারাপ মানের খাবার ... হতাশ...	

Fig. 1. Dataset Sample

Besides, different public pages [16], groups [17][18][19] of Facebook are used to collect data. As most part of the dataset collected manually some negligible inconsistencies may occur in it. Collected text documents are used to train the system which will be used for further classification of reviews. Dataset statistics is summarized in Table I.

TABLE I: Data Summary

	Training set	Testing set
Number of documents	800	200
Number of sentences	2559	628
Number of words	8236	2497
Total unique words	1069	358

IV. Proposed System Framework

Fig. 2 shows an abstract view of proposed system. Different statistical machine learning techniques are used to train the system. The overall framework is divided into three phases: training set preparation, feature extraction and classification respectively.

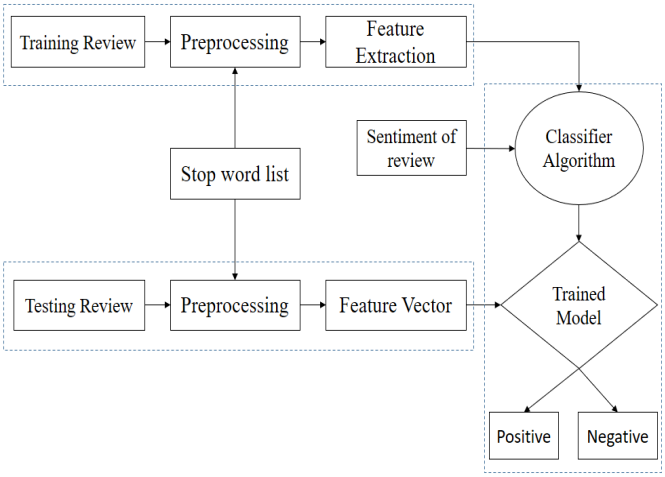


Fig. 2. Proposed System Framework

A. *Training Set Preparation*

Training set $R = \{r_1, r_2, r_3, ..., r_n\}$ consists of n training reviews. Each review either contains positive of negative sentiment where C_p and C_n are used to denote positive and negative class respectively. A review r_i with l words is represented by a word vector $W[] = \{w_1, w_2, w_3, ..., w_l\}$. All reviews are preprocessed to remove inconsistencies from dataset. Punctuation's are also removed in this step. Table II represents word vector$W[i]$ of a random review r_i.

TABLE II: Word vector of a review

r_i	w_1	w_2	w_3	...	w_l

A list $S[] = \{s_1, s_2, s_3, ..., s_t\}$ with t stopwords have been developed to remove the words w_i which has no contribution in deciding whether a review r_i conveys positive (C_p) or negative (C_n) sentiment. Stop words $s_1, s_2, s_3, ..., s_t$ are removed from a review by finding the match from stop word list $S[]$. Conjunctions, prepositions, interjections, pronouns, suffixes and prefixes are considered as stop words.

	Type	Example
s_1	pronoun	সে
s_2	conjunction	এবং
s_3	preposition	থেকে
...	...	...
s_r	interjection	সাবাশ

Fig. 3. Sample stop words

B. *Feature Extraction*

A vocabulary of Bengali words is created by tokenizing reviews of the corpus. "$tfidf$" : the term frequency-inverse document frequency statistic is used as features of the proposed model. Our used "$tfidf$" statistic is,

$$tfidf(w, r) = tf(w, r) \log \frac{N}{|\{r \epsilon R : w \epsilon r\}|} \quad (1)$$

Here,
$tfidf(w, r)$ = value of word w in review r,
$tf(w, r)$ = frequency of word w in review r,
N = total number of reviews,
$|\{r \epsilon R : w \epsilon r\}|$ = number of reviews containing w.

TABLE III: A small fragment of feature matrix

r \ c	w_1	w_2	w_3	w_4	...	w_v
r_1	0.21	0.13	0.20	0.16	...	0.10
r_2	0.18	0.36	0.17	0.21	...	0.31
r_3	0.17	0.13	0.14	0.15	...	0.19
r_4	0.24	0.40	0.48	0.47	...	0.41
...	...	...	...	...	...	...
r_u	0.25	0.32	0.46	0.71	...	0.48

Table III shows the feature matrix of proposed system. Features $(F[][])$ of the reviews are represented by a two dimensional$(U \times V)$ feature matrix. Here, rows represent the reviews $r_1, r_2, r_3, ..., r_u$ and columns represents unique words $w_1, w_2, w_3, ..., w_v$ respectively. Each cell of the feature matrix $(F[][])$ represent the "$tfidf$" value of word w_v occurs in a specific review r_i. Each row represents the feature vectors $F[1], F[2], F[3], ..., F[n]$ for the reviews $r_1, r_2, r_3, ..., r_n$ of the corpus.

C. *Classification*

As, the goal of the proposed system is to classify the reviews thus classification is the most significant part of the system. Extracted features from the reviews are used to train the proposed model that could classifies reviews into positive and negative sentiment. Mainly, the proposed system is implemented by using multinomial naïve bayes.

Multinomial naïve bayes [20] performs really well when multiple occurrences of the words have significant impact in the classification problem. It can be defined by bayes theorem where all the variables $V_1, V_2, V_3, ..., V_n$ in a given class C are conditionally independent with each other given C. Using bayes rule for a text review (R) and class (C),

$$P(R|T) = \frac{P(R|C)P(C)}{P(R)} \quad (2)$$

Final equation,

$$C_{MAP} = argmax P(R_1, R_2, ..., R_n|C)P(C) \quad (3)$$

A **decision tree** [21] has external and internal nodes. Decision classes are represented by external nodes and internal nodes correspond to attribute which are used for making decision. Entropy calculation is used to build a decision tree,

$$E(S) = \sum_{c \epsilon X} P(C)E(C) \tag{4}$$

Random forest is mainly collection of decision trees with some differences. Decision trees make predictions by formulating some set of rules but in random forest algorithm decision trees are built by selecting features randomly and finally acquired results are averaged. But for real time predictions a large number of trees can make this algorithm slow and ineffective.

Trained classifier model is tested with a test set, $TR = \{tr_1, tr_2, tr_3, ..., tr_x\}$ with x test documents. Feature extraction method is used to extract features $(F[][] = F[1], F[2], ..., F[x])$ from the test set TR. Proposed model use these features to classify a review tr_i as positive(C_p) or negative (C_n).

V. Evaluation Measures

A number of statistical and graphical measures have been adopted for evaluating the performance of the proposed system. Statistical measures such as confusion matrix, precision, recall, f_1 score and graphical measures such as precision-recall curve, receiver operating characteristics (ROC) curve have used for the measurement.

- A **confusion matrix** is used for evaluating the performance of a classification model. As the proposed system is designed for binary classification, thus the confusion matrix of the system must have two columns and two rows. This matrix indicates the number of true positives, false positives, true negatives and false negatives.
  - True Positive (TP): Reviews that conveys positive sentiment and also classified as positive.
  - True Negative (TN): Reviews that conveys negative sentiment and also classified as negative.
  - False Negative (FN): Reviews that conveys positive sentiment but classified as negative.
  - False Positive (FP): Reviews that conveys negative sentiment but classified as positive.

  This rates are used to calculate other evaluation measures.
- **Precision** refers as positive predictive value. Precision can be obtained form the following equation.

$$Precision = \frac{TP}{FP + TP} \tag{5}$$

  If precision is high then the algorithm is doing well.
- **Recall** is the ratio of correctly classified positive reviews to the total number of positive reviews. Recall can be obtained from the following equation.

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

  High value of precision and Recall are vital for a model.
- f_1**-score** of algorithms need to be calculated for choose a particular learning algorithm between several algorithms. f_1-score can be obtained from the following equation,

$$f_1 - score = \frac{2 * precision * recall}{precision + recall} \tag{7}$$

VI. Results and Analysis

Proposed model has been tested with random forest, decision tree and multinomial naïve bayes algorithms. K-fold cross validation has been used to test the performance of the system. Table IV presents a comparison of performance of this algorithms for different values of K.

TABLE IV: Performance Comparison

K-value	Classifier	Precision	Recall	f_1 score	Accuracy
5-fold	Random Forest	0.75	0.76	0.75	75.51
	Decision Tree	0.71	0.70	0.71	70.40
	Proposed	0.81	0.80	0.79	79.59
6-fold	Random Forest	0.79	0.78	0.77	78.04
	Decision Tree	0.70	0.70	0.70	69.51
	Proposed	0.82	0.80	0.80	**80.48**
7-fold	Random Forest	0.70	0.70	0.70	70.00
	Decision Tree	0.71	0.72	0.71	71.42
	Proposed	0.80	0.79	0.78	78.57

Same number of training and test reviews used for all algorithms. Table IV shows that multinomial naïve byes with "tfidf" feature outperforms other algorithms for every fold size K. Different value of K has been used for finding the optimum value of K which gives maximum accuracy. Initially, 2-fold cross validation has been selected and then by increasing the value of k we observed that after a certain value of K the accuracy become turns down. The reason is that, the increasing the number of folding size will decrease the number of testing data in the individual fold that effects the accuracy. From this discussion it can be concluded that, proposed system gets maximum accuracy of 80.48% for 6-fold cross validation.

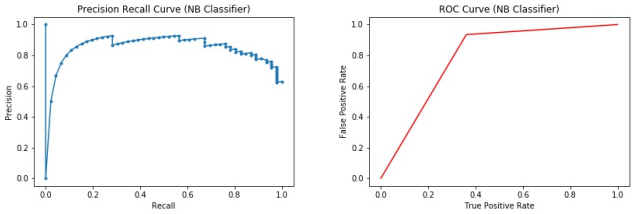
Precision, recall and f_1-score of each class can be presented by a classification report. This report helps for further analysis of an algorithm. In table V classification report of proposed algorithm is presented.

Classification report shows value of f_1-score is 0.74 and 0.84 for positive and negative class respectively. That means proposed system can correctly classify 74% of negative and 84% of positive reviews correctly. For the further evaluation

TABLE V: Classification Report (Proposed algorithm)

Class (C)	Precision	Recall	f_1-score	Support
Negative (C_n)	0.88	0.64	0.74	87
Positive (C_p)	0.77	0.93	0.84	113
avg./total	0.82	0.80	0.80	200

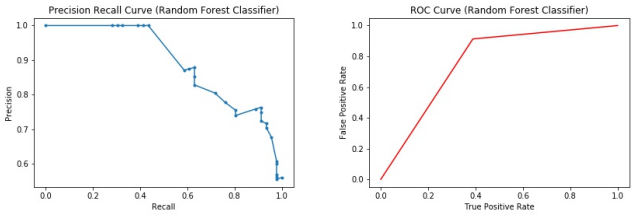
of proposed model graphical measures such as precision-recall curve and receiver operating characteristics curve have been used. Fig. 4 to Fig. 6 show these measures for the tested algorithms.



(a) precision-recall **(b)** ROC

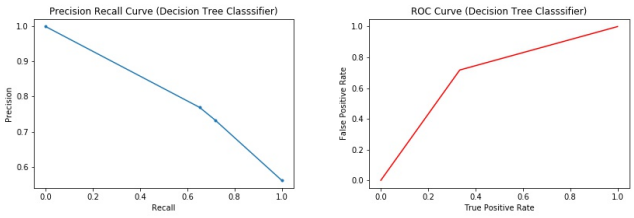
Fig. 4. Result of Multinomial Naïve Bayes

Fig. indicates the precision value of the multinomial naïve bayes is increasing for a certain margin of recall value. Fig. 5(a) and Fig. 6(a) show increase in recall make an adverse effect in precision.



(a) precision-recall **(b)** ROC

Fig. 5. Result of Random Forest



(a) precision-recall **(b)** ROC

Fig. 6. Result of Decision Tree

Fig. 4(b), 5(b) and 6(c) show that true positive rate linearly increase for certain value of false positive rate after that rate of increase becomes almost constant. The trade-off between

Sample Text = > এরাবিয়ান মাস্টারের অভ্যন্তরীণ সজ্জাটা অনন্য। এদের খাবার সুস্বাদু এবং একটু মসলাযুক্ত .. তাদের কর্মীরা খুব বন্ধুত্বপূর্ণ।
System output: Positive
Desired output: Positive

Sample Text = > অপূর্ব খাবার, গ্রেট সার্ভিস... !!
System output: Positive
Desired output: Positive

Sample Text = > আমার জীবনের খাওয়া সবচেয়ে খারাপ খাবার। পুরো টাকাটাই গোল্লায় গেল।
System output: Negative
Desired output: Negative

Sample Text = > পিজার স্বাদটা খুবই জঘণ্য ছিল।
System output: Negative
Desired output: Negative

Sample Text = > খাবারের পরিমাণ কম কিন্তু দাম খুব বেশী। খাবারটা আন্তর্জাতিক মানের নয় কিন্তু স্থানটা ভাল ছিল।
System output: Negative
Desired output: Neutral

Fig. 7. Sample Output

the true positive rate and the positive predicted value is summarized by precision-recall curve and receiver operating characteristics curve summarizes the trade-off between the true and false positive rate for different probability thresholds. Fig. 7 shows some sample restaurant reviews classified by the proposed system.

VII. Conclusion

In this paper a sentiment analysis scheme is presented for restaurant reviews on Bangla language. In this work, different machine learning algorithms have been employed to classify the underlying sentiment of the customer reviews for Bangla text. Besides this, an accuracy comparison of the used models has also been showed where multinomial naïve bayes obtained maximum 80.48% accuracy for 6-fold cross validation among all other used algorithms such as Random forest and Decision Tree. Sentiment analysis has become very effective for the restaurant owners as it can helps owner by providing an essence that what customers think about his restaurant. For sentiment analysis of restaurant reviews many works have been done in English. On the contrary, very few works are done in Bengali. This work is a little effort to compensate the deficiency. In future, more powerful algorithms can be used to find the semantic connection between words of a text that would help us to provide sentiment accurately for more sentiment category. To achieve this goal, increasing the size of dataset would be our prime concern.

References

[1] S. J. Wu, R. D. Chiang, and Z. H. Ji, "Development of a chinese opinion-mining system for application to internet online forums," *The Journal of Supercomputing*, vol. 73, no. 7, pp. 2987–3001, 2017.

[2] Z. Li, L. Liu, and C. Li, "Analysis of customer satisfaction from chinese reviews using opinion mining," in *2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, IEEE, 2015, pp. 95–99.

[3] H. S. Le, T. Van Le, and T. V. Pham, "Aspect analysis for opinion mining of vietnamese text," in *2015 international conference on advanced computing and applications (ACOMP)*, IEEE, 2015, pp. 118–123.
[4] S. Atia and K. Shaalan, "Increasing the accuracy of opinion mining in arabic," in *2015 First International Conference on Arabic Computational Linguistics (ACLing)*, IEEE, 2015, pp. 106–113.
[5] T. Chumwatana, "Using sentiment analysis technique for analyzing thai customer satisfaction from social media," 2015.
[6] Y. M. Aye and S. S. Aung, "Sentiment analysis for reviews of restaurants in myanmar text," in *2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, IEEE, 2017, pp. 321–326.
[7] S. Chowdhury and W. Chowdhury, "Performing sentiment analysis in bangla microblog posts," in *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*, IEEE, 2014, pp. 1–6.
[8] I. Perera and H. Caldera, "Aspect based opinion mining on restaurant reviews," in *Computational Intelligence and Applications (ICCIA), 2017 2nd IEEE International Conference on*, IEEE, 2017, pp. 542–546.
[9] T. Doan and J. Kalita, "Sentiment analysis of restaurant reviews on yelp with incremental learning," in *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, IEEE, 2016, pp. 697–700.
[10] K. Sarkar and M. Bhowmick, "Sentiment polarity detection in bengali tweets using multinomial naïve bayes and support vector machines," in *Calcutta Conference (CALCON), 2017 IEEE*, IEEE, 2017, pp. 31–36.
[11] Q. Gan and Y. Yu, "Restaurant rating: Industrial standard and word-of-mouth–a text mining and multi-dimensional sentiment analysis," in *System Sciences (HICSS), 2015 48th Hawaii International Conference on*, IEEE, 2015, pp. 1332–1340.
[12] P. Lak and O. Turetken, "Star ratings versus sentiment analysis–a comparison of explicit and implicit measures of opinions," in *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, IEEE, 2014, pp. 796–805.
[13] X. Fang and J. Zhan, "Sentiment analysis using product review data," *Journal of Big Data*, vol. 2, no. 1, p. 5, 2015.
[14] L. Banić, A. Mihanović, and M. Brakus, "Using big data and sentiment analysis in product evaluation," in *Information & Communication Technology Electronics & Microelectronics (MIPRO), 2013 36th International Convention on*, IEEE, 2013, pp. 1149–1154.
[15] Z. Singla, S. Randhawa, and S. Jain, "Sentiment analysis of customer product reviews using machine learning," in *Intelligent Computing and Control (I2C2), 2017 International Conference on*, IEEE, 2017, pp. 1–5.
[16] *Facebook page foodbank bd*. [Online]. Available: https://www.facebook.com/foodbankbd/.
[17] *Food bloggers bd*. [Online]. Available: https://www.facebook.com/groups/foodbloggersbd/.
[18] *Food lovers bd*. [Online]. Available: https://www.facebook.com/groups/230918710613193/.
[19] *The food guide chittagong*. [Online]. Available: https://www.facebook.com/groups/1879200579030221/.
[20] J. Y. Yoo and D. Yang, "Classification scheme of unstructured text document using tf-idf and naive bayes classifier," in *Computer and Computing Science 2015*, 2015, pp. 263–266.
[21] G. S. Chavan, S. Manjare, P. Hegde, and A. Sankhe, "A survey of various machine learning techniques for text classification," *International Journal of Engineering Trends and Technology*, vol. 15, no. 6, 2014.

Clinical Human Gait Classification: Extreme Learning Machine Approach

[1] Prithvi Patil, [2]K.Shusheel Kumar,[3]Neha Gaud and [4]Vijay Bhaskar Semwal*
[1]National Institute of Technology, Srinagar
[2]B.I.E.T. Jhansi
[3]Vikram University Ujjain
[4]Maulana Azad National Institute of Technology, Bhopal
prithvipatil1357@gmail.com,sus.iiita.932@gmail.com, gaud28neha@gmail.com,vsemwal@gmail.com*

Abstract: This study reports a novel approach for biometric gait pattern classification using Extreme Learning Machine (ELM) algorithm. Clinical gait analysis can be used for early detection of gait abnormality in brain or neurological disorder subjects. In many cases gait abnormality cannot be detected through visual observation alone, but becomes apparent only in a quantitative analysis of subject's gait. This can also help us understand the neuro-muscular mechanics associated with brain disorders. Human gait is also of profound interest to the research community in the field of biometric identification and bipedal robot locomotion due to its uniqueness and efficiency. This paper explores multi-class gait classification using four machine learning methods (KNN, SVM, ELM, MLP) and evaluates their performance for multi class gait classification. The proposed method achieves very good results. TheELM is used first time in to analyses the neuromuscular of patients suffering from multiple sclerosis and stroke

Keywords: Extreme Learning Machine (ELM), Biometric identification, Gait classification, Bipedal walk, Multiple sclerosis, Stroke gait, Clinical gait analysis, Cerebral palsy, Crouch gait, Anthropomorphic robot.

I. INTRODUCTION

Human gait is a complex locomotion for forward propulsion of center of gravity of the human body. It is achieved through a process of continuous learning as a child interacts with his environment [1]. It is highly nonlinear and complex due to varying configuration during different sub phases of human gait [2]. The human walk is designed as 8 discrete sub phases [3] with continuous behavior and researchers have started considering it as hybrid system [4]. Fig.1 shows the 8 different phases of bipedal walk. It is inherently unstable and nonlinear due to high degrees of non-linearity, high dimensionality, under actuation (in swing phase) and over actuation (in stance phase) [5]. Human gait is also unobtrusive compared to other biometric identifiers such as fingerprint [6]. Many Researchers have used human gait for purposes such as biometric identification, understanding the problem of elderly people, to understand the biomechanics of human walk, artificial limb development and bipedal walk [7] [8].

As on today, due to its inherent complexities, bipedal robots are not efficient to work outside laboratory environment i.e. unstructured environments designed for humans and are controlled as fully-actuated system, which is not energy efficient [9]. Among various types of humanoid locomotion, bipedal walking is the most natural, energy efficient and interesting, since bipedal robots have more potential to move in rugged terrains or complex environments, where wheeled or tracked robots cannot operate. On the other hand, bipedal robots are less stable and prone to falling down. To meet this challenge, over the years, many solutions have been proposed [10].

Gait can be studied as non-linear time series of kinematic trajectories [11],[12]. The analysis involves the investigating sensory motor interaction and understanding the biomechanics of locomotors system [13]. Gait classification allocates walking patterns into groups that can be identified and differentiated from one another based on a set of defined variables or features. We believe the problem being addressed so far using conventional mechanics based model and automated control theory can effectively be addressed using data driven computational theory. Throughout this work we tried to validate this hypothesis using machine learning techniques including ELM for patients suffering from multiple sclerosis and stroke. The ELM is used first time in to analyses the neuromuscular of patients suffering from multiple sclerosis and stroke

The paper is organized as follows. Section II is about methodology. Section III introduces ELM and briefly describes other machine learning approaches used in this study. Section IV discusses performance evaluation and results achieved by different machine learning algorithms. Section V is about Conclusion and Future work.

II. METHODOLOGY

A. The Proposed System:

Proposed system consists of eight phases which is shown in the Fig.1. They are gait data collection, gait data detection, trajectories smoothing, feature extraction, feature selection and classification using ML algorithms [14]. Figure 2 depicts a systematic approach for proposed gait recognition process.

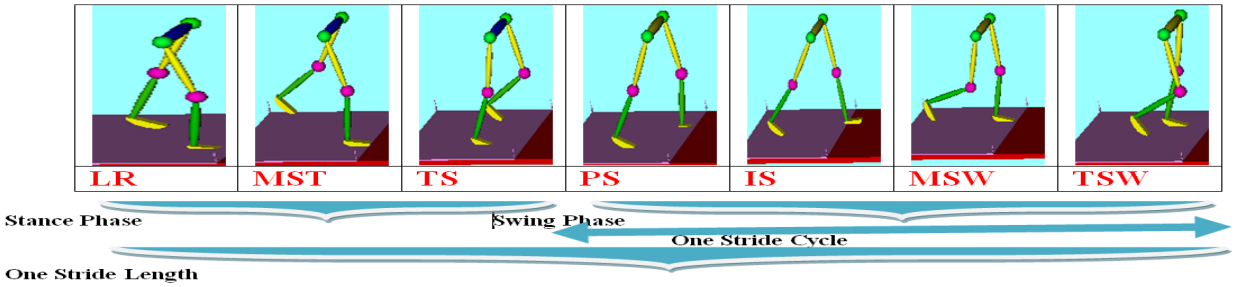


Fig.1. Breakdown of Human Gait into different Discrete Sub Phases



Fig.2. Proposed Architecture of the System

The Gait is divided into 2 phases, namely stance phase and swing phase. The swing phase is critical since only one leg needs to carry the entire body load. We consider it as discrete phase and the configuration during this phase will be over actuation. The swing phase can be considered as continuous phase as well. The configuration is under actuation during this phase.The Percentage-wise distribution of gait time series as following:

Stance phase:

1. Initial Contact– IC[0-2%]
2. Loading Response– LR[2-10%]
3. Mid Stance– MS[10-30%]
4. Terminal Stance– TS[30-50%]

Swing phase:

5. Pre Swing– PS[50-60%]
6. Initial Swing– IS[60-73%]
7. Mid Swing - MS[73-87%]
8. Terminal Swing – TS[87-100%]

B. Data Collection:

For class Multiple sclerosis and Stroke gait, data was collected and recorded using accelerometer for capturing acceleration in all the three planes for the 6 joints (left hip, right hip, left knee, right knee, left ankle, right ankle).This was later converted into joint angles using inverse kinematic solution [17].

C. Data Smoothing:

To remove unwanted noise from the data collected through accelerometers, we have cubic spline interpolation.

D. Dimension Reduction:

After data smoothing, we employed PCA (Principle Component Analysis) and IFA(Incremental Feature Analysis) to compute covariance between different features incrementally.

E. Feature selection:

After analyzing covariance between different set of features, we then proceed to select set of 6 features out of 10, which have highest covariance. This not only reduces the chance of over fitting, but also leads to improved generalization and classification accuracy as discussed in [15],[16].Table-1 below shows all the 10 kinematic features. F1-F6 represents the selected features.

Table1: List of 10 kinematic features

Feature Category	Feature Name
F1	Left Ankle
F2	Right Ankle
F3	Left Knee
F4	Right Knee
F5	Left Hip
F6	Right Hip
F7	CoP*
F8	CoG*
F9	GRF*
F10	Velocity

III. ALGORITHMS:

A. Extreme Learning Machines(ELM)

ELM was proposed by Huang [18], [19]. It is different from MLP and SVM. MLP used to consider many hidden layer neuron and considered whole networks as a black box. ELM theories show that hidden neurons do not need tuning because its hidden nodes parameters (c_i ,a_i) are randomly assigned [20]. For N arbitrary distinct samples $(x_k, t_k) \in R^n \times R^m$, the single ELM classifier with N hidden nodes becomes a linear system as,

$$\sum_{i=1}^{N} \beta_i G(x_k;\ c_i;\ a_i) = t_k, k = 1, \ldots \ldots . N. \quad (1)$$

where $c_i \in Rn$ and $a_i \in R$ are the learning parameters of hidden nodes and randomly assigned weight β_i connecting the ith hidden node to the output node, xk are the training examples, t_k is the target output for k = 1, ..., N, and G (x_k ; c_i ; a_i) is the

output of the ith hidden node with respect to the input xk. The output weights can be described in matrix form as

$$\beta = \left[\beta_1^T \ldots . \beta_N^T\right]_{mxN}^T \quad (2)$$

Equation (1) can be rewritten as:

$$H\beta = T \quad (3)$$

Where H (c_1…….c_N, a1 ……, x_1, ….., x_N) =

$$\begin{bmatrix} G(x_1; c_1, a_1) & \ldots & G(x_1; c_N, a_N) \\ \vdots & \ddots & \vdots \\ G(x_N; c_1, a_1) & \ldots & G(x_N; c_N, a_N) \end{bmatrix}_{NxN} \quad (4)$$

$$T = \left[t_1^T \ldots t_N^T\right]_{mxN}^T (5)$$

The weights β can be obtained using equation 6 by taking the least-square solution:

$$\hat{\beta} = H^{\dagger}T \quad (6)$$

Here, H† is the Moore-Penrose generalized inverse [21] and H is the output matrix of the hidden layer. The ELM used to implement the multi-class classification using a network architecture which has output nodes equal to the number of pattern in the classes. The network output can be written as O = $(O_1; O_2; \ldots O_n)^T$. For each training example say x, the target output tg is coded into n bits: $(tg_1; \text{--- } tg_n)^T$ .For a pattern of class k, only the target output t_k is "1" and the rest is "-1"

B. MLP, SVM and KNN Classifiers:

Multi-layer perceptron (MLP) is the simplest neural network architecture which is used most frequently. It has configuration flexibility, good representational capabilities and programmable algorithms. The Complexity of neural network depends of no. of hidden layer and no. of neurons. The best model of neural network is derived by trial and error. The major challenge of neural network is to decide the no. of hidden layer and no. of neurons, as there is no as such mechanism of optimal selection of neurons and hidden layers. Most MLP are based on gradient descent techniques that require iterative tuning which makes MLP extremely slow and the algorithm can also converge to local minima and provide sub optimal results.

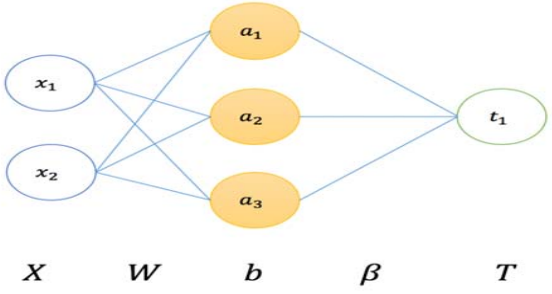


Fig.3. SLFNarchitecture of ELM

The objective of next classifiers SVM is to find the optimal decision boundary with maximum margin between classes in higher dimension feature space. It used to map the lower dimension data into higher dimension feature space using kernel function. The kernel function used to take linear time for this transformation. It converts the nonlinearly separable data into linearly separable data in high dimension feature space. Using SVM, an optimal separating hyper plane in the higher dimensional feature space can be computed by using kernel functions in the input space. In this study, we have used a Support Vector Machine classifier based on LibSVM. It uses a RBF(Radial Basis Function) kernel.

KNN used to match the similarity measure to classify the new class. It is simple algorithm. It is a simple non-parametric technique for classification. When the new testing data is fed, KNN computes the distance between the query data and training samples. Based on the defined threshold, number-k, k samples with least distances are selected and the class with more samples in the boundary is the result.

IV. Results and Discussions

Our experiments were conducted on a Windows 10 machine with a 2.50-GHz 7200U i5 CPU and 8.0-GB RAM. To find the best parameters for each of the classifiers, a grid search was run exhaustively with 10-fold cross validation to reduce variation associated with randomness of data shuffling and arrive at mean scores. KNN, SVM and MLP were implemented using scikit-learn python module. The best set of parameters was found to be as under. Table 2 represents the best set of parameters for MLP.

Table2: Set of Parameter for MLP

Hidden Layer Neurons	50
Activation function	'relu'
solver	'Lbfgs'
alpha	0.0001
momentum	0.9
iterations	1000

'relu' is an activation function such that,f(x)=max(0,x).'lbfgs' is an optimizer in the family of quasi-Newton methods.'alpha' is the L2 penalty parameter.'multiquadric' is an Radial basis activation function suchthat,f(x)=√1+(∈x)2. table 3,4 and 5 represents the best set of parameters for KNN, SVM, ELM.

Table 3:Set of Parameter for KNN

Neighbors	4
weights	'uniform'
metric	'manhattan'

Table 4:Set of Parameter for SVM

Parameter C	32768
kernel	'rbf'
gamma	0.001953125

Table 5: Set of Parameter for ELM

Hidden neurons	370
Activation function	'multiquadric'
Rbf-width	0.4

To evaluate results we computed confusion matrix, precision, recall, f1 score for each of the four classifiers and calculated overall classification accuracy. Overall accuracy is simply number of correct predictions divided by total number of predictions. F1 score is the harmonic mean of precision and recall. Additionally, we also recorded time taken for each classifier to train and predict labels on testing set. All performance measures were generated using scikit-learn python module. In order to eliminate high variation from results we have employed a 10-fold cross validation.

The data set consists of 5 classes namely,(Normal, Crouch1,Crouch2) obtained from OpenSim [22], Multiple sclerosis and Stroke. Data for classes Multiple sclerosis and Stroke were collected clinically capturing the same kinematic features as in the first three classes. Subjects didn't suffer from any other medical condition besides the said diseases. There are, a total of 945 samples across all 5 classes with their individual numbers given under 'support' column in the classification report tables. Crouch gait is a common movement abnormality among children with cerebral palsy, decreases walking efficiency due to the increased knee and hip flexion during the stance phase of gait [23][24].Crouch gait is of four types, out of which we have included type 1 and type 2 for classification. Table 6 and 7 represents the confusion matrix and classification report for ELM.

Table 6: ELM Confusion Matrix

	Normal	Crouch1	Crouch2	MS	Stroke
Normal	202	0	0	0	0
Crouch1	0	54	19	0	0
Crouch2	0	25	45	0	0
MS	0	0	1	299	0
Stroke	0	9	7	0	284

Table 7: Classification Report using ELM

	Precision	Recall	F1-score	Support
Normal	1.00	1.00	1.00	202
Crouch1	0.61	0.74	0.67	73
Crouch2	0.62	0.64	0.63	70
MS	1.00	1.00	1.00	300
Stroke	1.00	0.95	0.97	300

The Overall classification accuracy using ELM is 93.54%. Table 8 and 9 represents the confusion matrix and classification report for MLP.

Table 8: MLP Confusion matrix

	Normal	Crouch1	Crouch2	MS	Stroke
Normal	197	2	3	0	0
Crouch1	0	26	47	0	0
Crouch2	2	42	26	0	0
MS	0	0	0	300	0
Stroke	0	0	0	0	300

Table 9: Classification Report using MLP

	Precision	Recall	F1-score	Support
Normal	0.99	0.98	0.98	202
Crouch1	0.37	0.36	0.36	73
Crouch2	0.34	0.37	0.36	70
MS	1.00	1.00	1.00	300
Stroke	1.00	1.00	1.00	300

The Overall classification accuracy is 89.84% achieved using MLP. Table 10 and 11 represents the confusion matrix and classification report for KNN..

Table 10: KNN Confusion matrix

	Normal	Crouch1	Crouch2	MS	Stroke
Normal	191	5	6	0	0
Crouch1	20	20	33	0	0
Crouch2	19	50	1	0	0
MS	0	0	0	300	0
Stroke	0	0	0	0	300

Table 11: Classification Report using KNN

	Precision	Recall	F1-score	Support
Normal	0.83	0.95	0.88	202
Crouch1	0.27	0.27	0.27	73
Crouch2	0.03	0.01	0.02	70
MS	1.00	1.00	1.00	300
Stroke	1.00	1.00	1.00	300

The Overall classification accuracy=85.92% using KNN. Table 12 and 13 represents the confusion matrix and classification report for SVM.

Table 12: SVM Confusion matrix:

	Normal	Crouch1	Crouch2	MS	Stroke
Normal	202	0	0	0	0
Crouch1	0	53	20	0	0
Crouch2	0	15	55	0	0
MS	0	0	0	300	0
Stroke	0	0	0	0	300

Table13: Classification Report using SVM

	Precision	Recall	F1-score	Support
Normal	1.00	1.00	1.00	202
Crouch1	0.78	0.73	0.75	73
Crouch2	0.73	0.79	0.76	70
MS	1.00	1.00	1.00	300
Stroke	1.00	1.00	1.00	300

We have achieved the overall classification accuracy 96.29% using SVM.

Comparative analysis

We have used three figures as shown below to bring out a comparative analysis of the four classifiers. From figure-4 we can see that although numerically the poorest performing classifier-KNN, achieves a good overall accuracy of 85.92%, but it's classification is very poor in the case of Crouch 1 and Crouch2.MLP too, fails to have a decent f1 score in the same two classes, while SVM and ELM have evidently performed well across all 5 classes. The overall accuracy comparison is provided by figure-5.Interestingly, although KNN is the fastest of all four algorithms, it's also the least accurate, while ELM, despite having a far more complex Neural network of 370 neurons against MLP's 50, predicts labels in less than $1/10^{th}$ the time.

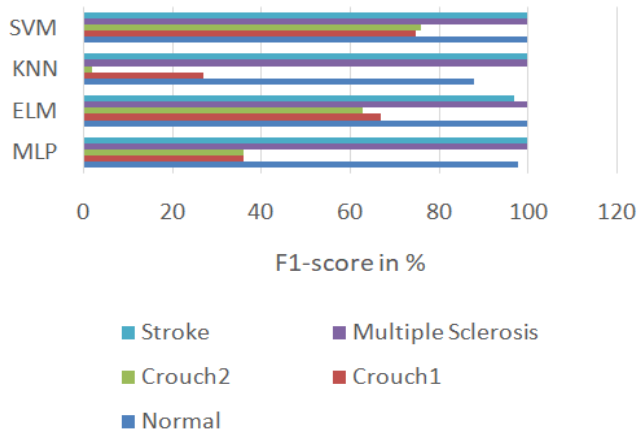


Fig.4. Comparison across classes

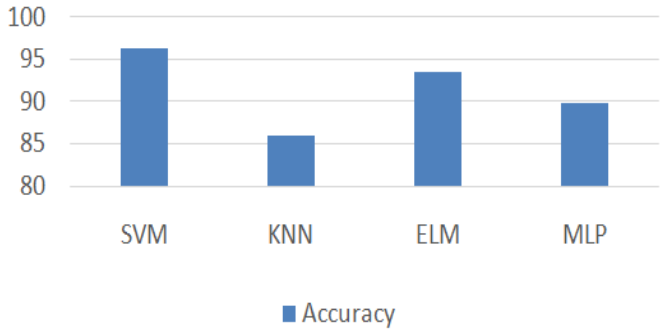


Fig.5. Different Classifier Classification Accuracy

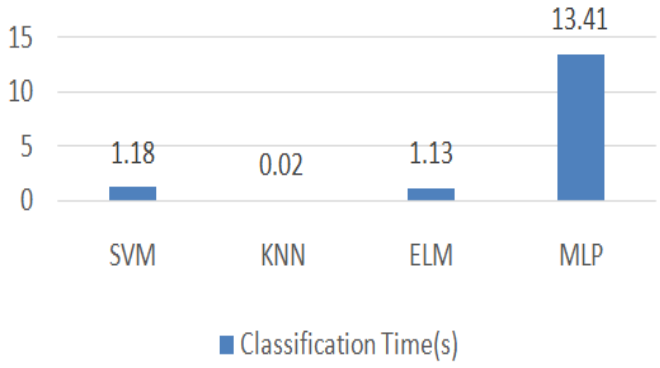


Fig.6. Different Classifier Classification time in seconds

VI. CONCLUSION

This paper presented the performance of ELM based classifiers with other machine learning classifiers named SVM, MLP and KNN for multi class gait classification. The paper has used kinematic features. The results show that ELM performs very good classification accuracy compare to other classifiers. The ELM has achieved the 93.54% overall classification accuracy. As the performance of machine learning algorithms dependent on the selected feature and parameter used, to improve the performance the paper has implemented increment feature analysis technique for kinematics features selection and various parameters is adjusted. In the future one could experiment with dynamic and kinematic features as this study was limited to kinematic gait characteristics alone.

References:

[1] Semwal, Vijay Bhaskar, et al. "Biped model based on human Gait pattern parameters for sagittal plane movement." IEEE,*International Conference onControl, Automation, Robotics and Embedded Systems (CARE)*, 2013.

[2] Kale, A.;Sundaresan, et.al. "Identification of humans using gait." Image Processing, IEEE Transactions on 13, no. 9 (2004): 1163-1173.

[3] Kiichi_Katoh,, Makoto Katoh.: Using FS in Tuning a Biped Locomotion Surrogate Robot for Walking Training. In: 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI).

[4] Nandi, Gora Chand, et al. "Modeling bipedal locomotion trajectories using hybrid automata." IEEE TENCON,2016.

[5] Mitsuru Yoneyama, Yosuke Kurihara, KajiroWatanabe,: Accelerometry-Based Gait Analysis and Its Application to Parkinson's Disease Assessment—Part 2: Ieee Transactions On Neural Systems And Rehabilitation Engineering, Vol. 21, No. 6, November 2013

[6] Venkataraman, S. T. "A simple legged locomotion gait model." Robotics and Autonomous Systems 22.1 (1997): 75-85.

[7] Semwal, Vijay Bhaskarand Gora Chand Nandi. "Biometric gait identification based on a multilayer perceptron." *Robotics and Autonomous Systems* 65 (2015): 65-75.

[8] Watelain, Eric, Franck Barbier, Paul Allard, André Thevenon, and Jean-Claude Angué. "Gait pattern classification of healthy elderly men based on biomechanical data." Archives of physical medicine and rehabilitation 81, no. 5 (2000): 579-586.

[9] Huang, Guang-Bin, Lei Chen, and Chee Kheong Siew. "Universal approximation using incremental constructive feedforward networks with random hidden nodes." IEEE Trans. Neural Networks 17.4 (2006): 879-892.

[10] Khusainov, R., A. Klimchik, and E. Magid. "Comparison of kinematic and dynamic leg trajectory optimization techniques for biped robot locomotion." Journal of Physics: Conference Series. Vol. 803. No. 1. IOP Publishing, 2017.

[11] Semwal, Vijay Bhaskar, et al. "Design of Vector Field for Different Subphases of Gait and Regeneration of Gait Pattern." IEEE Transactions on Automation Science and Engineering (2016).

[12] Khusainov, Ramil, et al. "Toward a human-like locomotion: modelling dynamically stable locomotion of an anthropomorphic robot in simulink environment." Informatics in Control, Automation and Robotics (ICINCO), 2015 12th International Conference on. Vol. 2. IEEE, 2015.

[13] Semwal, Vijay Bhaskar, and Gora Chand Nandi. "Generation of joint trajectories using hybrid automate-based model: a rocking block-based approach." IEEE Sensors Journal 16.14 (2016): 5805-5816.

[14] Huang, Guang-Bin, et al. "Extreme learning machine for regression and multiclass classification." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 42.2 (2012): 513-529.

[15] Semwal, Vijay Bhaskar, et al. "An optimized feature selection technique based on incremental feature analysis for bio-metric gait data classification." Multimedia Tools and Applications (2016): 1-19.

[16] Vijay Bhaskar Semwal and G.C. Nandi, Robust and more accurate feature and classification using deep neural network, Neural Computing and Application, Springer.

[17] Semwal, Vijay Bhaskar, and Gora Chand Nandi. "Toward developing a computational model for bipedal push recovery–a brief." *IEEE Sensors Journal* 15.4 (2015): 2021-2022.

[18] Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew. "Extreme learning machine: a new learning scheme of feedforward neural networks." Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on. Vol. 2. IEEE, 2004.

[19] Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew. "Extreme learning machine: theory and applications." Neurocomputing 70.1 (2006): 489-501

[20] Huang, Guang-Bin. "An insight into extreme learning machines: random neurons, random features and kernels." Cognitive Computation 6.3 (2014): 376-390.

[21] Rao, Calyampudi Radhakrishna, and Sujit Kumar Mitra. "Generalized inverse of matrices and its applications." (1971): 601-620.

[22] Delp, Scott L., et al. "OpenSim: open-source software to create and analyze dynamic simulations of movement." IEEE transactions on biomedical engineering 54.11 (2007): 1940-1950.

[23] Wren, Tishya AL, Susan Rethlefsen, and Robert M. Kay. "Prevalence of specific gait abnormalities in children with cerebral palsy: influence of cerebral palsy subtype, age, and previous surgery." Journal of Pediatric Orthopaedics 25.1 (2005): 79-83.

[24] Semwal, Vijay Bhaskar, Neha Gaud, and G. C. Nandi. "Human gait state prediction using cellular automata and classification using ELM." Machine Intelligence and Signal Analysis. Springer, Singapore, 2019. 135-145.

A Performance Improved Kretschmann Configuration Based Surface Plasmon Resonance (SPR) Sensor

Tauhidul Haque * and Hasan Khaled Rouf †
Department of Electrical & Electronic Engineering
University of Chittagong
Chittagong-4331, Bangladesh
E-mail : * tauhidtuhin670@gmail.com, † hasan.rouf@cu.ac.bd

***Abstract*— A surface plasmon resonance (SPR) sensor has been designed having improved sensor performance by taking the advantage of higher sensitivity of silver (Ag). Integration of a titanium (Ti) layer in the sensor configuration improves the adherence of metal on prism glass. Inclusions of an air gap also increase the sensor performance. Addition of a thin dielectric layer of indium phosphide (InP) enables surface plasmons (SPs) to spread along it and support a guided mode. Thereafter SPR sensor performances are measured with different performance parameters including sensitivity factor or sensitivity, full width at half maximum (FWHM) value, detection accuracy and quality factor. Numerical simulation has been carried out by using finite difference time domain (FDTD) method to study the reflectance curve and also metallic film optimization has been performed. In a comparative study between proposed SPR sensor and Kretschmann configuration, the proposed sensor provides better sensitivity of 70.65^0/RIU and higher quality factor of 365.9 RIU^{-1} which is 2.4 times higher than the Kretschmann configuration.**

Keywords—surface plasmon; SPR sensor; sensitivity; quality factor; Kretschmann configuration; FDTD.

I. INTRODUCTION

Surface plasmon resonance (SPR) is a charge density oscillation wave arising from the interaction of light with free electrons at metal - dielectric interface. At specific resonance wavelength or resonance angle of incident light, the photons energy is transferred to surface plasmon (SP) when the wave vector of incident light and SP are matched. The resulting SPs are excited and produce a surface plasmon wave (SPW) which is considered as electromagnetic surface waves. This wave propagates along the interface and decay exponentially with distance normal to the interface [1]. Generally, SPW propagation constant is always higher than the incident optical wave propagation constant. Therefore, to excite SP it is necessary to be enhanced photon momentum of the incident light beam to the level of SP momentum. Conventionally, prism coupling is used for optical excitation of SPs by attenuated total reflection (ATR) method which was demonstrated by Kretschmann. In this case a light is totally reflected at prism - metal interface by generating an evanescent wave when it intends to pass through the prism. This evanescent wave penetrates through the metallic layer and propagates with the propagation constant of $k_x = n_p(2\pi/\lambda)sin\theta$, where, n_p is the refractive index of the prism and λ and θ are wavelength and angle of incident light, respectively. The propagation constant k_x can be adjusted in order to match with that of the SPs by controlling the angle of incidence (θ) [2].

As gold (Au) and silver (Ag) provides low losses and strong resonance responses so they are widely used in plasmonic applications than other metals. Ag provides higher sensitivity by producing narrower SPR curve and sharp SPR resonance dip while Au produces wider response curve which limits the sensitivity. On the other hand, Ag is more susceptible to oxidation while Au is impermeable to oxidation and chemically stable. Therefore, in terms of sensitivity Ag may be a promising candidate if it is possible to insert a chemically inert layer on its surface [3]. A dielectric layer like indium phosphide (InP) can be used on Ag surface to protect it from undesired oxidation. InP has a large real part of dielectric constant which can helps to absorb light energy in the metal. It provides larger fraction of the evanescent field in the analyte region which enables SPs to spread along it and support a guided mode. To improve the SPR sensor performance of the existing Kretschmann configuration several works have been conducted by using Ag metal [2, 3, 4, 5]. A. Verma et al. proposed graphene based SPR biosensor with air gap providing 2.35 times higher sensitivity than the configuration without air gap [2]. In [3] T. Haque et al. proposed a graphene based SPR sensor which provides higher sensitivity and it is 1.4 times better than [4] where in [4] S. H. Choi et al. proposed a graphene on silver film based SPR sensor which provides more than 3 times higher sensitivity and higher imaging detection for graphene monolayer. A. K. Sharma et al. proposed a graphene on silver based SPR sensor by using ZBLAN fluoride glass substrate. They observed the impact of graphene chemical potential (μ) with layer numbers on sensor performance and found higher sensitivity for $0.7 < \mu < 0.1$ eV and graphene monolayer [5].

In this work we have carried out a comparative study between proposed and basic Kretchmann configuration shown in Fig. 1. In the proposed SPR sensor a titanium (Ti) layer is used for better adherence on prism glass. Silver as metal layer is used to get higher sensitivity. The thickness of Ag is optimized at 47 nm. To get a larger fraction of the evanescent field in the analyte region we have also used a thin dielectric layer of InP which enables SPs to spread along it and support a guided mode. Inclusion of air gap layer in the proposed sensor increases the sensor performance. Thereafter SPR

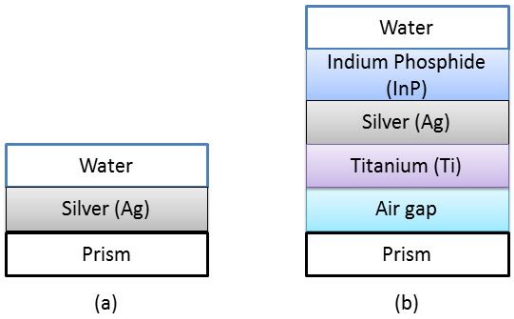


Fig. 1. Schematic diagram of (a) Kretschmann configuration and (b) proposed SPR sensor.

sensor performances are measured with different performance parameters including sensitivity factor or sensitivity, full width at half maximum (FWHM) value, detection accuracy and quality factor. Simulation has been carried out by using finite-difference time-domain (FDTD) method to study the reflectance curve and also metallic film optimization has been performed. Proposed sensor structure provides better sensitivity (70.65^0/RIU against 52.82^0/RIU) and quality factor of the proposed sensor configuration is 2.4 times higher than that of the Kretchmann configuration.

II. Theoretical Modeling

For the calculation of reflectivity of the SPR sensor, matrix method for an N-layer model is used. Considering the SPR sensor configuration of Fig. 1 as an N-layered multilayer system the tangential fields at the first boundary are related to those at the final boundary by

$$\begin{bmatrix} U_1 \\ V_1 \end{bmatrix} = M_2 M_3 M \dots M_{N-1} \begin{bmatrix} U_{N-1} \\ V_{N-1} \end{bmatrix} = M \begin{bmatrix} U_{N-1} \\ V_{N-1} \end{bmatrix} \quad (1)$$

where, U_1 and V_1 represent the tangential components of electric and magnetic fields, respectively, at the boundary of the first layer and U_{N-1} and V_{N-1} are the corresponding fields for the boundary at N^{th} layer [6].

For p-wave at boundary k

$$\begin{aligned} U_k &= H_y^T + H_y^R \\ V_k &= E_y^T + E_y^R \end{aligned} \quad (2)$$

and

$$M_k = \begin{bmatrix} cos\beta_k & -i \sin(\beta_k/q_k) \\ -i\, q_k\, sin\beta_k & cos\beta_k \end{bmatrix} \quad (3)$$

where

$$\begin{aligned} q_k &= (\mu_k/\varepsilon_k)^{1/2} cos\theta_k \\ \beta_k &= \frac{2\pi}{\lambda} n_k\, cos\theta_k (d_k) \end{aligned}$$

From Eq. (3), after some straightforward steps, the reflection coefficient for p-wave incident light is given by

$$r_p = \frac{(M_{11} + M_{12} q_N) q_1 - (M_{21} + M_{22} q_N)}{(M_{11} + M_{12} q_N) q_1 + (M_{21} + M_{22} q_N)} \quad (4)$$

From Eq. (4), M_{ij} is the characteristics matrix of the combined structure of the sensor and for p- polarized light it is given by

$$M_{ij} = \left(\prod_{k=2}^{N-1} M_k \right)_{ij}, i,j = 1,2,\dots \quad (5)$$

Finally, the reflectivity R of a given multilayer structure is given by

$$R_p = |r_p|^2 \quad (6)$$

Several parameters are used to analyze the performance of a SPR sensor. Sensitivity (S) of the SPR sensor describes the change in resonance angle ($\delta\theta_{res}$) per unit change in the refractive index of the sensing medium (δn_s) [2]:

$$S = \frac{\delta\theta_{res}}{\delta n_s} (deg/RIU) \quad (7)$$

The detection accuracy (DA) describes the ratio of shift in resonance angle ($\delta\theta_{res}$) to the full width at half maximum (FWHM) of the reflectance curve [7]:

$$DA = \delta\theta_{res}/\text{FWHM} \quad (8)$$

The quality factor (QF) is defined as the ratio of sensitivity (S) to the FWHM of the reflectance curve [7]:

$$QF = S/\text{FWHM} \quad (9)$$

III. Proposed SPR Sensor Design

In Fig. 1(a) we have shown the schematic of basic Kretschmann configuration. The schematic diagram of the proposed SPR sensor is shown in Fig. 1(b). This SPR sensor is comprised of six layers where it includes a glass prism, air gap layer, titanium (Ti) layer, silver (Ag) layer, indium phosphide (InP) layer and water as sensing layer. To convey this sensor design we have used an operating wavelength of 750 nm which is optimum. The design consideration of proposed SPR sensor with thicknesses of all layers and refractive indices are described below briefly. A prism is used as a first layer that indicates a substrate for the design with high RI of 1.723.

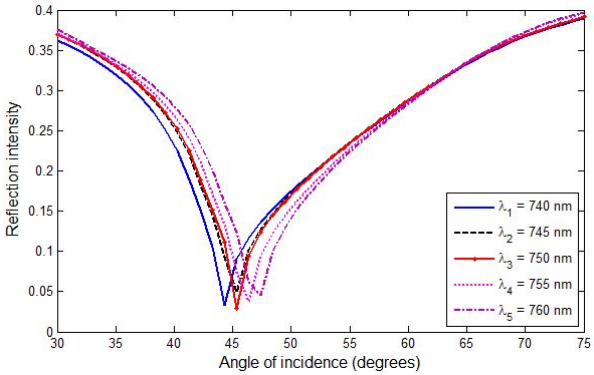


Fig. 2. Reflectance curves variation as a function of incidence angle at different operating wavelengths (λ).

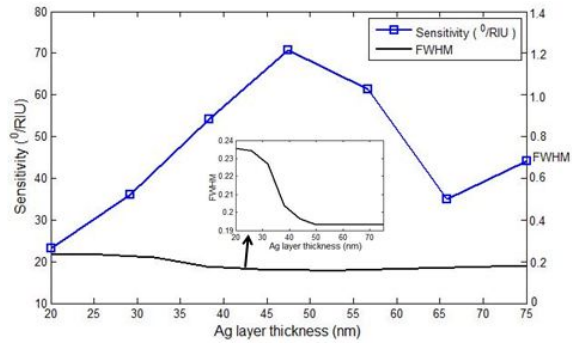


Fig. 3. Variation of sensitivity and FWHM as a function of Ag layer thickness.

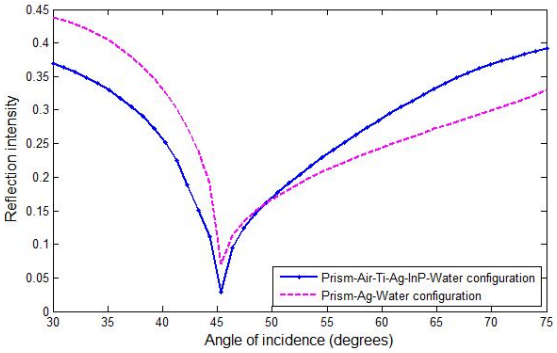


Fig. 4. Reflectance curves as a function of incidence angle for two different configurations.

Inclusion of an air gap layer between prism and metal increases the sensitivity of the sensor where RI of air is 1 and thickness is optimized. As silver has low adherence with prism glass so it must be needed to insert a titanium (Ti) layer between them to enhance the adherence of silver layer on prism glass having refractive index of 2.9838 + i4.0042 [8] and thickness is 15 nm. A silver (Ag) metallic layer is deposited on the titanium layer to take the advantage of higher sensitivity. In this case the refractive index of Ag is 0.1468 + i4.8604 [9] where optimized thickness is 47 nm. Thereafter a thin dielectric layer of indium phosphide (InP) having RI of 3.4719 + i0.23111 [10] and thickness 20 nm is used to support a guided mode. The last and sixth layer is water having RI of 1.33 is used as sensing medium.

IV. Result and Discussion

Numerical simulation using finite-difference time-domain (FDTD) method is used to study the proposed SPR sensor and optimize the thickness of the metallic layer. There exists an operating wavelength at which maximum energy transfer is occurred from incident light to SPs. So, from the very beginning it is necessary to find out this specific resonance wavelength. In Fig. 2 we have shown reflection intensity as a function of incidence angle for different wavelengths. At specific incidence angle, SP is resonantly coupled with incidence angle which is known as resonance angle (θ_{res}). When perfect coupling between SP and incident light wave is occurred then SPR condition takes place at θ_{res} with a fixed

TABLE I. Performance of Sensor at Different Wavelengths.

Wavelength λ	θ_{res} (deg)	R_r	Sensitivity (^{0}/RIU)	Detection accuracy	Quality factor (RIU^{-1})
740	44.32	0.0325	69.85	0.71	227.8
750	45.34	0.0288	70.65	0.85	365.9
760	47.39	0.0454	67.77	0.62	207.5

TABLE II. Performance Comparison of Proposed SPR Sensor Against Kretschmann Configuration.

Configuration	θ_{res} (deg)	R_r	Sensitivity (^{0}/RIU)	Detection accuracy	Quality factor (RIU^{-1})
Proposed SPR sensor (P-Air-Ti-Ag -InP-W)	45.34	0.0288	70.65	0.85	365.9
P-Ag-W	45.34	0.0708	52.82	0.55	155.1

P=Prism, Ti= Titanium, Ag = Silver, InP=Indium Phosphide, W=water

operating wavelength and a dip is appeared due to the absorption of light energy in the metal which is decreased to nearly zero. Therefore from Fig. 2 it is observed that at the operating wavelength of 750 nm and for resonance angle of 45.34^0 the dip of the reflectivity curve is lowest. So 750 nm is an operating wavelength for this proposed design. When the wavelength is increased from 740 nm to 760 nm it is observed that the dip of the reflectivity curve is going up with broadening which decreases the sensitivity. In Table I performance parameters at different wavelengths are summarized. Full width at half maximum (FWHM) describes the width and sharpness of the reflectivity curve. A sensor provides better sensitivity at lowest FWHM value. So among the entire wavelengths 750 nm provides lowest FWHM value (0.19308497^0) which indicates its better sensitivity than other wavelengths of 740 nm, 745 nm, 755 nm and 760 nm. Therefore, at the operating wavelength of 750 nm the proposed sensor provides a better performance (S, DA, QF).

The maximum energy transfer from incident light to SP also depends on an optimum metallic layer thickness. So it is important to find out optimum metal layer thickness. From Fig. 3 we can clearly observe that at 47 nm of Ag layer thickness the sensitivity of sensor is highest (70.65^0/RIU) and FWHM is almost constant after 47 nm. As Ag dielectric constant imaginary part is increased with the increment of Ag layer thickness, so we can observe from Fig. 3 that for 20 nm to 47 nm thickness of Ag layer the FWHM value is drastically decreased due to absorbing incident light photon energy by the metal. But after 47 nm of Ag layer thickness the metal do not absorb significant amount of energy and the curve shows almost constant. Hence, we get lowest value of FWHM (0.19308497^0) at 47 nm of Ag layer thickness with highest sensitivity.

A comparative study between proposed sensor (Fig. 1(b)) and Kretchmann configuration (Fig. 1(a)) has been carried out where proposed sensor shows improved performance. In Fig. 4 we have observed that the lowest dip of reflectivity curve is appeared for proposed sensor structure at an operating wavelength of 750 nm. The proposed SPR sensor provides better sensitivity of 70.65^0/RIU against 52.82^0/RIU for Kretschmann configuration. The proposed one also provides higher quality factor of 365.9 RIU^{-1} (against 155.1 RIU^{-1}) which is 2.4 times higher than Kretschmann configuration. All of the performance parameters are summarized in Table II. By evaluating Table II and Fig. 4 it is undoubtedly clear that the proposed SPR sensor enhance the performance.

V. Conclusion

A SPR sensor has been designed having improved sensitivity, quality factor and other performances. Inclusion of an air gap enhanced the sensor performance. Using Ti as adherence layer and InP as dielectric layer we have achieved an enhanced performance. InP protect the Ag metal from oxidation. It also enables SPs to spread along it and support a guided mode. A comparison between proposed SPR sensor and Kretschmann configuration has been carried out where proposed one provides better sensitivity 70.65^0/RIU than another and proposed SPR sensor provides 2.4 times higher quality factor.

Acknowledgment

This work was supported by ICT Fellowship (FY 2017-2018, 2nd round) provided by the Information and Communication Technology (ICT) Division, Government of the People's Republic of Bangladesh.

References

[1] Ibrahim Abdulhalim, Mohammad Zourob and Akhlesh Lakhtakia, "Surface plasmon resonance for biosensing : A mini-review," Electromagnetics, vol. 28, pp. 214-242, 2008.

[2] Alka Verma , Arun Prakash and Rajeev Tripathi, "Sensitivity enhancement of surface plasmon resonance biosensor using graphene and air gap," Optics Communications, vol. 357, pp. 106-112, 2015.

[3] Tauhidul Haque and Hasan Khaled Rouf, "Graphene based surface plasmon resonance (SPR) sensors : An approach to enhance the performance," Proceedings of the International Conference on Innovations in Science, Engineering and Technology (ICISET), 4 pages, October 27-28, 2018.

[4] Seung Ho Choi,Young L. Kim, and Kyung Min Byun, "Graphene-on-silver substrates for sensitive surface plasmon resonance imaging biosensors," Optics Express, vol. 19, pp. 458-466, 2011.

[5] Anuj K. Sharma and Anumol Dominic, "Influence of chemical potential on graphene-based SPR sensor's performance," IEEE Photonics Technology Letters, vol. 30, no. 1, pp. 95-98, January 1, 2018.

[6] Anuj K. Sharma and B. D. Gupta, "On the performance of different bimetallic combinations in surface plasmon resonance based fiber optic sensors," Journal of Applied Physics, vol. 101, pp. 093111, 2007.

[7] J. B. Maurya, Y. K. Prajapati, V. Singh, J. P. Saini, and R. Tripathi, "Performance of graphene-MoS_2 based surface plasmon resonance sensor using silicon layer," Applied Physics A, vol. 121, no. 2, pp. 525–533, 2015.

[8] P. B. Johnson and R. W. Christy, "Optical constants of transition metals: Ti, V, Cr, Mn, Fe, Co, Ni, and Pd," Physical Review B, vol. 9, pp. 5056-5070, 1974.

[9] P. B. Johnson and R. W. Christy, "Optical Constants of the Noble Metals," Physical Review B, vol. 6, p. 4370-4379,1972.

[10] D. E. Aspnes and A. A. Studna, "Dielectric functions and optical parameters of Si, Ge, GaP, GaAs, GaSb, InP, InAs, and InSb from 1.5 to 6.0 eV," Physical Review B, vol. 27, pp. 985-1009, 1983.

Leakage Detection in Water Pipeline Using Micro-controller

Muhammad Arman Uddin
Deptertment of Computer Science and Engineering
BRAC University
66 mohakhali, Dhaka, Bangladesh
muhammad.arman.uddin@g.bracu.ac.bd

Mohammad Mohibul hossain
Deptertment of Computer Science and Engineering
BRAC University
66 mohakhali, Dhaka, Bangladesh
mohamed.mohibul.hossain@g.bracu.ac.bd

Akil Ahmed
Deptertment of Computer Science and Engineering
BRAC University
66 mohakhali, Dhaka, Bangladesh
akil.ahmed@g.bracu.ac.bd

Hasibul Hasan Sabuj
Deptertment of Computer Science and Engineering
BRAC University
66 mohakhali, Dhaka, Bangladesh
hasibul.hasan.sabuj@g.bracu.ac.bd

Samin Yasar Seaum
Deptertment of Computer Science and Engineering
BRAC University
66 mohakhali, Dhaka, Bangladesh
samin.yasar.seaum@g.bracu.ac.bd

***Abstract*—This methodology uses information from the water flow sensors which identifies leakage in water pipeline progressively. Testing the leakage in pipeline is done on every sensor hub and data is then sent to the Arduino UNO micro-controller for further procedure. The output of this proposed strategy can be accomplished by contrasting different water flow reading with the sensors. The threshold value is set by taking examination of the sensor reading. If the water flow reading is not exactly limit esteem, at that point leakage is available in the pipeline is appeared on the yield show. Consistent flow is set for no leakage condition.**

Keywords— Water Flow Sensor, Arduino UNO, Pipeline Leakage, Flow Monitor.

I. INTRODUCTION

Water is a significant element for each country. Water leakages are being one of the serious issues which each country is confronting [1]. Financially, leaks are costing where in addition to wasting water; it devours significant capacity to treat water or to exchange it starting with one point then onto the next [8]. In underground pipeline leakage identification is a troublesome assignment. Now a days so many countries are suffering in terms of financial loss, fitness, energy and surroundings due to pipe leakages. In a study of smart water systems, it is seen that the leakage rate percentages in Bangladesh is the highest among many countries [1]. According to Dhaka Water and Sewer Authority (WASA), they faced system loss (non-revenue water) 22% in 2015 [2]. Also, Dhaka WASA shows that water supply demand is 2250 MLD, where the supply is 2420 MLD but every citizen of this city does not get safe water from Dhaka WASA though there is no shortage of supply of safe water [2,9]. In Bangladesh along with water leakage, unauthorized connections also make the situation worst for Dhaka WASA [7]. As of late, however, leakage checking hypothesis and application has extraordinary advancement [10,11].

There are several methods in literature for detection of leakages in pipelines such as infrared, SVM Classification and automated meter reading [3-5]. The Infrared technique is utilized to recognize pipe leakages moderately in the nations which contain deserts using impact of the difference in the temperature of the ground because of the leakages. But the disadvantage of this method is the countries in which even least precipitation occurs, this method does not work properly. However, the automated meter reading method measures meter read using a vehicle equipped with a communication technology like GSM-GPRS. But the constraint of the method are slow data rate and high complexity. In the contrary, In SVM Classification method, detection of pipe leakage and location is predicted from a dataset of CSIR- Central Electronics Engineering Research Institute (CEERI). Without this particular dataset it won't give exact output for detecting leakage and location. Nevertheless, there are few methods where wireless communication system is introduced [12-14] which is a future plan of this project.

To recognize the constant water leakage this project is proposed which utilizes two water flow sensors, when water flows through the rotor, rotor rolls. Its speed changes with various rate of flow. This project utilizes water flow sensor as it is effectively accessible in the market, it gives extraordinary productivity and savvy. Sensor detected information is put away in Arduino UNO micro-controller in which pressing is finished. This information is exchanged to the PC or head office through wire. In background of PC different strategies were utilized for examination of sensor information, for example, Arduino IDE.

The rest of the paper is organized as follows. Section 2 describes the theoretical study and methodology that we are going through. Section 3 describes the proposed model with a block diagram, circuit components with detailed description. Additionally, hardware implementation of the proposed model is demonstrated in this section. Moreover, Section 4 shows the output data and analyzes it. Finally, section 5 concludes the paper.

II. METHODOLOGY AND THEORY DISCUSSION

The data collected from water flow sensor are taken into micro-controller. Precise water flow estimation is a crucial step both in the terms of subjective and financial perspectives. Water flow sensor have proven excellent machine for estimating water flow, and now it is anything but difficult to construct a water monitoring system utilizing the G 1/2" Water Flow sensor. This sensor sits in accordance with the water line and contains a pin wheel sensor to quantify how much water has traveled through it. There is an integrated Magnetic Hall-Effect sensor that outputs an electrical pulse with each revolution. Here we have decided flow rate by the change in velocity of water. Velocity relies upon the pressure that forces the through pipelines. As the pipe's cross-sectional area is known and stays steady, the normal speed means that

the stream rate. In such cases, the principle for deciding the water flow rate is

$$Q = V * A \quad (1)$$

Where, Q is flow rate/total flow of water through the pipe, V is average velocity of the flow and A is the cross-sectional area of the pipe [6].

Next, the data from two points are analyzed to determine the result. The water flow of these two points can be denoted as Q_1 and Q_2 where Q_1 is source flow and Q_2 is ending flow of the two points.

$$\Delta Q = |Q_1 - Q_2| \quad (2)$$

The data is taken from all the points of the systems-Central Supplier, Hubs and end-users. The expected average flow of water from Central Supplier to Hub and Hub to end users are different. When the flow is less than the expected average flow the system detects a leakage as ΔQ will be greater. And if Q_2 is close to zero then there might be a clog between the two points. A threshold will be determined for normal flow of water by the minimum and maximum value of ΔQ denoted as T_{max} and T_{min} respectively. The decisions that can be taken are

- $T_{max} > \Delta Q > T_{min,} \rightarrow$ Normal flow
- $\Delta Q > T_{max} \rightarrow$ Leakage
- $Q_2 \approx 0 \rightarrow$ Clogged pipe

A. Hardwear

1) Sensors utilized:

Water flow sensor comprises of a plastic valve body, a water motor, and a hall-effect sensor [13]. At the point when water moves through the motor, it rolls. Later on, Its speed changes with various rate of flow. The hall-effect sensor yields the relating pulse signal. It utilizes a basic turning wheel that pulses a hall effect sensor. Flow rate range is 1～60L/min. Operating temperature is ≤80℃. Water pressure tolerance ≤1.75 MPa (Max 2MPa). Storage temperature needs to be -25℃～+80℃.

2) Micro-controller

Arduino Uno is a micro-controller board dependent on the ATmega328P. It has 14 digital input/output pins (of which 6 can be utilized as PWM outputs), 6 analog input pins, a 16 MHz quartz crystal, a USB association, a power jack, an ICSP header and a reset button. It contains everything expected to help the micro-controller; just associate it to a PC with a USB link or power it with an AC-to-DC connector or battery to begin.

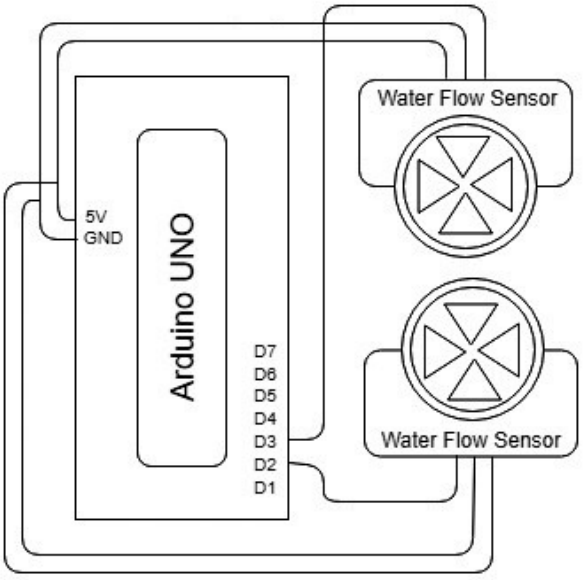


Fig. 1. Water flow sensor pin layout required for the system.

3) LCD display

The LCD display is used to observe the data which are useful for users for analysis. 16*2 LCD display is used here. The data pins 4-7 are connected to the Arduino UNO.

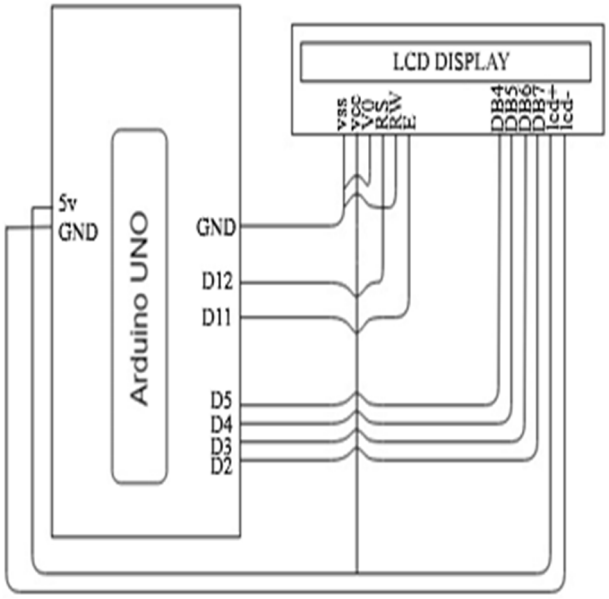


Fig. 2. LCD Pin layout of the system

B. *Flow chart*

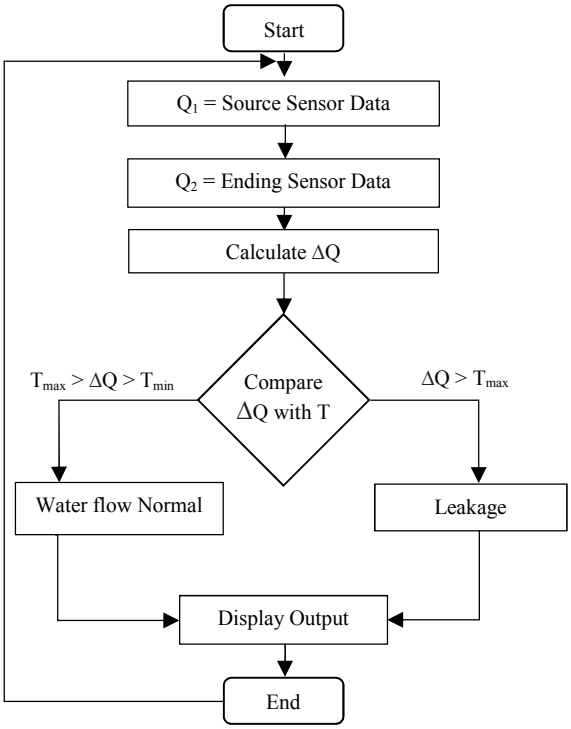


Fig. 3. Flowchart

The above Flow chart shows how leakage detection functions in the system. Leakage detection happens between every two nodes of WDS. To start with, flow sensor information from source between the two nodes is taken. Next, the ending sensor's information is taken. At that point, ΔQ is calculated and compare it with threshold T and see whether there is a leakage or the flow is normal. At last, the decision output is appeared in LCD display. The checking proceeds in a loop.

C. *System Design*

This system contains hardware part like water flow sensors, power supply, microcontroller and LCD display.

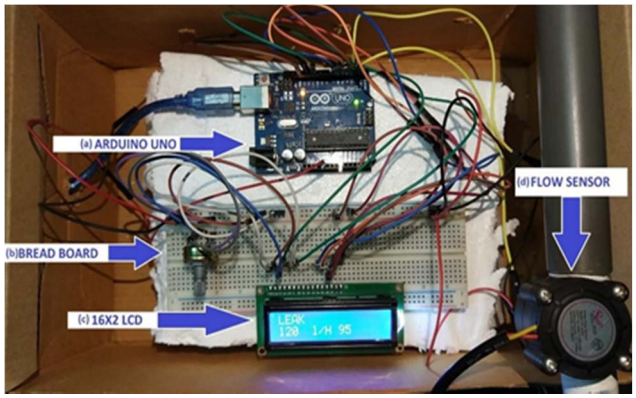


Fig. 4. The components of our project are shown in the figure. (a) is the Arduino Uno (b) is the bread board (c) is the 16x2 LCD (d) is the flow sensor.

III. PROPOSED MODEL

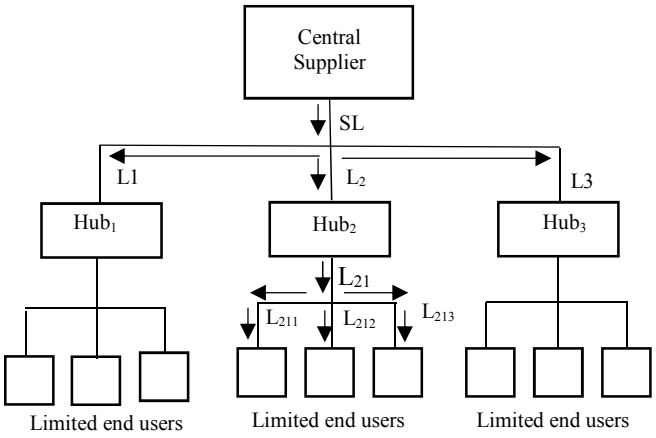


Fig. 5. Proposed Model

The whole system architecture follows the water distribution system (WDS). In earlier approaches, other researches were not able to detect the leakage in real time [8]. The WDS approach makes it possible to get the real time data and detects the exact location of the leakage within the shortest possible time. The central water supplier is divided into hubs which distributes water to the end users. Each hub distributes water to the limited number of end-users. The device is designed to be placed on each node shown in fig. 5. The system can detect leakage in between two nodes. When the node in line L_{213} gives an alert for water leakage the system determines the position of leakage in two ways. The leakage may be in line L_{21} or in line L_{213}. When detecting the problem, ΔQ calculated from Hub_2 to its end-users are same then the problem is in line L_{21}. However if only ΔQ calculated in line L_{213} determines leakage value and other two values are expected then the problem is in line L_{213}.

IV. RESULT ANALYSIS AND DISCUSSION

Fig. 4 shows the experimental test system implemented to monitor the test section of a pipe line using two water flow sensors. The system consists of two water flow sensors, bread board, Arduino UNO, some wires and LCD display to show the outputs. The test was done connecting a pipe with flow sensors and creating water flow through the pipe to simulate actual buried pipeline. A split was made in the pipe to reproduce leakage. For simulation, some pipes are used which have piped grade standard ANSI Sch40 and the size of the pipe was NPS1. The pipe's water velocity in the environment was 0.5~0.7 ft/s. The average water flow was found 1204 lit/hr on average. The proper threshold was found as 0 lit/hr~200 lit/hr. Therefore, if ΔQ greater than 200 lit/hr is found between two nodes then there is a leakage between the two nodes.

TABLE I. FLOW RATE IN DIFFERENT CONDITION

Scenarios	Q1 (L/H)	Q2 (L/H)	ΔQ (L/H)	T (L/H)	Output
1	1200	1100	100	0~200	Normal Flow
2	1200	800	400	0~200	Leakage
3	1200	0	-	0~200	Clogged

[a] L=Liter, H=Hour, T=Threshold

V. CONCLUSION

This paper exhibits water pipeline leakage detection model. Leakage is identified utilizing water flow sensor and correspondence is done remotely. By comparing different reading of sensors leakage is identified. Arduino UNO is used to show the consequence of sensors. This model gives ongoing leakage present in the pipeline. Utilizing this model water misfortune and water misfortune causes, water contamination issues that are effectively recuperated. However, there is not much cloud computing, IoT and wireless based method has been implemented yet to detect the leakage in pipeline. There is various opportunity of research in this field that could be focused.

REFERENCES

[1] G. Gupta, A., Mishra, S., Bokde, N., Kulat, K.: Need of smart water systems in India. Int.J. Appl. Eng. Res.11(4), 2216–2223 (2016)

[2] Dwasa.org.bd. (2014). *DHAKA WASA*. [online] Available at: https://dwasa.org.bd/annual-reports/ [Accessed 11 Apr. 2019].

[3] Shakmak, B., & Al-Habaibeh, A. (2015). Detection of water leakage in buried pipes using infrared technology; A comparative study of using high and low resolution infrared cameras for evaluating distant remote detection. 2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT).

[4] Porwal, S., Akbar, S. A., & Jain, S. C. (2017). Leakage detection and prediction of location in a smart water grid using SVM classification. 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS).

[5] Farah, E., & Shahrour, I. (2017). Smart water for leakage detection: Feedback about the use of automated meter reading technology. 2017 Sensors Networks Smart and Emerging Technologies (SENSET).

[6] Sood, R., Kaur, M., & Lenka, H. (2013). Design and development of automatic water flow meter. International Journal of Computer Science, Engineering and Applications, 3(3), 49.dfasd

[7] Alam, H. (2017, June 01). Water crisis all around Dhaka. *The Daily Star*. Retrieved April 11, 2019, from https://www.thedailystar.net/city/water-crisis-all-around-1413664

[8] H. L. Zhao, J. K. Wang, and X. Jiang, "A Real-time Monitoring System for the Pipeline Network of Coalmine," J. Phys. Conf. Ser., vol. 364, p. 012107, May 2012.

[9] Khan, T. A. (n.d.). *Page 6 of 13 Dhaka Water Supply and Sewerage Authority: Performance and Challenges*.

[10] Tang Xioujia. Location theory about unequal thermal pipe. Journal of BeiJin University,1997,33(5):575-579.

[11] Jin Shijiu, Tang Xioujia etc. Crude oil pipe leakage and location. Chinese journal of scientificinstrument,1997,18(4):343-348.

[12] Saraswati Vasantrao, K., & Rajbhoj, S. M. (2017). WSN Based Water Pipeline Leakage Detection. 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA).

[13] Y. Jin, and A. Eydgahi, "Monitoring of distributed pipeline system by wireless sensor networks,"ResesrchGate publication, In: proceeding of the 2008 IAJC-IJME international conference, 2008.

[14] I. Stoianov, L. Nachman, and S. Madden, "PIPENET: A wireless sensor network for pipeline monitoring," In :6th international symposium on information processing in sensor network (IPSN), Vol. 07, pp. 25-27, April 2007

Smart Gas Furnace

Md. Mominur Rahman, Abu Shufian, Farzana Akter, Md. Tarikul Islam Juel, and Riadul Islam
Department of Electrical and Electronic Engineering
American International University-Bangladesh, Dhaka-1229, Bangladesh
mominur18aiub@gmail.com, shufian2aiub@gmail.com, sumaaiub138@gmail.com, md.tarikulislamjuel@gmail.com, and riadul999@gmail.com

***Abstract*—In recent times, the use of furnace in industrial sectors has increased. Furnace deals with high temperature for melting and die casting purpose. The smart gas furnace basically made by G.I metal sheet (outer part of the furnace), solenoid valve, ceramic fiber, Plaster of Paris, black fire cement, crucible and gas as a fuel. Among different types of furnace, higher efficiency can be achieved from smart gas furnace. A smart gas furnace is the minimum expensive melting process introduced for melting any types of metal or alloy compared with other furnace. The objectives of this project work are to control the flow of gas for maintaining the required temperature inside the furnace and melt the material. The goal of this project work is to build a gas fired crucible furnaces and develop a controlling system which provide more benefit compare with other furnace. In this project Arduino is used to maintain the required temperature inside the furnace by switching the solenoid valve which controls the flow of gas. Whenever the temperature reaches the desired temperature the gas flow will automatically stopped by solenoid valve resulting in minimum gas loss. Which will reduce fuel cost. Thermocouple is used as a sensor. It continuously sends data (temperature) to Arduino, and by comparing the both data (required temperature and available temperature inside the furnace) the Arduino switches the solenoid valve.**

Keywords—Crucible, G.I metal sheet, solenoid valve, ceramic fiber, Plaster of Paris, Black fire cement, Arduino Uno.

I. INTRODUCTION

Material melting plays a very significant role in different types of melting industry for variety of purposes such as making steel and heat treating of materials to change their molecular structure to give them a perfect shape and this process is performed by the furnace. The furnace is a melting machine which produces heat by burning different types of fuels or electricity or other means. The name of furnace is originated from the Latin word fornix, which implies oven [1]. Material reusing is particularly critical for industries and this difficult work is performed by the furnace. The use of furnace in industrial sectors has been increased in the past few years. There are different types of furnaces are available for melting materials in industrial scale such as blast furnace, open hearth furnace, electric arc furnace, etc. In 1970s electrical furnace started replacing furnace due to energy crisis but the popularity of the gas furnace is still high. And it is one of the most effective method of melting materials in remote areas. Most of the furnace use gaseous fuel like as natural gas, manufactured gas, liquefied petroleum (Lp) as the energy input. Smart gas furnace can reduce the wasting of gas to save the natural resource.

II. LITERATURE REVIEW

The idea of melting metal by using furnace is centuries old. In early years people mainly used wood to generate heat in furnace. 1776 is the year from where the roots of crucible can be traced, when a steel mill named Naylor and Sanderson steel Mill was built in Sheffield, England to produce fine quality tools steel using the crucible method. Sanderson Brothers Steel Company built a new gas fired crucible melting furnace in 1883. Up to 30 graphite's furnace was held by this high temperature furnace, called crucible. Each charged with 60 to 100 pounds of iron plus selected alloy additions [3].

A. State of the art technology

Furnace is a melting machine which deals with high temperatures more than 400 degrees Celsius for melting materials. In industrial furnace heat generally generated by burning fuel with the help of air or by electrical energy. Furnace designs varies with its function, heating duty, types of fuel and method of introducing combustion air [4].

B. Furnace categories:

Different kinds of furnace have been using in industries from many years. Furnace can be classified into four categories based on its design and efficiency.

1) Natural draft: This type of furnace consisted of cast-iron. It is very simple in design and carries few controls. For operating this system operation wood, coke, coal, trash, paper, natural gas, fuel, oil can be used [4].

2) Forced air: Forced air is another types of furnace, atmospheric burner style with a cast-iron. Big natural draft type furnace was replaced by it. But stills its operations can be found in large brick chimneys [4].

3) Forced draft: Forced draft third categories of furnace with a steel heat exchange and multi-speed blower. It is much more compact than the previous styles. This kind of furnace is called mid-efficiency furnace [4].

C. Combustion controls:

There are many types of combustion controllers is used for controlling the flow of gas. Among them main three types are discussed below.

1) Single stage or ON/OFF control: It can operates only one stage operation; it is either on or off. So, it is called the simplest controller. It operates with noise and always running highest speed and pumping out the hottest air with highest velocity [4].

2) Two stage or HIGH/LOW/OFF controls: It is slightly more complex system compare with single stage. It has two firing rate, one slower firing rate and another is full firing as needed. But most of the time it operates at slower speed. Operating system is quieter than one stage and moves the air at velocity [4].

In this project two stage or HIGH/LOW/OFF controller were used.

III. METHODOLOGY

The figure in below explain the basic structure and methodology of smart gas furnace.

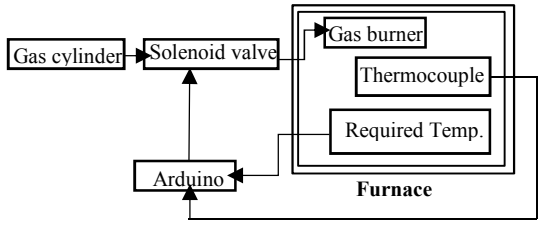


Fig. 1: Block diagram of a gas-fired crucible furnace

This project is done based on three important parts,

- Sensor
- Mechanical parts
- Controlling parts

A. Sensor: In this project thermocouple is used as a sensing device. Details about thermocouple device is given below,

1. Thermocouple: Thermocouple is a sensing device which is used to measure temperature. It has two different metal wires which are connected together to create a junction [5]. In this project thermocouple will sense the amount of heat inside the furnace and send the data to the Arduino.

B. Mechanical parts: For building the mechanical part G.I metal sheet, solenoid valve, gas burner torch has been used. Details about mechanical parts are given below,

1. Outer body frame: At First, the outer frame of a furnace body was made. The outer frame is cylindrical in shape which is made by GI (Galvanized iron) metal sheets. This metal sheet is one kind of sheet where the thickness of steel is 5mm.This metal sheet is very high temperature resistive. The total height of the cylindrical shape is 16.5 inch and only body height is 15 inch and total side insulation is 2 inch. The furnace will be open or close from the upper side of the cylinder where there will be coverage at the top of the furnace. There is one thermocouple is used in the chamber vertically for measuring temperature. The thermocouple sends data continuously in an Arduino. There is also a gas burner which is placed at the bottom of the furnace body which will rise the temperature at a very high level by burning the gas into the chamber.



Fig. 2: Outer body of the furnace.

2. Insulation inside the furnace body:
The main and most impotent part of the furnace is internal structure of the furnace body. There are some insulating materials like as ceramic fiber blanket, plaster of Paris which are adjusted at the inner side of the outer frame where the thickness of the ceramic fiber 1.5 inch and the thickness of the plaster of Paris is 05 inch. The ceramic fiber can protect the outer part of the furnace up to 1200-1250 degree Celsius from the effect of high temperature operation and the plaster of paris can protect up to 1200 degree Celsius [6]. Double layer ceramic fiber was used in this system and connected one with other using black fire cement which tolerates 1200 degree Celsius. Plaster of Paris is a soft mixture which tolerates 1200 degree Celsius was used for joining ceramic fiber with furnace body to build a hard surface. The thickness of Plaster of Paris is 0.5 inch. Now lower insulation of the crucible furnace body is 4 inch, where ceramic fiber Lear is 3 inch and Plaster of Paris Lear is 1 inch.



Fig. 3: insulation inside the furnace

3. Solenoid valve: Solenoid valve is known as an electromechanical control valve which has electric coil with ferromagnetic core in its center. This device is used to control the liquid or gas flow. A 2/2-way solenoid valve has inlet and outlet ports and also have two position (open or closed) [6]. In this project solenoid valve controls the flow of gas inside the furnace. Arduino controls the solenoid valve for maintaining the required temperature inside the furnace.
4. Gas burner torch: Gas burner torch is a heating device through which gas will be brunt and heat will produce inside the furnace.

C. Controlling parts : In this project Arduino is used as a controlling device. Details about Arduino is given below,

1. Arduino Uno: Arduino is an open-source platform which is used for building electronics project [7]. In Arduino, open-source means that its can modify and produce different Arduino boards which provide different factor and functionality. In this project Arduino is used as a controlling device. It takes data of furnace inside temperature from sensing device (thermocouple) and compare it with the required temperature. And maintain the required temperature

inside the furnace by switching off/on the solenoid valve.

D. *Measurements and Equations:*

The internal diameter = 7 inch

And, external diameter = (internal diameter + thickness of the refractory)

= (7 inch +4 inch) = 11 inch

Internal height of the furnace = 10 inch
External height of the furnace = (Internal height + metal sheet thickness + Ceramic fiber thickness + Plaster of Paris thickness)

= (10 inch + 2 inch + 3 inch + 0.5 inch)
=16.5 inch

To maintain the required temperature inside the furnace, the following equations were used.

- Heat produce to reach melting point temperature to melting point of the desire material

  $Q1=ms\Delta t$

  Here,
  m= Mass of melting product
  s = Specific heat
  Δt= Temperature difference
- Heat produce for melting material at melting point

  $Q2=mlf$
- Heat produce to increase 1℃ of the desired material,

  $Qint=ms*1$
- Final heat produced, Q= Q1 + Q2
- Required temperature,

  $T=Q/Qint$

IV. HARDWARE AND SOFTWARE INTERFACE

The main idea of this research is to automate the gas furnace operations. For this reason, this smart automation process is done by using a microcontroller Arduino which will control the operation of a solenoid valve.

A. *Overview of the system*

This complete automation system is done by following the flowchart below. This flowchart also displays the interconnection between the hardware and software section of the project.

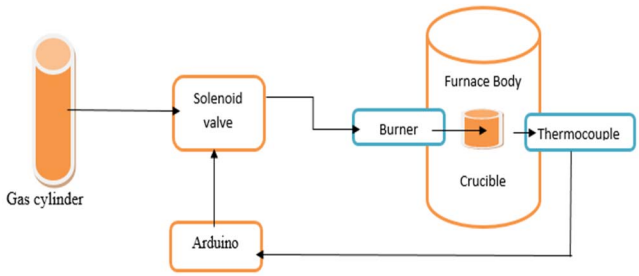


Fig. 4: Flowchart of interfacing the smart gas furnace system.

B. *Simulation and results:*

The schematic diagram of our project is as follows:

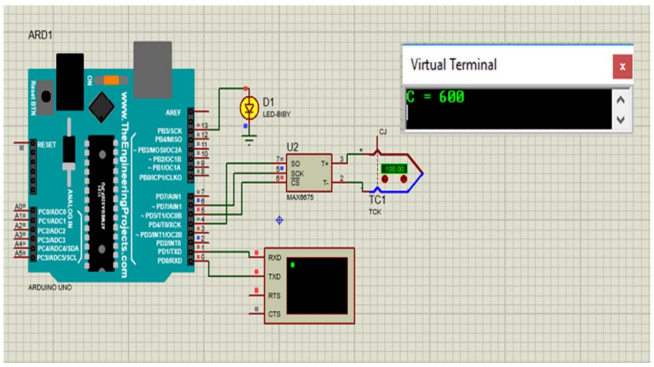


Fig. 5: Schematic diagram of Smart Gas Furnace

This figure represents the simulated circuit of controlling portion of the title project. Here virtual terminal is used for showing the temperature. It is connected with the serial pins of the Arduino and k type thermocouple is connecter with the digital pins of the Arduino. Because of unavailability of solenoid valve a LED light has been used. Whenever the temperature reaches the required temperature the led will turn on. Turning on the LED is the indication of switching off the solenoid valves furnace can reduce the wasting of gas to save the natural resource.

C. *Full hardware impletentation process:*

This whole procedure is done to build up the hardware portion of the project

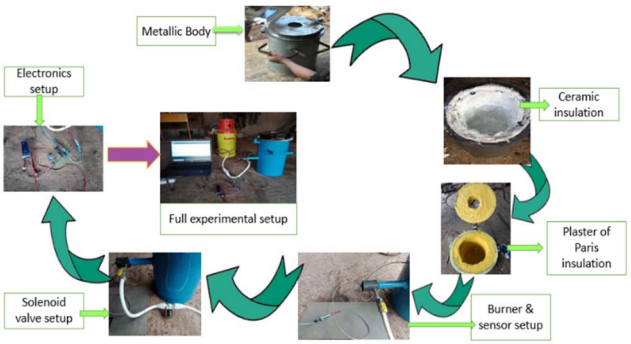


Fig. 6: Diagram of smart furnace building procedure

The main and most impotent part of the furnace is internal structure of the furnace body. There are some insulating materials like as ceramic fiber blanket, plaster of Paris which are adjusted at the inner side of the outer frame where the thickness of the ceramic fiber 1.5 inch and the thickness of the plaster of Paris is 05 inch. The ceramic fiber can protect the outer part of the furnace up to 1200-1250 degree Celsius from the effect of high temperature operation and the plaster of Paris can protect up to 1200 degree Celsius. A double layer ceramic fiber was used in this system and connected one with other using black fire cement which tolerates 1200 degree Celsius. The lower insulation of the crucible furnace body is 4 inch, where ceramic fiber layer is 3 inch and Plaster of Paris layer is 1 inch.

D. *Final combination of hardware and software*

Full experimental setup are shown below:

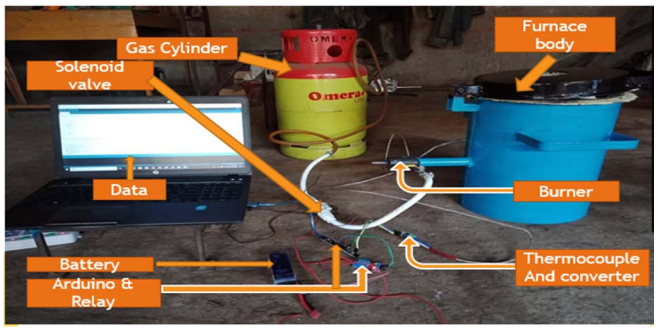


Fig. 7: Full experimental setup of the project.

Here, Arduino is connected with solenoid valve and thermocouple so that it can control the inside temperature of furnace body. Arduino take data from thermocouple (temperature sensor) and maintain the required temperature inside the furnace body by switching the solenoid valve. When the inside temperature reaches to the required temperature the Arduino switch off the solenoid valve and whenever the temperature remain below the required temperature the solenoid valve is open. This is how a smart gas furnace is built which can melt any type of metal within its maximum temperature range. Actually a metal will be kept inside the crucible. According to the metal and furnace heat capacity a required temperature will be set and after that gas will flow up to required temperature and the metal will be melted where melting metal keep any type of mold and also making any types of shape.

V. Result Analysis

Table 1: Valve operation according to required temperature (considering required temperatures 600°C)

Serial number	Temperature (°C)	Time(sec)
1	565.00	64.18
2	573.50	65.22
3	582.50	66.25
4	592.50	67.29
5	601.75	68.32
6	607.25	69.36
7	605.25	70.39
8	601.75	71.42
9	597.50	72.46
10	603.25	73.49
11	610.00	74.53
12	605.25	75.56
13	601.25	76.59
14	592.25	77.62

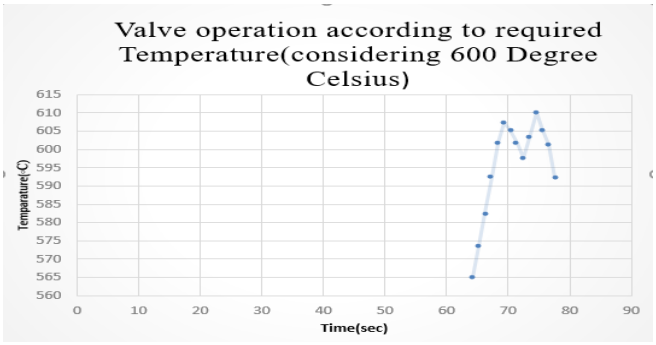


Fig. 8: Graphical representation of valve operation

From the graph it is clearly seen that when the temperature is below required temperature (600°C) then the valve opened and temperature inside the furnace body start increasing rapidly But when the temperature has reached the required temperature the Arduino switched off the valve. But due to enclosed surface temperature goes beyond the required level. As the valve remain closed the temperature inside the furnace started decreasing again. Whenever its goes below the required level the Arduino will switch on the solenoid valve and similar procedure will be followed again.



Fig. 9: Material (aluminium) melted in Smart Furnace

The above figure shows the clear image of melting procedure of desired material in Smart Gas Furnace. For melting aluminium (91gm), 1000°C is kept fixed inside the furnace.

VI. Energy Consumption Calculation

Some electronics devices like Arduino, solenoid valve, thermocouple and rely are used in this furnace designing project.

Table 2: Energy consumption of the proposed model

Electronics Devices	Ratings	Power consumption (W)
Arduino	5V, 13.92mA	0.348
Solenoid Valve	12V, 320mA	3.84
Thermocouple	5V, 30mA	0.15
Relay		0.45
	Total:	4.778W

A solar panel (12V, 1A) charge a battery for 6 hours can supply 72W. By using this solar panel furnace can be

operated for, (72/4.778) = 15.07 hours. A battery of (12V, 1A) can be used for storing the electricity from solar panel. So this proposed model consumes very low electricity. For this reason, it is a very effective for melting industries located in remote areas.

VII. Novelty Of This Project

- This proposed model will be very effective for jewelry shops and melting industries which are situated in the remote areas. Because in many country it is not possible to supply high voltage current in every corner of the country, so smart gas furnace will be very effective for that areas.
- Big melting industry can also keep this model as a backup. During the electricity failure and maintenance period they can continue their production by using smart gas furnace.
- This proposed model needs very small amount of power to operate its operations and this little amount of current can also be provided from solar energy, so this proposed model will play an effective role in remote areas.
- For insulation, double layer of ceramic fiber, plaster of Paris and black fire cement has used which made the system highly heat resistive, so high amount of heat can be produced inside the furnace which will make the melting process much faster.
- Easily transportable due to small in size, so bringing it to any remote areas will not be a big challenge.

VIII. Limitations Of This Project

In this research work some limitations are found such as, the weight capacity of the crucible is very low, high amount of heat loss due to using graphite crucible which takes huge amount of heat to heat up itself before transferring the heat to the material inside it, cannot maintain the required temperature precisely inside the furnace.

IX. Future Enhancement

Although the objective was fulfilled, but there are some update that can be done. Here, the weight capacity of crucible need to increase. As there is an issue of high amount of heat loss so it need to be reduced to increase the efficiency of the furnace. For getting higher demand in market the required temperature need to maintain precisely inside the furnace.

X. Conclusion

This system is used to melt any type of metal within its heating capacity. In our project material selection were carefully done so as to enable it operate the system it was designed for. It was manufactured following to the design specification. It produced high temperature to melt the metal. This is electronics based project. Arduino is the main controlling part in our project and solenoid valve which taken command from the Arduino to on/off the gas flow. As a result, we successfully control the gas flow and taken rising temperature data with time duration.

References

[1] How furnace is made - making, history, used, components, structure, industry, Raw Materials, Design", Madehow.com, 2019. [Online]. Available: http://www.madehow.com/Volume-7/Furnace.html. [Accessed: 05- Feb- 2019].

[2] "Different Types of Industrial Furnaces | Furnace | Heat Treating", Scribd, 2019. [Online]. Available: https://www.scribd.com/document/124299262/Different-Types-of-Industrial-Furnaces. [Accessed: 05- Feb- 2019].

[3] American Made Steel Manufacturers- Crucible Industries : New West KnifeWorks", Newwestknifeworks.com, 2019. [Online]. Available: http://www.newwestknifeworks.com/content/information/american-made-steel-manufacturers-crucible-industries-. [Accessed: 05- Feb- 2019].

[4] "Single Stage, Two Stage, and Modulating Furnaces", Information Library, 2019. [Online]. Available: https://hvacdirect.com/info/single-stage-vs-two-stage/. [Accessed: 05- Feb- 2019].

[5] "Type K Thermocouple - Type K Thermocouples - K Type Thermocouples - K Type Thermocouple", Thermocoupleinfo.com, 2019. [Online]. Available: https://www.thermocoupleinfo.com/type-k-thermocouple.htm. [Accessed: 05- Feb- 2019].

[6] "Insulation Materials: Fiberfrax Ceramic Fiber", Thermaxx Jackets, 2019. [Online]. Available: https://www.thermaxxjackets.com/products/insulation-materials/ceramic/. [Accessed: 05- Feb- 2019].

[7] "Tameson", Tameson.com, 2019. [Online]. Available: https://tameson.com/solenoid-valve-types.html. [Accessed: 05- Feb- 2019].

[8] "What is an Arduino? - learn.sparkfun.com", Learn.sparkfun.com, 2019. [Online]. Available: https://learn.sparkfun.com/tutorials/what-is-an-arduino/all. [Accessed: 05- Feb- 2019].

DOORMOR: A Fully Functional Prototype of a Search and Rescue Robot

Fahim Faisal
Dept. of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh.
embeddedfahim@gmail.com

Farhana Sharmin Tithi
Dept. of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh.
farhanasharmin3tithi@gmail.com

***Abstract*—In this paper, we have discussed about DOORMOR - a standalone, fully functional search and rescue robot prototype which can come to aid for humans in many cases. With the help of Wi-Fi technology, this robot can be controlled within up to hundred meters. It can search for alive human bodies underneath a pile of debris, clear out debris and other obstacles, stream live video of the surrounding environment wirelessly and assist people in search and rescue operations. In this paper, we have tried to give a brief idea about DOORMOR, it's architecture, facts and features, used components and techniques used to build up the whole robot.**

Keywords—Bot, Live Video Streaming, Control Application, Living Body Search, Debris Clearance, Microcontroller, PIR Sensor

I. INTRODUCTION

Nowadays, robotics and mechatronics is considered as one of the most prominent topics in research. In robotics, architecture, design, manufacturing is important as well as implementation of software to give robots better artificial intelligence to make them appropriate for the tasks which they should handle. Mechatronics is the consolidation of mechanical, computer and electronic engineering in the design of integrated systems with high capabilities that accommodate numerous intelligent traits [1]. It's a derived branch of engineering which is getting more and more popular in this era. From defense to public sectors, robots are being used everywhere. Today, due to unplanned civilization and industrialization, environment pollution has become earth's greatest enemy and a threat to our existence. We have never been more prone to natural calamities as we are today. Earthquakes, cyclones, tsunamis are becoming common incidents day by day. But, often due to lack of proper disaster management and implementation of technology, people die in large numbers. In this particular area, robots can help save many valuable lives. Often, humans face a situation while rescuing that, they can't reach the affected people due to lack of proper path and there is always a risk of the rescuers getting hurt themselves. Robots on the other hand don't have these limitations, they can slip through even the smallest and slender paths, and they can operate without anyone getting hurt.

II. RELATED WORKS

The concept of search and rescue robots is still relatively new. It emerged roughly at the beginning of the 1980s. Nevertheless, actual developments in this field began after the 1990s. Although robots are not widely used in actual search and rescue operations at the moment, the development in this field seems quite promising. In Eastern Japan Earthquake, various robot technologies were applied and achieved productive results for nuclear plant accident response, port inspection, victim search, structural inspection, etc. [2].

Trupti B. Bhondve, Prof. R. Satyanarayan, et al proposed a rescue robot, which can search for alive humans in disaster events in both auto and manual mode. They used a Radio Frequency (RF) modem to maintain connection to the bot from the control device [3]. Albert Ko and Henry Y. K. Lau proposed an autonomous bot circumscribing a wireless network of sensors specially designed for rescue operations. This robotic system can navigate autonomously into rubbles and search for living human body heat using its thermal array sensor [4]. Zachary Cody Hazelwood and Dr. Saleh M. Sbenaty proposed a robot that can move autonomously and avoid obstacles while moving. It also has a user input mode in which the robot takes instructions from the user [5]. Siobh´an Grayson proposed the deployment of multiple robots in disaster events, each having different functionalities. The robots each would have their own assigned tasks [6]. Luis Pineda, Takeshi Takahashi, et al introduced a smart, sophisticated algorithm which is capable of automatically detecting distressed people in the event of a disaster and ease the work of the rescuers [7]. Jingchao Zhao, Junyao Gao, et al introduced a robot which is capable of sensing various kinds of gases in coal mines. Live video of the surrounding environment is captured through a webcam and then broadcasted over a local wireless area network [8]. Jaeeun Shim and Ronald C. Arkin introduced deception of robots in SAR contexts and presented a novel computational approach for an autonomous rescue robot's deceptive action selection mechanism [9]. Johannes Maurer, Gerald Steinbauer, et al tried to document the state of the art in disaster robotics in their paper, in order to provide an overview of the technology already available or currently under development [10]. Illah R. Nourbakhsh, Katia Sycara, et al, in their paper, discussed about the automatic determination of disaster victims using sophisticated algorithms [11]. J. Ju´arez-Guerrero, I. Pissokas, et al, in their paper, discussed about the implementation of smart algorithms for determination of disaster victims in unknown environments without human assistance [12].

III. DESIGN AND BOT MECHANICS

At the very core of the bot design, four wheels have been used for moving the bot. On the front side, two wheels have been used, and on the back, another two, thus two wheels have been used per side. The chassis was designed in a way that it can carry all the electronic equipment like batteries, circuits, drivers as well as the arm and other components. Full dimension of the bot is $9.5'' \times 7'' \times 7''$. The dimension of the chassis of the bot is $9.5'' \times 7'' \times 2''$. The body is built of stainless steel to ensure maximum distortion tolerance. When we try to turn the bot in the left direction, the right wheels go

forward and left wheels go backward thus creating an uneven pressure in the bot's body and the bot turns left. When we try to turn the bot in the right direction, the left wheels go forward and right wheels go backward, and the bot turns right. In this bot, four gear motors [13] have been used so that moving on an uneven surface becomes easy.

The bot's arm has three parts: base, elbow and wrist. The base and the elbow of the arm can rotate up to 180 degrees. The bot uses two servo motors [14] to perform this task. The servo motors [14] are placed respectively in the base and elbow, which gives a sophisticated hand movement. Using two servo motors [14] gives us two Degrees of Freedom (DOF). The claw [15] used in DOORMOR is two fingered, which also holds a servo [14] for materializing the grabbing action. The bot is designed this way so that it can overcome any hazardous situation.

IV. Bot Electronics

The main controller of the bot is the microcontroller *Atmel ATMega328P*. In other words, it's the brain of DOORMOR. An *Arduino Nano* [16] has been used for utilizing this microcontroller and programming it easily. A microprocessor was needed for rendering live video data as microcontrollers don't possess such powerful CPUs. So, a *Raspberry Pi 3* [17] has been used which has a quad-core *Cortex A57* CPU, more than capable of rendering HD video data.

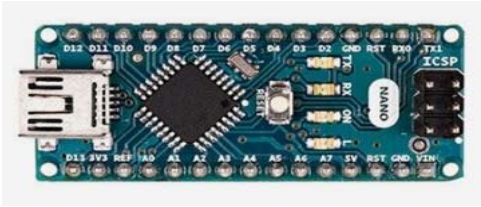


Fig. 1: Arduino Nano.

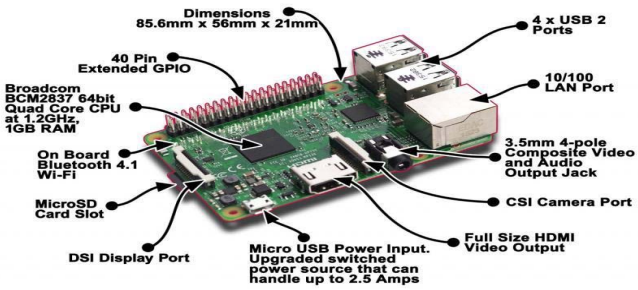


Fig. 2: Raspberry Pi 3 Model B.

Two 5V Lithium-Ion rechargeable power supplies [18] have been used to power the bot. One for powering all the power-hungry components such as, gear motors [13], servo motors [14] and the motor driver [19]. The other one is for powering the *Raspberry Pi* [17] and the *Arduino Nano* [16]. The *Raspberry Pi* [17] and the gear motors [13] require 5V to perform properly and the optimum voltage for the servo motors is also 5V (Input Range: 4.8V – 7.2V). The *Arduino Nano* [16] receives 5V directly from the *Raspberry Pi* [17]. No separate power supply was used for the Passive Infrared Ray (PIR) sensor [20] as it receives power directly from the *Arduino Nano* [16] pins. There is also a 5MP camera module [21] which is powered through the Camera Serial Interface (CSI) port on the *Raspberry Pi* [17]. The L293D dual-channel motor driver [19] has been given 5V as input because the gear motors [13] that we have used, perform best under 5V. Four gear motors [13] have been used to power up four wheels and all of them are given power through the motor driver [19]. Two motors are connected in parallel with reverse polarity on each channel of the motor driver [19]. As both of the power supplies [18] have onboard voltage regulators, additional voltage regulation wasn't required. In average, the total power consumption of the whole bot is around 5A in medium terrain. The power supplies that have been used, can handle the loads flawlessly. The *Arduino Nano* [16] was used to control everything and collect sensor data because it's easier to connect different modules to it as it has 5V output whereas the *Raspberry Pi* [17] has an output of 3.3V, which some modules and sensors don't accept. Also, generating Pulse Width Modulation (PWM) signals with various duty cycles for servo control with the *Pi* is a bit inefficient.



Fig. 3: PIR Sensor Module.

V. Communication System

FCC standards and regulations were followed regarding the communication between the bot and the control device. A 2.4 GHz Wi-Fi router [22] has been used to transfer video data from the bot to the network and control signals from the control device to the bot. The onboard Wi-Fi module of the *Raspberry Pi* [17] which connects to the Wi-Fi router [22], was used to transmit video information to the router and receive control signals from the control device. IEEE 802.11 b/g/n wireless distribution was utilized in all the network components. An android app (DOORMOR Controller) was used to maintain the connection from the control device to the bot. The control commands are collected from the app and then sent to the *Arduino Nano* [16] by the android device through the Pi's onboard Wi-Fi module. The Passive Infrared Ray (PIR) sensor is hooked up to the *Arduino Nano* [16] and feeds data directly to the microcontroller. The microcontroller then redirects the data to the control device through the Wi-Fi module and Wi-Fi router [22]. The redirected data is then translated in the app.

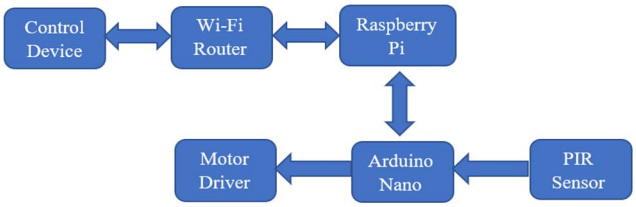


Fig. 4: Diagram of the Communication System.

VI. Live Video Streaming

The 5 megapixels camera module [21] has been used to remotely see what is going on at the desired location. It is

connected to the mainboard [17] via the Camera Serial Interface (CSI) port. It was set up in a way so that we can avoid data loss and other consequences. It can capture full HD (1920x1080) videos at 30 frames per second. The video data is transmitted over a Wireless Local Area Network (WLAN) created by the Wi-Fi router [22] and can be viewed from any smart device connected to the network through a web browser. The video encoding has been implemented by the FFMPEG [23] open source video library. Video data is temporarily stored in a webserver inside the Pi and then it is broadcasted over the WLAN using User Datagram Protocol (UDP) for fast data transmission.



Fig. 5: Video Data Transmission Path.



Fig. 6: Raspberry Pi Camera Module (5 MP).

VII. Bot Software and Associated Hardware

The microcontroller (*Atmel ATMega328P*) was programmed in the *Arduino* language (a subset of C++) for ease of programming. The *Raspberry Pi* [17] was programmed in Python as it is the official programming language for the device. The *Raspberry Pi* [17] works as a server and the android device acts as a client. PHP and HTML languages were used to create a webpage, which is hosted on the *Raspberry Pi* [17] server. LIGHTTPD [24] was used in this regard. It is a light, flexible, fast and less resource-consuming webserver which is very easy to use. Single-board computers like the *Raspberry Pi* [17] are usually low on hardware resources and as such, we need lightweight software to compensate for the hardware limitations. Additionally, CRTMPSERVER [25] was set up for temporarily storing the video data coming from the camera module [21].



Fig. 7: DOORMOR Controller (Control App).

The bot software or control app establishes a central control system to control the bot from a remote distance with the help of an android device. It offers a user-friendly Graphical User Interface (GUI), which includes all the necessary tools and information display areas.

A. Bot Movement Controller

The motor driver [19] has six control pins. Four are for controlling the direction in which the motors [13] should rotate, e.g. clockwise or anti-clockwise. They are also used to start or stop the motors [13]. So, these pins are pulled HIGH and LOW with a combination from the L293D's truth table [26] to execute different actions. Six *Arduino Nano* [16] digital pins are connected to the motor driver [19]. These pins generate high and low logic to materialize the desired bot movements. The *Raspberry Pi* [17] has been programmed to act as a server here. Server-client communication is done through Transmission Control Protocol (TCP) for secure communication. The client software named "DOORMOR Controller" was written in the Java programming language. It has a GUI which consists of FORWARD, BACKWARD, LEFT, RIGHT and STOP buttons as well as a status box to display the state of the bot. These buttons have predefined values which are sent in packets. With the help of sockets [27], the *Raspberry Pi* [17] receives these values and sends appropriate signals to the microcontroller. A query is sent to the bot per second, in order to see if the connection to the bot is maintained or not. If no reply is received, all the *Arduino Nano* [16] pins turn low in order to protect the bot from malfunctioning.

Input A	Input B	Motor State
High	Low	Turns clockwise
Low	High	Turns anti-clockwise
High	High	Braking occurs
Low	Low	Braking occurs

Fig. 8: Single Channel Truth Table of L293D.

B. Arm Movement Controller

A portion of the 'DOORMOR Controller' app has the necessary tools to control the arm movement. There are three sliders for controlling the three servos [14]. The sliders have values from 0° to 180°. These sliders need to be moved in a systematic way in order to obtain desired arm movements. The status area of the control app shows in which degree the servo [14] is at, whenever the concerned slider is highlighted.

On the programming side, these sliders have different predefined underlying values at different points, e.g. the first slider has a value range of 1000-1180, the second slider has a value range of 2000-2180 and the third one has a value range of 3000-3180. These values are sent as integers which are stored inside arrays in the microcontroller. The first digit represents the servo number and the rest three digits represent the degree at which the servos need to rotate and stay. After receiving signals from the control device, the microcontroller generates appropriate Pulse Width Modulation (PWM) signals for driving the servos [14]. PWM signal generation is handled by the Servo library of the *Arduino* language.

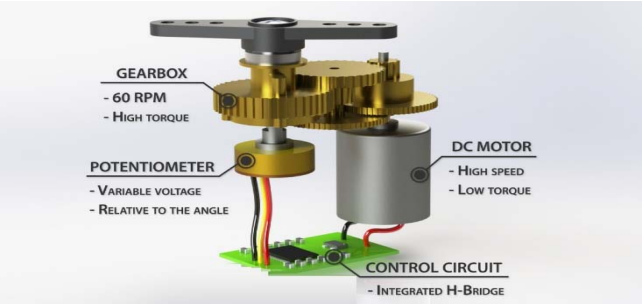


Fig. 9: Servo Mechanism.

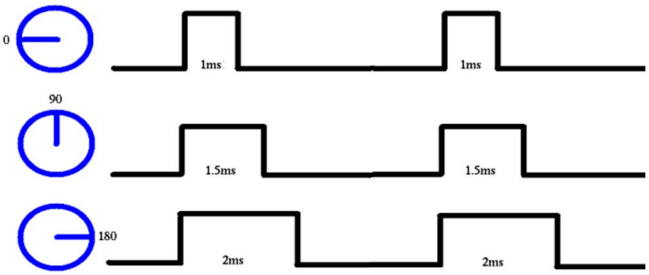


Fig. 10: Different PWM Signals for Servo Control.

```
if(realservo >= 1000 && realservo < 1180)
{
  int servo1 = realservo; //storing values for servo 1..
  servo1 = map(servo1, 1000, 1180, 0, 180); //mapping down to suitable degree values..
  myservo1.write(servo1); //sending positional information to servo 1..
  Serial.println("SERVO 1 ON"); //prints confirmation message on the serial monitor..
  delay(10); //waits for 10 millisecond for the servo to adjust it's position..
}

if(realservo >= 2000 && realservo < 2180)
{
  int servo2 = realservo; //storing values for servo 2..
  servo2 = map(servo2, 2000, 2180, 0, 180); //mapping down to suitable degree values..
  myservo2.write(servo2); //sending positional information to servo 2..
  Serial.println("SERVO 2 ON"); //prints confirmation message on the serial monitor..
  delay(10); //waits for 10 millisecond for the servo to adjust it's position..
}

if(realservo >= 3000 && realservo < 3180)
{
  int servo3 = realservo; //storing values for servo 3..
  servo3 = map(servo3, 3000, 3180, 0, 180); //mapping down to suitable degree values..
  myservo3.write(servo3); //sending positional information to servo 3..
  Serial.println("SERVO 3 ON"); //prints confirmation message on the serial monitor..
  delay(10); //waits for 10 millisecond for the servo to adjust it's position..
}
```

Fig. 11: Partial Source Code.

C. PIR Sensor Status

Another portion of the 'DOORMOR Controller' app displays the status of the PIR sensor [20]. The PIR sensor [20] can detect Passive Infrared Rays (PIR) emitted from living beings, such as, alive humans and other animals [28]. A dead body doesn't emit such rays. Inside the PIR sensor, all the thermal energy is at first focused on to the Pyroelectric sensor through the focusing lens in front of the sensor. If PIR is available, the voltage across the sensor changes. This change in voltage is too delicate to be detected by the microcontroller, so we need an amplifier circuit which will amplify the change. The output of the amplifier is connected to a voltage comparator, which compares voltages from the amplifier with a reference voltage and generates HIGH/LOW logic [28]. So, the module sends a HIGH logic to the microcontroller whenever it detects a living being within 1 meter. The microcontroller then sends a signal to the control device through the Wi-Fi network. This signal is translated in the app. The app shows "Human Detected" in the sensor status area, whenever a living body is detected. If there's no living body, the app shows "No Human Detected". The mechanism of the PIR sensor [20] is shown in Fig. 12.

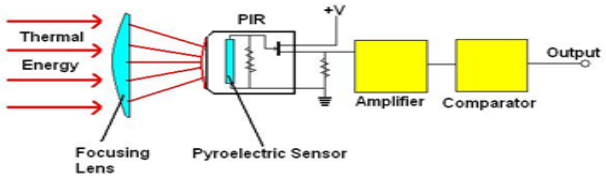


Fig. 12: Passive Infrared Sensor Mechanism.

VIII. DESCRIPTION OF THE TASKS

The bot can perform several tasks. They are respectively:

1. Searching for alive humans under pile of debris.
2. Picking up and clearing debris using the arm.
3. Capturing live video of surrounding environment and streaming it through a wireless network.

Task 1 is shown in Fig. 13, where the bot has detected the existence of a living body.
Task 2 is shown in Fig. 14, where the bot is picking up demo debris.
Task 3 is shown in Fig. 15, where the bot is streaming live video to a Windows PC connected to the bot's wireless network.



Fig. 13: Live Human Detection.

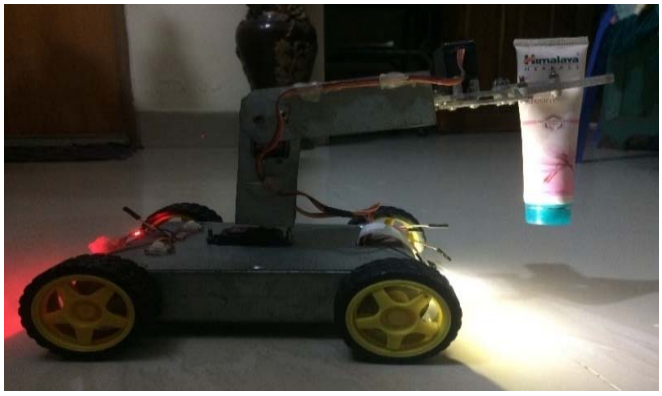


Fig. 14: Demo Debris Cleaning.



Fig. 15: Live Video Streaming.

IX. Observations and Experiments

Several experiments were conducted to ensure proper functionality of the bot. This helps to accurately determine the future complications that may occur in actual field. The experiments can be split into several steps:

A. Live Video Streaming Test: Success scenarios of the bot during the live video streaming test are shown in Table I.

B. Traversing Test: Success scenarios of the bot's traversing test are also shown in Table I.

C. Arm Test: Success scenarios of the bot's arm test are shown in Table II.

D. Sensor Test: Success scenarios of the bot's PIR sensor test are shown in Table III.

TABLE I. Bot Test Results

Test Run	Terrain	Vision	Control	Status
1	Plane	Operator's eye	With bot	Success
2	Medium	Bot camera	Blind from operator	Success
3	Plane	Operator's eye	With bot	Success
4	Medium	Bot camera	Blind from operator	Success

TABLE II. Arm Test Results

Test Run	Type of Object	Weight	Action Performed	Status
1	Demo Debris	50 gm	Picked up & moved	Success
2	Demo Obstacle	100 gm	Picked up & moved	Success

TABLE III. Sensor Test Results

Test Run	Sample	Output	Status
1	Living body	'Human Detected' shown in the control application	Success
2	Living body	'Human Detected' shown in the control application	Success
3	Living body	'No Human Detected' shown in the control application	Failure
4	No living body	'No Human Detected' shown in the control application	Success
5	No living body	'No Human Detected' shown in the control application	Success

As seen in the tables above, DOORMOR showed success scenarios when tested. The sensor test returned a success rate of 80%. We conducted further investigations on the signal range of the bot and found out that DOORMOR sometimes does not respond to control signals at distances more than 90 meters and stops responding completely at distances greater than 95 meters. So, it is safe to say that, the range efficiency of DOORMOR is 90%.

X. Advantages of DOORMOR

The main advantage of DOORMOR is that it doesn't require any additional remote controller or wireless network for operation. Any android device can be used to control the bot and any existing Wi-Fi network can be used to transfer video data and control signals between the control device and the bot. Wi-Fi repeaters can be used anytime to expand the range of the bot without any kind of hassle. These features also add to the overall cost reduction of DOORMOR and make the robot distinguishable from rest of the similar systems.

XI. Areas of Further Development

As this is only a prototype, the actual robot will be superior to the prototype in many ways. First, the chassis will be a lot stronger compared to the prototype's chassis with better metal alloy construction. Metal chain will be added on the wheels for traversing on rough terrain for hours. The wheels will also be made out of metal. The metal alloy will be selected in such a way that the robot can withstand extreme pressure and heat for a descent amount of time. Better motors will be used to ensure maximum torque and Rotation Per Minute (RPM) under heavy load. The motors will be picked in a way so that they can be powered through the available power source. The power source will also be upgraded to ensure maximum current supply to the components for a longer period of time. The camera module will be upgraded to a night vision camera module to ensure clear imagery even in the dark. The arm will be redesigned with better gears to ensure maximum performance. The Degree of Freedom (DOF) of the arm will be upgraded for better arm movement. Face and object recognition will be added in order to recognize people and objects while searching an affected area. This will help the rescuers to efficiently locate the distressed people.

The robot can be slightly modified for military use, specially bomb disposal. The claw [15] can be replaced with a cutting tool that will allow the robot to cut wires for disposing bombs. Also, the robot can be used as a spy robot.

Another interesting addition would be live GPS location streaming which would help us pinpoint the bot, and addition of sonar sensors which would stop the bot automatically before it accidentally crashes into something. Thus, some sort of Artificial Intelligence (AI) would also be in action.

XII. Conclusion

In this modern era, robotics has become such a field which has made our lives easier, safer and efficient through variety of its applications. In almost every aspect of our daily life, we can see the benefits of this particular field. However, our effort was to develop a life saver. Its applicability is not confined within the pre-specified scientific tasks. Again, its software and the integration amongst the associated hardware components made it equally capable of performing other similar tasks with minimum modification. In the previous section, we have discussed areas of further development. Concentration on traversing task gave the bot such an infrastructure that can withstand expected adversities up to a certain limit and made it capable of exploring places that are almost out of bounds for humans. The compilation of the mechanics, self-developed control system, software and hardware provides the blueprint of such a standalone system that can accommodate certain utilities in favor of assisting humans in various horizons. As there is no end of perfection, with more resources and logistics support plus keen observation of its performance on different sectors, the scope of improvement of the bot is still on the table.

XIII. Acknowledgement

The authors would like to express profound gratitude to Global Emerging Technology Network, Dhaka, Bangladesh for their endorsement as well as technical supervision. Also, the authors of this paper are very much thankful to the Department of Computer Science and Engineering, Bangladesh Army University of Engineering and Technology, Natore, Bangladesh and Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh, for their support throughout the build of this robot. Without their support, we could never pursue such a complicated project.

References

[1] "Robotics and Mechatronics," Last accessed on February 1, 2019, at 12:55:00PM [Online]. Available: https://sst.nu.edu.kz/robotics/

[2] "Disaster Robotics," Last accessed on February 1, 2019, at 1:00:00PM [Online]. Available: http://irides.tohoku.ac.jp/eng/organization/regeneration/04.html

[3] Trupti B. Bhondve, Prof. R. Satyanarayan & Prof. Moresh Mukhedkar, "Mobile Rescue Robot for Human Body Detection in Rescue Operation of Disaster", published on the International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, ISSN: 2320 – 3765, Vol. 3, Issue 6, June 2014

[4] Albert Ko and Henry Y. K. Lau, "Robot Assisted Emergency Search and Rescue System with a Wireless Sensor Network", published on the International Journal of Advanced Science and Technology, Vol. 3, February 2009

[5] Zachary Cody Hazelwood and Dr. Saleh M. Sbenaty, "Designing, Building, and Testing an Autonomous Search and Rescue Robot — An Undergraduate Applied Research Experience", published on the 121st ASEE Annual Conference & Exposition 2014, Paper ID #10699

[6] Siobh´an Grayson, "Search & Rescue using Multi-Robot Systems", published by Semantic Scholar, 2014.

[7] Luis Pineda, Takeshi Takahashi, Hee-Tae Jung, Shlomo Zilberstein and Roderic Grupen, "Continual Planning for Search and Rescue Robots", published on the 2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids), DOI: 10.1109/HUMANOIDS.2015.7363542

[8] Jingchao Zhao, Junyao Gao, Fangzhou Zhao and Yi Liu, "A Search-and-Rescuen Robot System for Remotely Sensing the Underground Coal Mine Environment", published on the Sensors article by MDPI, Sensors 2017, 17, 2426; DOI: 10.3390/s17102426

[9] Jaeeun Shim and Ronald C. Arkin, "The Benefits of Robot Deception in Search and Rescue: Computational Approach for Deceptive Action Selection via Case-based Reasoning", published on the 2015 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), DOI: 10.1109/SSRR.2015.7443002

[10] Johannes Maurer, Gerald Steinbauer, Peter Lepej and Suzana Uran, "State of the Art in Search and Rescue Robots", published on the TEDUSAR White Book version 1.0 by TEDUSAR (www.tedusar.eu)

[11] Illah R. Nourbakhsh, Katia Sycara, Mary Koes, Mark Yong, Michael Lewis and Steve Burion, "Human-Robot Teaming for Search and Rescue", published by the IEEE CS and IEEE ComSoc, 1536-1268/05 © 2005 IEEE

[12] J. Ju´arez-Guerrero, I. Pissokas & D. Farinha, "An Autonomous Team of Search and Rescue Robots", published on AAAI Technical Report WS-01-01, Compilation copyright © 2001, AAAI (www.aaai.org)

[13] "Gear Motor," Last accessed on February 1, 2019, at 1:05:00PM [Online]. Available: https://www.techshopbd.com/product-categories/motors/1101/gear-motor-techshop-bangladesh

[14] "Servo Motor," Last accessed on February 1, 2019, at 1:10:00PM [Online]. Available: https://www.techshopbd.com/product-categories/motors/1337/servo-motor-mg996-techshop-bangladesh

[15] "Claw," Last accessed on February 1, 2019, at 1:15:00PM [Online]. Available: https://www.techshopbd.com/product-categories/others/1297/robotic-claw-techshop-bangladesh

[16] "Arduino Nano," Last accessed on February 1, 2019, at 1:20:00PM [Online]. Available: https://www.arduino.cc/en/Guide/ArduinoNano

[17] "Raspberry Pi," Last accessed on February 1, 2019, at 1:25:00PM [Online]. Available: https://www.raspberrypi.org/products/raspberry-pi-3-model-b/

[18] "Power Supply," Last accessed on February 1, 2019, at 1:30:00PM [Online]. Available: https://www.alibaba.com/product-detail/5v-li-ion-battery-pack_1246735134.html

[19] "Motor Driver," Last accessed on February 1, 2019, at 1:35:00PM [Online]. Available: https://www.techshopbd.com/product-categories/drivers/1749/l293d-dual-dc-motor-controller-techshop-bangladesh

[20] "PIR Sensor," Last accessed on February 1, 2019, at 1:40:00PM [Online]. Available: https://www.techshopbd.com/product-categories/motion/1336/hc-sr501-pir-sensor-module-techshop-bangladesh

[21] "Camera Module," Last accessed on February 1, 2019, at 1:45:00PM [Online]. Available: http://bdspeedytech.com/index.php?route=product/product&product_id=1878

[22] "Wi-Fi Router," Last accessed on February 1, 2019, at 1:50:00PM [Online]. Available: https://ryanscomputers.com/network/router/tp-link-tl-wr840n-v2-300mbps-router.html

[23] "FFMPEG," Last accessed on February 1, 2019, at 2:05:00PM [Online]. Available: https://www.ffmpeg.org/documentation.html

[24] "LIGHTTPD," Last accessed on February 1, 2019, at 2:25:00PM [Online]. Available: https://redmine.lighttpd.net/projects/lighttpd/wiki

[25] "CRTMPSERVER," Last accessed on February 1, 2019, at 2:30:00PM [Online]. Available: https://github.com/j0sh/crtmpserver

[26] "L293D's Truth Table," Last accessed on February 1, 2019, at 2:35:00PM [Online]. Available: https://roboticbasicsblog.wordpress.com/2016/03/30/l293d-motor-driver/

[27] "Sockets," Last accessed on February 1, 2019, at 2:45:00PM [Online]. Available: https://docs.python.org/2/library/socket.html

[28] "Mechanism of the PIR Sensor," Last accessed on February 1, 2019, at 2:50:00PM [Online]. Available: https://howtomechatronics.com/tutorials/arduino/how-pir-sensor-works-and-how-to-use-it-with-arduino/2

VADet: An Arduino based automated vehicle accident detection and messaging system

Nazir Ahmmed
Dept. of CSE
United International University
Dhaka, Bangladesh
goodvubo@gmail.com

Nusrat Jahan Jenny
Dept. of CSE
United International University
Dhaka, Bangladesh
jen.jahan@gmail.com

Most Fowziya Akther Houya
Dept. of CSE
United International University
Dhaka, Bangladesh
fowziyaboshra@gmail.com

Anika Ibnat Binte Alam
Dept. of CSE
United International University
Dhaka, Bangladesh
anikaalamsuhi@gmail.com

Md. Adnan Arefeen
Dept. of CSE
United International University
Dhaka, Bangladesh
adnanarefeen@cse.uiu.ac.bd

***Abstract*—With the advancement of technology and rapid overpopulation, people are constantly trying to ease their daily life by adapting with newer technology. Transport is one of the daily amenities for us to depend on, yet not sufficient enough for the growing population. This is one of the common issues which people try to solve with vehicle purchasing. The increasing amount of vehicles create mismanagement in controlling traffic leading to accidents. Although accidents happen due to various factors other than traffic management, such as unstable weather, reckless driving, faulty vehicles or maybe road conditions. But the most important part after an accident is to detect the accident and take immediate action upon detection. Bangladesh has registered a significant increase, 19.58%, in road accidents in 2018 compared to 2017 and 3,472 accidents in 2017, where there were 4,317 deaths in 2018. All this time we overlooked the fact that immediate aid to an accident scene can reduce the number of people getting traumatized, disabled or lose their precious lives due to lack of emergency facilities. Also, delayed action can lead to criminal activity of the accident doer as the person gets enough time to flee from the scene. We built this system to solve these issues by detecting and confirming the accident using the MPU6050 module, and notifying the proper authorities by sending the location of the incident using SIM808 module.**

***Index Terms*—accident, detection, notification, message, MPU6050, SIM808, GPS, GSM**

I. Introduction

The increasing amount of automobiles causing frequent hazardous traffic situations leading to life-threatening accidents. In most cases, poor emergency facilities increase the level of fatality. Our system provides an optimal solution to an accident scene. This system is able to detect any causality in every possible direction and confirms if the accident occurred or not. On the findings, the exact location in longitude and latitude form is sent to the authorized number through SMS. On any nonfatal situation, drivers have the full privilege on cancelling the notification sending system.

II. Motivation

Our motivation for working on this project came from a recent issue of two students who died in a road accident [1]. Therefore, it came to our mind that accidents can occur in any situation, but the golden time is to take action immediately after the accident as people are traumatized to decide what action they should do next and fall into grave risks. If all vehicles are set to detect accidents and send automated SMS containing location to help the injured right away, the chances of such unfortunate events can be reduced.

The objective of our project is to prevent fatal situation for lack of emergency facilities when an accident occurs. Using this system, if anyone faces an accident, his or her close relatives and nearby help centre will get a notification so that they can immediately come to aid and help them as our system provides the exact location of the accident spot.

III. Related Works

This section demonstrates some of the previous research works and these co-respond to our proposed work. C. Prabha et al. [2] worked with GSM and GPS modem for detecting an accident and messaging system, whereas some works had been done with new methods which do not require GSM and GPS modem. J. R. Barfield, et al. [3] worked with on board devices to detect car crashes. R. C. McKown and J. T. Mader [4] focused on notification and they did it by producing a decision of 3-digit emergency call from an in-vehicle TCU to a local PSAP operator, but before notifying they qualify vehicle crash events.

IV. Abbreviations and Acronyms

SMS(Short Message Service): SMS is a text messaging service in cellular-device systems and the internet. To exchange short text messages it follows standardized

communication protocols.

GSM(Global System for Mobile communication):
GSM is a mobile communication modem, transmits mobile voice and data service. It is an open and digital cellular technology operating at the 850MHz, 900MHz, 1800MHz and 1900MHz frequency bands. It has the ability to carry 64 kbps to 120 Mbps of data rates.

GPS(Global Positioning System): A global navigation satellite system which provides Geo-location (latitude and longitude) and time information to a GPS receiver if there is an unobstructed line of sight to four or more GPS satellites.

V. Components

A. Arduino NANO:

The Arduino Nano uses the Atmega328p Processor. There are total 8 Analog pins and 14 digital Pins on our Nano board. To interface sensors as input pins or drive loads as output pins we used the digital pins. Functions like pinMode and digitalRead, digitalWrite performs the desired actions. [5].

B. SIM808:

SIM808 is integrated with a high-performance GSM/GPRS engine, a GPS engine and a BT engine which meets almost all the type of space requirements, such as M2M, smart phone, PDA, tracker and, other mobile devices. [6].

C. MPU6050:

The MPU-6050 is a device which combines a 3-axis gyroscope, 3-axis accelerometer, and a Digital Motion Processor (DMP). It has dedicated I2C sensor bus, which directly accepts inputs from an external 3-axis compass to provide a complete 9-axis Motion-Fusion output. The MPU-6050 features three 16-bit ADC for digitizing the accelerometer outputs and three 16-bit analog-to-digital converters (ADCs) for digitizing the gyroscope outputs. For precision tracking of both slow and fast motions, the parts feature a user-programmable accelerometer full-scale range of $\pm 2g$, $\pm 4g$, $\pm 8g$, and $\pm 16g$ and a user-programmable gyroscope full-scale range of 250, 500, 1000, and $2000/sec(DPS)$ [7].

D. 16×2 LCD Display:

A 16×2 LCD (Liquid Crystal Display) screen is an electronic display module which can display 2 such lines having 16 characters in 5×7 pixel matrix per line [8].

E. Buzzer:

A buzzer is a small component, adds sound features efficiently to any system. It needs DC power supply ranging from 4V to 9V. it is usually associated with a switching circuit to turn ON or OFF at required time and interval.

F. Push Button:

Push-Buttons are normally-open tactile switches, allows to power any circuit or make any particular connection on pressing the button.

G. Potentiometer:

Potentiometers are nothing but variable resistors which also known as POT. By simply rotating the knob on top of its head, they can provide variable resistance. Potentiometer standards values are $500\,\Omega$, 1KΩ, 2KΩ, 5KΩ, 10KΩ, 22KΩ, 47KΩ, 50KΩ, 100KΩ, 220KΩ, 470KΩ, 500KΩ, 1 MΩ. Potentiometer used in an LCD to adjust the contrast level.

H. Breadboard:

Breadboards consist of four rows of 63 vertical columns with 5 connected holes which is the circuit area. Also, 8 rails (or distribution strips) connected horizontally for power and ground or to carry a signal if it is not needed for power or ground.

I. Jumpers:

A jump wire (also known as jumper wire, or jumper) is an electrical wire used to interconnect the components without soldering, or group of them in a cable, with a connector or pin at each end (or sometimes without them simply ”tinned”).

VI. Proposed System

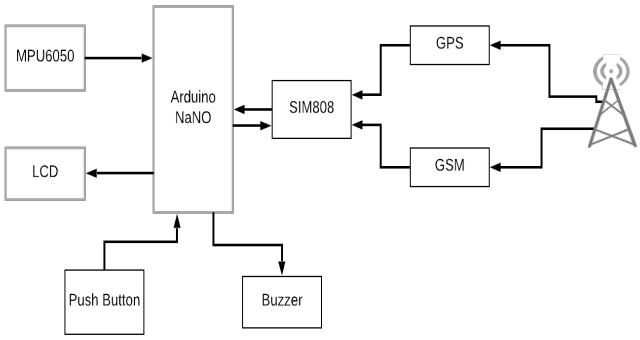


Fig. 1. Block Diagram of the system

Figure 1 shows the working process of the system. Values taken from MPU6050 and SIM808, Arduino has processed them to detect an accident. Upon detection, Buzzer has turned on. Based on the decision, GPS has traced the location and an SMS is sent using GSM. Each status of action is displayed on the LCD.

Figure 2 breaks down the continuous workflow of the system. It continuously checked the conditions of the detection scheme from current values with the previously stored values to identify any type of anomalies. The system is terminated when the notification message is sent.

VII. System Mechanism

For our system, we have tried to consider all possible ways to detect an accident based on the values of acceleration and deceleration of X-axis, Y-axis and Z-axis, pitch and roll and angular offset for rotation of the vehicle. In reality, an accident has many factors but our system has been able to detect most accidents in any situation using these conditions.

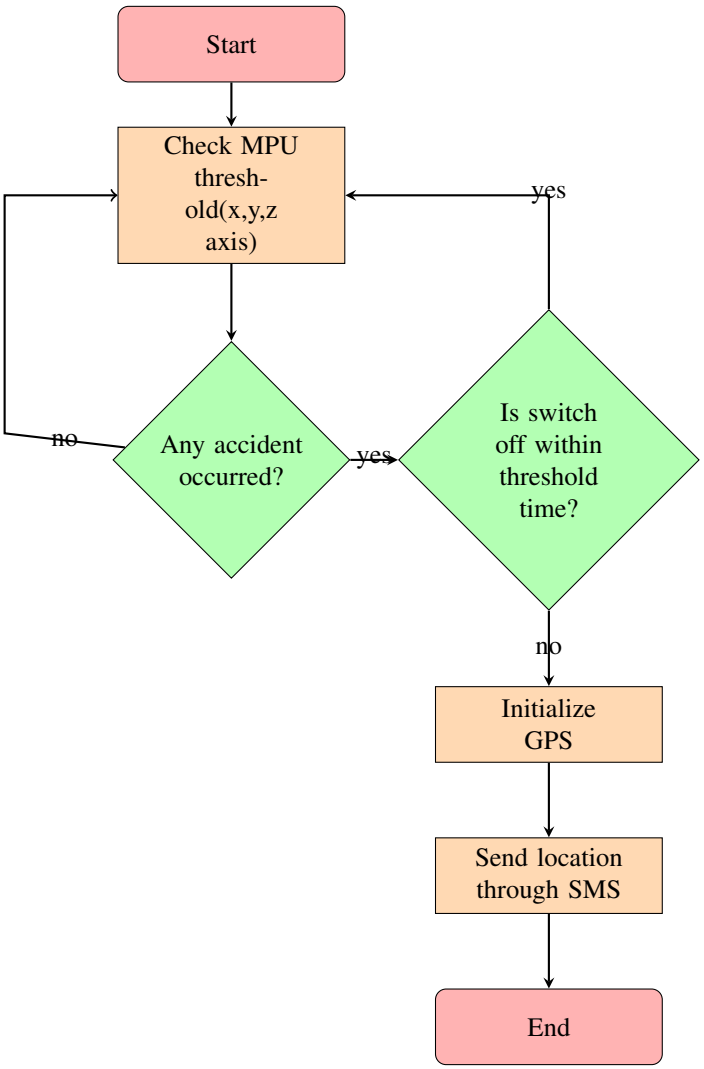


Fig. 2. Flow chart of the system

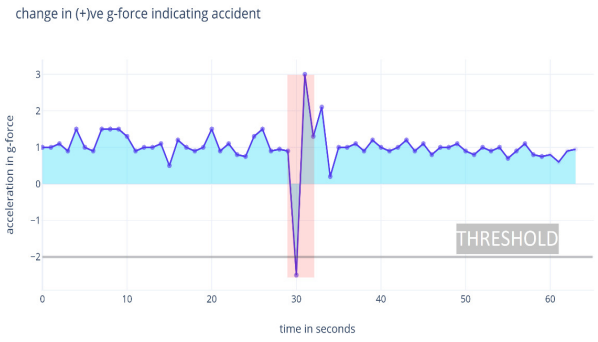


Fig. 3. Spike shows the anomaly of acceleration

Initially, we have set some values such as threshold values lets the system detects an accident, a cellphone number for sending SMS if an accident occurs and initialize an array to check any disruption in the acceleration of the vehicle.

In Figure 3 and 4, we have provided two examples of accident occurring situation. In Figure 3, the acceleration of the vehicle is relatively close to each other but at a point, the acceleration suddenly spikes and we have consider this as an accident. Moreover, a spike point is an anomaly and this usually happens when an accident occurs. Figure 4, shows the deceleration situation of the vehicle.

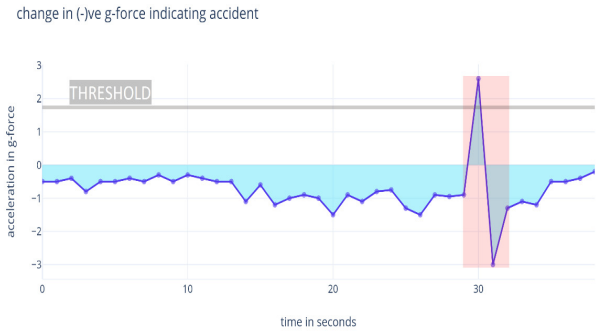


Fig. 4. Spike shows the anomaly of deceleration

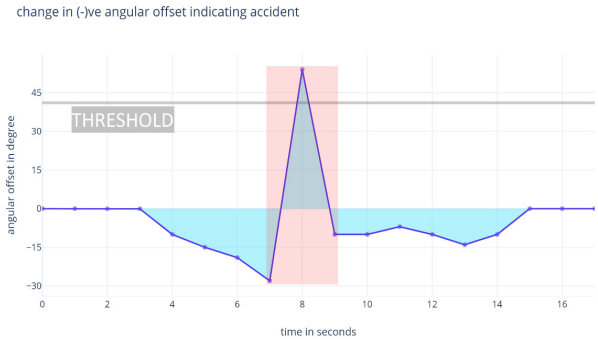


Fig. 5. Angular offset change

We have normalized the raw accelerated values we got from the moving vehicle and have calibrated the values of a gyroscope. Comparing these values with our preset thresholds, the system has decided on any accident occurrence. Moreover, the pitch and roll values were calculated using the following two equations to let the system detect an accident.

$$pitch = \frac{180}{\pi} * \tan^{-1}\left(\frac{aX}{\sqrt{aY^2 + aZ^2}}\right) \quad (1)$$

[10]

$$roll = \frac{180}{\pi} * \tan^{-1}\left(\frac{aY}{\sqrt{aX^2 + aZ^2}}\right) \quad (2)$$

[10]

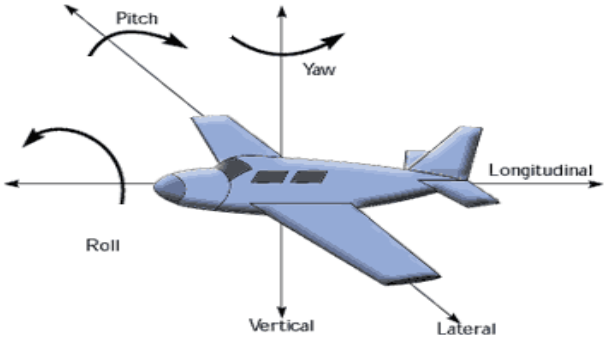


Fig. 6. Movement of an object in three different axis [9]

Here,

$$pitch = \text{rotation of lateral axis}$$

$$roll = \text{rotation of longitudinal axis}$$

$$aX = \text{acceleration about X axis}$$

$$aY = \text{acceleration about Y axis}$$

$$aZ = \text{acceleration about Z axis}$$

$$\pi = 3.1415926535, \text{mathematical constant}$$

Figure 5 and 7 shows the graph of the sudden change on any angular offset of a vehicle. We had situations where a vehicle can take a 360-degree turn but the speed of the car would be within a range. If the acceleration of the vehicle is very much high compared to the previous acceleration rates, in that case, we have considered it as an accident.

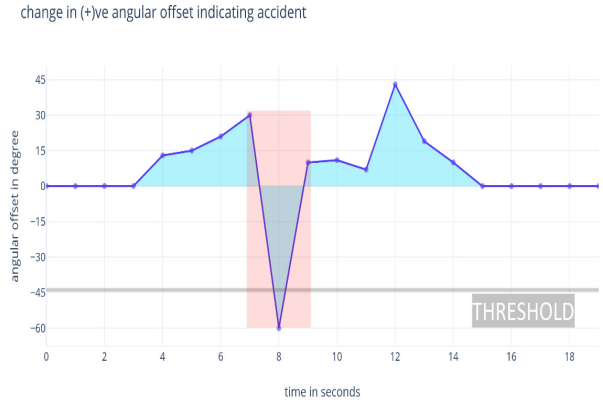


Fig. 7. Angular offset change

Our special technique was to detect an accident in Z-axis. Naturally accidents on Z-axis are the underwater situations, where any vehicle might fall from any cliff and fall into open water. In real life any vehicle accident detection methods are not applicable to the underwater situation.

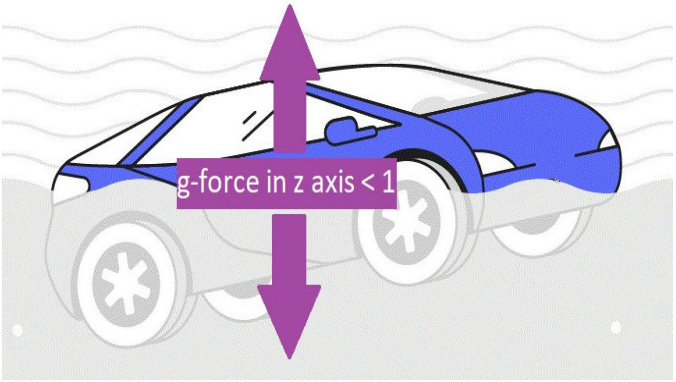


Fig. 8. g-force value change in underwater

If a vehicle faces an accident due to a height such as a fall from a bridge or hill, and falls into water. In these types of situation, our system has managed to detect it is an accident and the techniques for this is g-force checking. Usually, g-force is always 1 for any flat space but in underwater it becomes less than 1, which is shown in Figure 8. Our system has overcome the limitation to detect an accident in underwater.

If an accident is detected by our system, the buzzer is immediately turned on which we initially affix with the vehicle. In case of less severity of the accident or no causality, if the driver wishes not to notify anyone, we have given the privilege to cancel out the notification sent. If the buzzer is turned off within a given threshold time, no message is sent to the authorized number.

After any accident detected, our system waits for the threshold time to run out. If the threshold time has run out but the buzzer is not turned off indicates an accident actually occurred and authorized persons must be notified. Thus an automated message is generated containing a Google map address of the accident location and sent to the authorized number from the number which is already inserted in our system. Since the system is able to identify the exact location of the accident through GPS, thus the accident spot is easily traceable and can be reached within short time.

VIII. Limitations

We have designed the system in a way that can trace the location and send notification only by mobile network. It is not possible for the system to send notification using another network.

IX. Conclusion

Detecting accident and notifying immediately after the accident through SMS provides emergency facilities for any kind of unfortunate situation. In this paper, we have considered all possible ways to detect an accident which usually occurs. We have analyzed different situations logically and mathematically calculated them to make it accurate enough to serve its users.

X. FUTURE WORK

The future direction of our work is to reduce the severity of after accident causalities using more robust modules so that the detection system can become more accurate. We plan on improving the notification system, where we can send a message to the closest response sites i.e. Police station, hospitals. If an accident occurs, our system will match the location and search for nearby help stations and notify them This will help to use the time after accident more efficiently and possibly will save more lives.

REFERENCES

[1] ProthomAlo.com, 'Bus runs over students waiting beside city road, 2 dead', 2018. [Online]. Available: https://en.prothomalo.com/bangladesh/news/180644/2-school-students-killed-as-city-bus-hit-a-dozen.

[2] C. Prabha, R. Sunitha, and R. Anitha, "Automatic vehicle accident detection and messaging system using gsm and gps modem," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 3, no. 7, pp. 10 723–10 727, 2014.

[3] J. R. Barfield, S. C. Welch *et al.*, "Automatic vehicle crash detection using onboard devices," Jul. 12 2016, uS Patent 9,392,431.

[4] R. C. McKown and J. T. Mader, "Qualifying automatic vehicle crash emergency calls to public safety answering points," Apr. 28 2015, uS Patent 9,020,690.

[5] Components101.com,'arduino nano', 2018. [Online]. Available: https://components101.com/microcontrollers/arduino-nano

[6] Cdn-shop.adafruit.com,'SIM808_Hardware Design_V1.00', 2014. [Online]. Available: https://cdn-shop.adafruit.com/datasheets/SIM808_Hardware+Design_V1.00.pdf

[7] Invensense.com,'MPU-6050Product SpecificationRevision 3.4', 2013. [Online]. Available: https://www.invensense.com/wp-content/uploads/2015/02/MPU-6000-Datasheet1.pdf

[8] Engineersgrage.com,'lcd-016m002b specification datasheet', 2002. [Online]. Available: https://www.engineersgarage.com/electronic-components/16x2-lcd-module-datasheet

[9] "Attitude-Pitch/Roll/Yaw." [Online]. Available: https://www.novatel.com/solutions/attitude/

[10] Calculating pitch, yaw, and roll from mag, acc, and gyro data. [Online]. Available: https://engineering.stackexchange.com/questions/3348/calculating-pitch-yaw-and-roll-from-mag-acc-and-gyro-data?answertab=active#tab-top

ParkEasy: An Embedded Model to Mitigate Car Parking Problems Using IoT Technology

Samrat Kumar Dey
Dept. of CSE
Dhaka International University
Dhaka, Bangladesh
samratcsepstu@gmail.com

R. M. Rubayet Shamim
Dept. of CSE
Dhaka International University
Dhaka, Bangladesh
mddigonto3@gmail.com

Md. Atiqul Islam
Dept. of CSE
Dhaka International University
Dhaka, Bangladesh
atiqul58@gmail.com

Md. Mahbubur Rahman
Dept. of CSE
Military Institute of Science and Technology
Dhaka, Bangladesh
mahbubcse@yahoo.com

***Abstract*— In the passage of time, the number of population in Dhaka city is increasing very rapidly. Most parts of this city are highly affected by the huge traffic jam. Peoples of all ages have to suffer and valuable times are being wasted for this long-standing hobble. The inadequate and inadvertent parking system is one of the main reasons for causing traffic congestion. Our prime aim is to develop a smart-phone based car parking model to minimize traffic jam and to ensure efficient car parking facilities in Dhaka city. To attain our aim we have designed and simulated a prototype of reservation based car parking system. Our developed smart-phone application will help the users to find the available parking spot in the nearest location. To maintain the parking details of different users a Java SE application has been developed.**

Keywords*— *IoT, car parking, double-layer security, request process, park share

I. INTRODUCTION

Dhaka is one of the most densely populated cities in the world whose population growth rate is increasing rapidly. At the time of Bangladesh coup, the population of Dhaka was just 3 million, which increased from 48 years to 20.28 million [1]. The population density of the city is approximately 13854/km^2, where the area is 1464 km^2. In this vastly populated area, traffic jam is one of the most burning questions to us. The average speed of automobiles in this city has decreased from 21 km/h to 7 km/h and it is estimated that in 2035 this speed will be reduced to about 4 km/h which is less than mass people walking speed [2].

Over the previous decades, the amount of automobiles has been increased tremendously in our country, especially in Dhaka city. These excessive increases of automobiles work as a vital reason for the city's parking problems, fuel wastage, time wastage, and user frustration which have a greater impact on the urban ecosystem of Dhaka. According to the BRAC Institute of Government and Development, Dhaka city has lost approximately 5 million working hours per day for this traffic congestion, which damages $11.4 billion per year [2]. To overcome this situation, the Bangladesh government (Revised Strategic Transport Plan, 2014-15) has taken various effective steps including five metro rail lines, two rapid bus routes, 1200 km of new roadways, 6 existing elevated expressways and three ring roads [3].

Although these steps can play an effective role in solving traffic congestion, if we considering Dhaka city then only 7 % of the roads are available compared to the area [4]. On the other hand, Paris and Vienna have 25% and Washington DC and Chicago have 40% roads. In Dhaka, 70% of this small number of roads is occupied by automobiles. Moreover, about 3700 new automobiles are being added to the city in every year [4]. In modern metropolitan cities, it takes an average of time to get a user's parking facility which is analyzed, 35% of the cases require 10 to 15 minutes, in which 30% of the cases increase by more than 20 minutes [9].

In Dhaka city, where there is a lack of parking reservation policy, our main aim is to increase the parking reservations as well as modernize the system with an Internet of Things (IoT) technology. On the contrary, accelerating its development by providing parking spot share facilities is also a key factor. Often the users faced the problem of finding a parking spot to their desired location. However, they are unable to manage a parking position according to their needs and therefore they cover a long distance from that position to the workspace. Although the numbers of new automobiles are being added to the city every year, there has been no development of the parking system. As a result, many of these automobiles are parked on either side of the streets. There are also a few parking reservation systems, but they do not have enough modernization, hence searching for a desired parking spot at the correct time has always been a challenge in Dhaka with impact on traffic, pollution level, wastage of fuel, and user frustration. Our main focus is to develop a convenient car parking mechanism in order to alleviate the traffic congestion at Dhaka city. Moreover, we have portioned our model into threefold: Android Application, Embedded Controller and Java Desktop Application.

II. RELATRD WORK

We have studied various related articles for the development of our proposed architecture. In this segment, we will briefly discuss some relevant work that helps us to gain additional information that is available in literature.

Kianpisheh *et al.* [5] proposed a Smart Parking System (SPS) Architecture using Ultrasonic Detector which is actually works on echolocation (the location of objects by reflected sound), emits a sound and then it is reflected and return to the sensor. In this paper authors only provides information on parking lots and the available numbers of parking. It can be defined as a general Informative Parking System (IPS) rather than Smart Parking System (SPS).

In another study, Kaarthik *et al.* [6] showed an Image Processing Based Intelligent Parking System where they developed an empty lot detection technology by processing of image and displayed the data through 7 segments LED display. In this method, they showed only parking lot in a

specific parking zone and they were not concerned regarding data storage.

A. Nandugudi *et al.* [7] proposed a crowd sourcing-based automated system called PocketParker for informing the free parking available to the users. However, this system is unable to provide a probability of free parking spot and accurate information about free parking or street parking.

In another paper, Rachapol et al. [8] also proposed a Smart Parking Using IoT Technology and showed an approach by using a mobile application to know vacant parking lots. The significance of this article is, they added a mechanism for detecting vehicle plate number to monitor vehicles in the parking area. They also designed the payment process using a mobile application. But in certain cases, if any user needs urgent parking, their system won't be able to perform accordingly to address the issue.

Authors in [9] proposed an incentive-based crowdsource bidding service for parking reservation. Here they presented a method to solve parking problem by using crowd source based approaches. In their paper, the central system only collects information on free parking with the help of third-party users called bidders and provides services to the users. This system is smart and cost-effective to simulate. If bidders are dishonest or absence, the whole process will fail to execute.

O. Orrie *et al.* [10] presented a wireless smart parking system providing location and availability of parking spots using android application. Their proposed system informs users about the nearest lots status within 2 km away and they used proximity sensors to read current data in every 2 minutes.

Chungsan Lee *et al.* [11] addressed a unique approach in terms of smart parking system by employing the Bluetooth communication between smartphone and wireless sensor. They used Ultrasonic sensor and Magnetic sensor to detect vehicle in indoor or outdoor parking field.

Authors in [12] proposed an algorithm to increase the efficiency of crowd source based smart parking system. Here, they proposed a system that helps users to find free parking using least cost by considering the distance and a total number of free places in each car parking reservation system. By this system, the user can get information about the nearest free parking but in crowded cities where finding free parking is a great challenge, this system will not function properly.

Chieh-Hsun Huang et al. [13] designed an intelligent parking system where drivers can get parking information like vacant parking space, parked vehicle status, remote end monitoring etc. through Micro Aerial Vehicle (MAV). Here they also proposed QR code scanning method for finding vacant space or parked vehicle position. In their system, authors focused only on a large parking reservation system rather than providing information about all nearby parking location.

In another exploration, Vanessa *et al.* [14] proposed a wireless sensor networks based car parking model by deploying wireless sensors into car parking field. They only designed a parking spot by providing availability information, security management system and static report.

III. System Architecture

In this segment, we will describe in detail regarding our system architecture. Following Fig.1. illustrates our proposed architectural diagram. Each module of our designed model is interconnected with other components, which can easily be analyzed from Fig. 1. According to our proposed architecture user needs to complete their registration with a unique username. After successful registration, users are allowed to log into their account and be able to access all facilities of our designed system. By using our developed mobile application, users can book a parking spot, share their free parking spot, request for urgent parking and so on. In order to ensure safety, we also designed a Double-layer security mechanism.

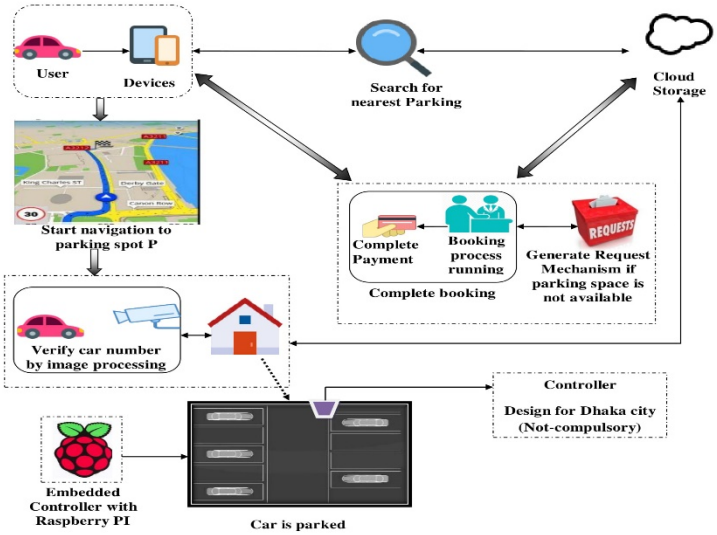


Fig. 1. Architectural diagram of our proposed model

A. Overview of our proposed system

In this section, we will briefly discuss the working procedure of our proposed architecture. We have segmented each module into individual processes. Following processes will provide us with an overview regarding the working approaches of this system.

Process (1): *Checking the nearest parking location*

After successfully login into android application, users can easily check for the vacant space through Google map using our developed mobile application. They also will be able to find all information including parking rate, vacant status, with the street address of the desired parking location and so on.

Process (2): *Request and booking procedure*

Based on availability of parking spot user will be able to confirm booking and they can request for urgent parking if there is no vacant space.

Process (3): *Booking confirmation and payment completion*

After selecting a vacant parking slot users need to complete the payment either the online payment gateway system or via cash to confirm their booking.

Process (4): *Driving to the parking lot and park the vehicle*

Users will be able to reach their desired location whenever the booking is confirmed. The process of number plate detection will be performed in order to entering to the parking spot.

Process (5): *Releasing the vehicle and Exit process*

In order to releasing the vehicle from parking space, number plate verification process will be performed to exit from the spot.

B. Elements of the system

The following section will discuss the interconnection between various components that we have proposed to build this Model. Following Fig. 2. clearly depicts the architectural diagram of our proposed model. All the connected elements of this model are discussed briefly in this segment. We have segmented our system elements as follows.

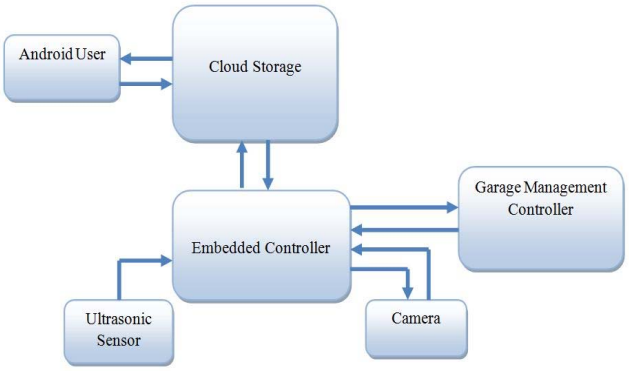


Fig. 2. Architectural diagram of interconnection of individual elements of our proposed system

Hardware Elements:

Sensor: We have used the ultrasonic sensor in order to detect the exact location and any displacement from previous location to new location of the respective car.

Camera: For capturing real time data of parked cars, we developed an image processing model. A quality camera will be used to capture the data of respective cars.

Embedded Controller: Among all components, it is considered as prime one. We have used Raspberry pi3 as an embedded controller.

Software Elements:

Cloud Control API : By Cloud Control API, the image of the number-plate is converted into text data through image processing and later the data is stored on the cloud server (Cloud Storage Databases).

Android Application: Through the mobile application, it provides information using the various Google API to find parking at the desired location.

Desktop Application: One of the features of our proposed ParkEasy is to control the garage through the desktop application. Therefore, we have developed a Java desktop based application to handle the real time data from the sensor.

Garage Management Controller (GMC): GMC is another controller that is designed for providing services both registered and unregistered users.

IV. METHODOLOGY

In this portion, we will explain each module of our proposed model with in-depth analysis. Moreover, our main methodology behind this effort will be enlightened here. To develop an effective and secure car parking system to alleviate traffic congestion at Dhaka city we have followed the following methods which derives us to develop a cost-effective and useable car parking model for Dhaka city. The following phases are obligatory in order to build our proposed propositions.

Garage Automation: Existing parking reservation systems of Dhaka city will be coordinated with the Ultrasonic sensor and digital camera which will be combined with IoT technology. The data from the camera and the sensor will be updated in real time to the cloud storage. As a result, it is possible to get all the information about parking from any place through our smart-phone application. Following Fig. 3. indicates a sample diagram of an automated garage. Aside from parking reservation systems, there will be an available parking lot display, where real-time data will be displayed at every moment. Also, we will include modern facilities like Indicator display, Garage map, Security alarm and so on.

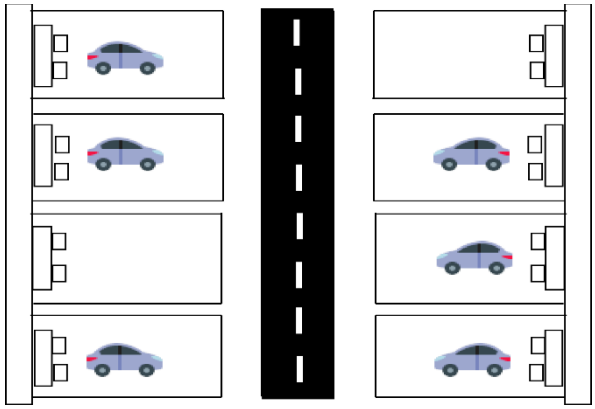


Fig. 3. Sample diagram of an automated garage

***Embedded Controller*:** IoT technology based Embedded Controller plays a vital role in building our model. In our system, we will use Raspberry pi3 as an embedded controller. This embedded controller fetches the data from the sensors and cameras. It also operates simultaneously collects real time data and transfers the processed data to the cloud server.

***Application Functionality and User Interface Design*:** Our developed smart-phone application is another prime component of the system. Through this application user will be able to find parking lot at desired location by the help of Google API. Following Fig. 4. shows the User Interface (UI) of our developed Mobile application. Users will be able to book and request parking places by clicking on the desired location on the Map. From the garage map, user also will be able to view the real-time image of the desired parking spot and its status.

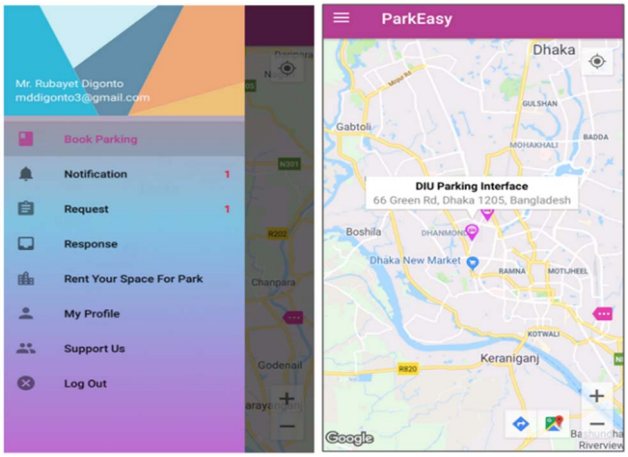


Fig. 4. User Interface (UI) of our developed application

GMC: By introducing the concept of GMC, we already have developed a Java SE based desktop application. Following Fig. 5. provides us with an overview that how our developed desktop application works.

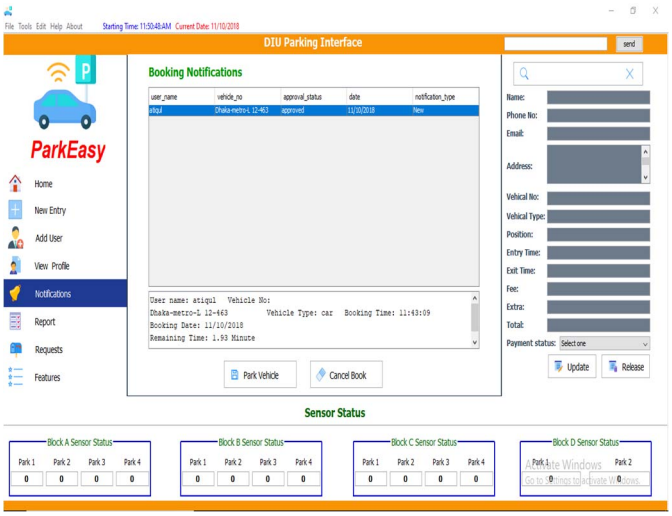


Fig. 5. Java based desktop application

***Number plate detection and recognition*:** Authentication will be provided through number plate detection for both car entry and exit. Following Fig. 6. depicts the captured images with different measurement for number plate verification purposes. To confirm the location of the car in the parking space, we will use the Cloud API. Image processing method will be applied through Number plate detection and respective data will be sent to Cloud server.



Fig. 6. Captured number-plate at different distance and angle(*44cm and 55cm*)

***Double-layer security system*:** To ensure the maximum security of parking cars, we have developed a double-layer security system that is depicted in the following Fig. 7. In the first layer, we will follow the number plate verification process and in the second layer, parking slots will be surrounded by sensors and cameras. So if the car moves away from the sensor's coverage area, the alarms will be ringing and also it will be notified to the user as well as GMC via our developed application.

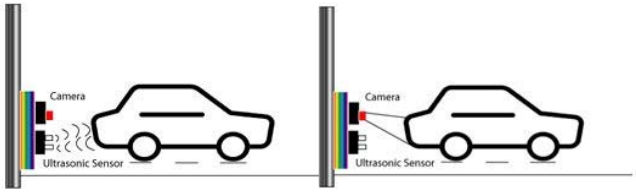


Fig. 7. Vehicle covered by ultrasonic sensor and camera

Share the Park or Rent your space: The unused basement of buildings like Shopping malls, Residence-buildings, Hospitals etc. that will be converted into a parking zone. Anybody interested can rent or share single to N numbers of unused spaces by using our developed mobile application. As a result, it is possible to increase parking lots without creating new parking reservation system. The complete process for sharing the parking lots is illustrated in following Fig. 8.

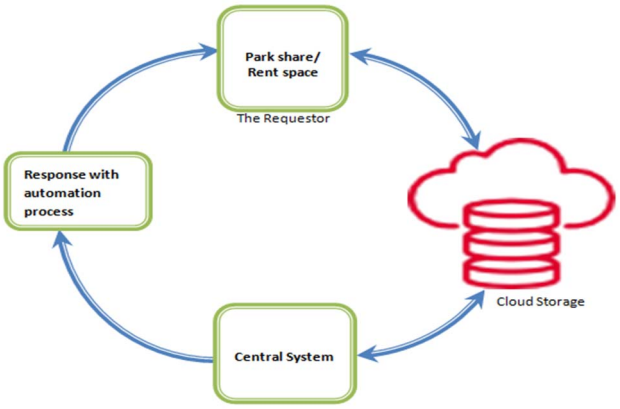


Fig. 8. Process diagram of parking zone sharing

Visualization of Parking Garage: Visualization of the parking garage helps the users to track the status of parking spot, mapping and direction to the slots inside the spot and so on. Following Fig. 9. shows how our developed mobile application interface will visualize the real time availability of parking slots.

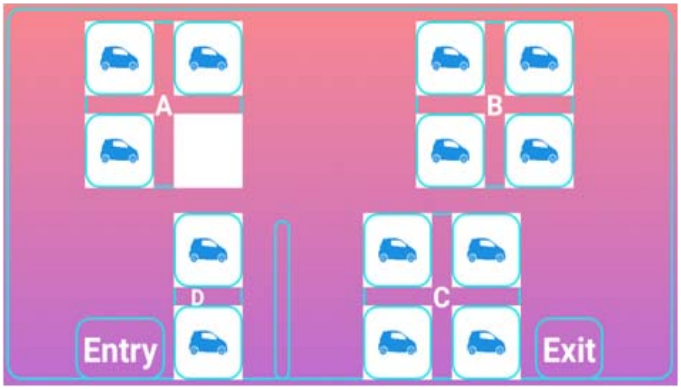


Fig. 9. Garage Map visulation of our designed parking garage

Request Mechanism: This mechanism specially is designed for those users who cannot manage parking spot at their desired location. In this case, users can request for urgent parking for their desired location. Fig. 10. depicts the UI of request and response process for urgent parking of our developed application.

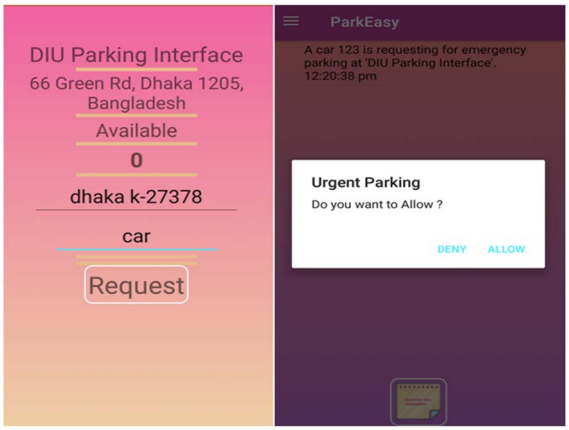


Fig. 10. Request and response process for urgent parking

V. IMPLEMENTATION

This segment will cover implementation procedure of proposed methodologies. In this portion, we will represent our simulation model with working procedure of our designed architecture. Apart from that, we will also explain how car booking process, parking and releasing process, and process of request mechanism works. Following Fig. 11 depicts the entire architecture of our simulation. In general we used some toy cars for simulating our proposed design with the support of developed android mobile application, garage management controller, and embedded controller. The later part of this section will emphasize on the various process and the sequential steps of its working procedure.



Fig. 11. Simulation of *Parkeasy* model

A. Booking process

TABLE I. SEQUENTIAL STEPS FOR BOOKING PROCESS

Step 1: Take user location
Step 2: Find nearest parking location
Step 3: Select the digit note node
Step 4: Book a new Space to park if available exist
Step 5: Request to a parking space if available does not exist
Step 6: Confirm the request if gets any response for urgent parking
Step 7: Get Notification

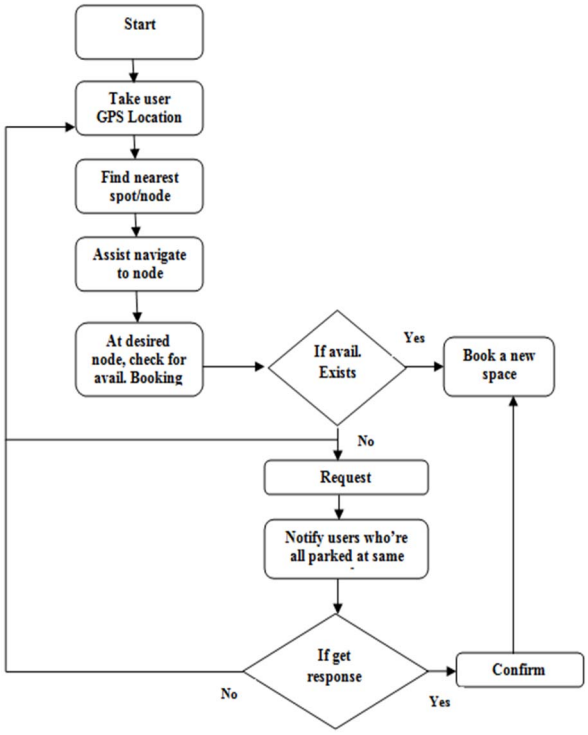


Fig. 12. Flowchart of Booking process

B. Parking and Releasing process

TABLE II. SEQUENTIAL STEPS FOR PARKING AND RELEASING PROCESS

Step 1: Complete brooking a space for parking
Step 2: Transaction of payment
Step 3: Getting the chosen or given slot number/opposition
Step 4: Verification by vehicle number plate detection and recognition
Step 5: Opening the gate to enter to the parking lot
Step 6: Notifying after successful parking
Step 7: Displacing of vehicle, notify to the user
Step 8: Repeat of step 4
Step 9: Release vehicle and exit
Step 10: Notify to the user

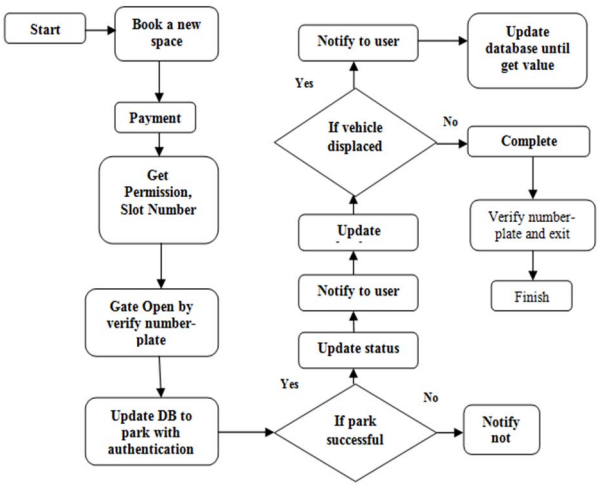


Fig. 13. Flowchart of Parking and Releasing process

C. Process of RequestMechanism

TABLE III. SEQUENTIAL STEPS FOR REQUEST MECHANISM

Step 1: Request to a parking space if available is zero
Step 2: Generate request to the all users who are currently parked at the same parking lot
Step 3: Notify the requestor if any response for urgent parking
Step 4: Release the donator
Step 5: Confirm the donator's slot for the requestor (here, first donator's slot will be booked for the requestor if donator releases himself)
Step 6: Notify the requestor

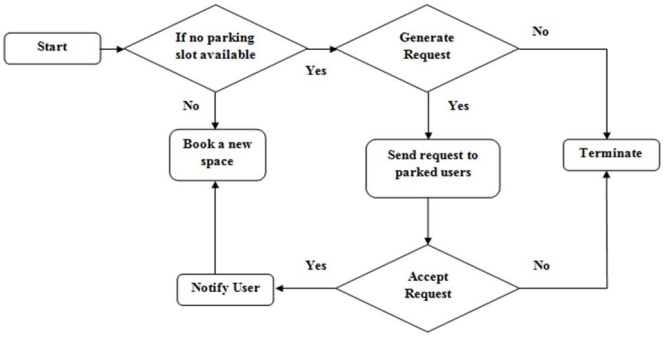


Fig. 14. Flowchart of Request Mechanism

VI. EXPERIMENTAL RESULT

In order to evaluate the effectiveness of our proposed model we have performed a street survey based on the different status shown in TABLE IV. We surveyed 100 people around the Dhaka city.

TABLE IV. SURVEY RESULTS ON CAR PARKING STATUS(BEFORE PROVIDING INFORMATION)

Status	Percentage
Do not know about parking location	60%
Do not know about parking system	17%
Cannot manage lot at desired parking location	10%
Not interested to park	8%
Don't care about parking	5%

After providing information by mobile application, the following users will be informed about the advantages. The number of users will know about parking location (60*90/100) or 54%, since 90% users of them have smartphone. Similarly, the other status of the TABLE IV will have different proportion and we have enlisted them in TABLE V.

TABLE V. SURVEY RESULTS ON CAR PARKING STATUS (AFTER PROVIDING INFORMATION)

Status	Percentage
The number of users will know about parking location	54%
The number of users will know about available and others information	15.3%
The number of users will know about urgent parking but failed to manage parking	9%
The number of users will know about the parking info though they are not interested	7.2%
The number of users will know about the parking info but they do not care	4.5%

So the total number of people will be benefited (54+15.3+9) or 78.3%. As 9% of users cannot manage the parking spot, the actual number of users will be (78.3-9) or 69.3%. At this initial stage, if we are able to provide service to 20% among the 69.3% users, then (69.3-(20/100*69.3)) or 55.4% will remain parked beside roads. By increasing from 20% to 50% and 80%, the number of parked vehicles beside roads will be reduced. The calculated value based on the result is summarized and shown in the following TABLE VI.

TABLE VI. COVERAGE USERS UNDER SERVICE ACCORDING TO EXPERIMENTAL DATA

After providing service to 20%, vehicles remain parked beside roads	After providing service to 50%, vehicles remain parked beside roads	After providing service to 80%, vehicles remain parked beside roads
55.44%	**34.65%**	**13.86%**

VII. CONCLUSION

In this era of rapid urbanization, designing the city system modern and systematic by maintaining everything is a major challenge. Under the pressure of increasing population, Dhaka is becoming tougher to change. Our proposed solution is to keep the city parking system as a fully automated and additionally use garage management controllers system so that user can get parking facilities with the android application as well as without android application. To best of our knowledge, we are the first to design an IoT based parking system with request mechanism and parking share facility. To make parking spaces more secure, we also developed a Double Layer Security System. Currently, we are working on it and in near future, we will be able to implement it for real-life development at Dhaka city.

REFERENCES

[1] "Dhaka Population 2019," Oct. 27, 2018. [Online]. Available: http://worldpopulationreview.com/world-cities/dhaka-population. [Accessed: Jan. 3, 2019].

[2] Abu Afsarul Haider, "Traffic jam: The Ugly side of Dhaka's development," The Daily Star, May 13, 2018. [Online]. Available: https://www.thedailystar.net/opinion/society/traffic-jam-the-ugly-side-dhakas-development-1575355. [Accessed: Jul. 5, 2018].

[3] Robert Gallagher, "Bangladesh Priorities: Urban Transport, Gallagher," Copenhagen Consensus Center, 2019. [Online]. Available: https://www.copenhagenconsensus.com/publication/bangladesh-priorities-urban-transport-gallagher. [Accessed: Jan. 5, 2019].

[4] Michael Hobbes, "Welcome to the Traffic Capital of the World," THE NEW REPUBLIC, Jul. 3, 2014. [Online]. Available: https://newrepublic.com/article/118416/what-dhaka-bangladesh-traffic-capital-world-can-teach-us. [Accessed: Nov. 11, 2018].

[5] A. Kianpisheh, N. Mustaffa, Norlia & Limtrairut, Pakapan & Keikhosrokiani, Pantea, "Smart Parking System (SPS) Architecture Using Ultrasonic Detector" 2012 International Journal of Software Engineering and Its Application, pp. 6.

[6] K. Kaarthik, A. Sridevi and C. Vivek, "Image processing based intelligent parking system," 2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE), Karur, 2017, pp. 1-4.

[7] A. Nandugudi, T. Ki, C. Nuessle, and G. Challen, "PocketParker: pocketsourcing parking lot availability" In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14). ACM, New York, NY, USA, pp. 963-973.

[8] R. Lookmuang, K. Nambut and S. Usanavasin, "Smart parking using IoT technology," 2018 5th International Conference on Business and Industrial Research (ICBIR), Bangkok, 2018, pp. 1-6.

[9] S. Noor, R. Hasan and A. Arora, "ParkBid: An Incentive Based Crowdsourced Bidding Service for Parking Reservation," 2017 IEEE International Conference on Services Computing (SCC), Honolulu, HI, 2017, pp. 60-67.

[10] O. Orrie, B. Silva and G. P. Hancke, "A wireless smart parking system," *IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society*, Yokohama, 2015, pp. 004110-004114.

[11] C. Lee, Y. Han, S. Jeon, D. Seo and I. Jung, "Smart parking system for Internet of Things," *2016 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, 2016, pp. 263-264.

[12] T. N. Pham, M. Tsai, D. B. Nguyen, C. Dow and D. Deng, "A Cloud-Based Smart-Parking System Based on Internet-of-Things Technologies," in *IEEE Access*, vol. 3, pp. 1581-1591, 2015.

[13] C. Huang, H. Hsu, H. Wang, T. Yang and C. Huang, "Design and management of an intelligent parking lot system by multiple camera platforms," *2015 IEEE 12th International Conference on Networking, Sensing and Control*, Taipei, 2015, pp. 354-359.

[14] V. W. s. Tang, Y. Zheng and J. Cao, "An Intelligent Car Park Management System based on Wireless Sensor Networks," *2006 First International Symposium on Pervasive Computing and Applications*, Urumqi, 2006, pp. 65-70.

DESIGN AND IMPLEMENTATION OF BOILER AUTOMATION SYSTEM USING PLC

S.M. Tahsin Labib
Department of Electrical & Electronics Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
tahsinlabib2@gmail.com

Sohan Ul Alam
Department of Electrical & Electronics Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
Sohanulalam1996@gmail.com

Shafayat Hossain
Department of Electrical & Electronics Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
shafayat212@gmail.com

Md.Iquebal Hossain Patwary
Department of Electrical & Electronics Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
iquebalsoykot@gmail.com

Rubel Ahmed
Department of Electrical & Electronics Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
jxrubelahmed@gmail.com

Md.Ashikul Islam
Department of Electrical & Electronics Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
ashiknishan12@gmail.com

***Abstract*— As the world is globalized day by day, so the demand of automated machines has increased rapidly. In industry or any power plant boiler hold a significant contribution. To control it industry need uninterrupted monitoring at frequent interim. But there have number of restriction to measure various stage of boiler by worker. In this paper there has been designed and implemented an automated boiler control system using programmable logic control system (PLC) to exceed respective errors. The main operation of this project is to start and stop boiler system automatically according to set value. There has used sensor to measure the temperature and control it by sending this signal to PLC. Moreover, the boiler system can run for a certain period using the ON-OFF push button.**

Keywords—Boiler Automation System; Programmable Logic Controller (PLC); NTC, SMPS.

I. INTRODUCTION

The demand for automated machines that can operate much efficiently to produce higher quality products has increased. In case of boiler plant, there needs proper control of sensors and proper monitoring of temperature regularly. Thus handling the plant manually is become arduous job.The idea of boiler automation was introduced to achieve the goals involves with the automation control system are operation development, improving management level and high efficiency process . As handling the boiler manually is become arduous job, industrial sector is adopting automation.

A programmable logic controller (PLC) is an industrial automated system that constantly monitors the input device states and makes decisions based on a custom program to control output device states[1].Basically boilers are used to produce steam which will be used for different operations like to run the turbine which in turn produces the electricity. During this process it is essential to monitor the temperature, steam pressure, water level, etc. for reliable operation and to avoid errors and damages. But with the human workers there are a number of possibilities of errors while measuring at different stages. The boiler automation techniques are very much important to avoid the errors due to human workers. Previously, a several number of studies has been done on boiler automation system. In industrial sector automation was done using various techniques. T. Karuppiah and Dr. Azha developed a system by using microcontroller and GSM technology for power plant boiler automation[2]. In their proposed system there have a number of boilers supplied by a main water tank. There have ultrasonic sensor for controlled the water level in the main water tank. There have some temperature sensor in every pipes of boiler which controlled the temperature of the system. The GSM module is used for give information to the consumer. Udaykumar S. Kulkarni & Venkat N. Ghodke proposed a design of boiler automation using FPGA & GSM[3]. In their proposed design a main water tank is connected to several boilers. water level sensor & temperature sensors are used to control the water level of the main tank & temperature of the boilers. GSM mobile phone are used to get information about the current temperature in any boiler. VHDL (VHSIC Hardware Description Language) is used and implemented in hardware using FPGA (Field Programmable Gate Array).Aaron James et. al. proposed a boiler automation system for a power plant using distributed control system (DCS)[4].The system conventionally used as a water tube boiler up to approx. 300 bar or as a quick steam boiler up to approx. 30 bar. Here with the assistance of DCS the output steam from the steam distribution header encompasses a desired temperature by proper controlling of temperature. However, the functions of this systems have not yet been clearly defined. The cost is so high in Distributed control System & in using FPGA. In a boiler system by using

microcontroller, response rate is low. Its operation is not flexible and durability is high.

The main aim of this paper is to design and implement a boiler automation system using PLC for ensure efficient operation of boiler and digitalize the industries. To achieve the goal of this project a LOGO PLC was employed to control the operation of the boiler. In this project a heater jug was used as a prototype of boiler. A NTC (Negative Temperature Coefficient) thermistor along with a temperature showing module is used to measure the temperature. Then the thermistor sends signal to the PLC through a regulator IC 7812. And therefore according to the predefined value set via ladder diagram into the logo PLC, it performs the on-off operation of the boiler accordingly.

II. Overview of the System

"Fig. 1", shows the block diagram of the complete system. 220v AC voltage is used to give the power supply from which necessary power is given to all equipment to activate it.In this system water heater tank is used as prototype of the boiler where there have temperature sensor which senses the temperature.This resulting temperature signal sends to PLC through op-amp as a voltage and PLC makes decision based on the predefined programing. The boiler system is shut down by PLC when the temperature exceeds set maximum value. At the same time, a cooling fan starts running to release the high temperature to a certain limit and then the cooling fan stops. Then when the temperature is below set minimum value, the boiler system is started by PLC again. Thus, the above process is continued. Timer is used for timing operation of the boiler and manual switch is used for emergency switching off the whole boiler system.

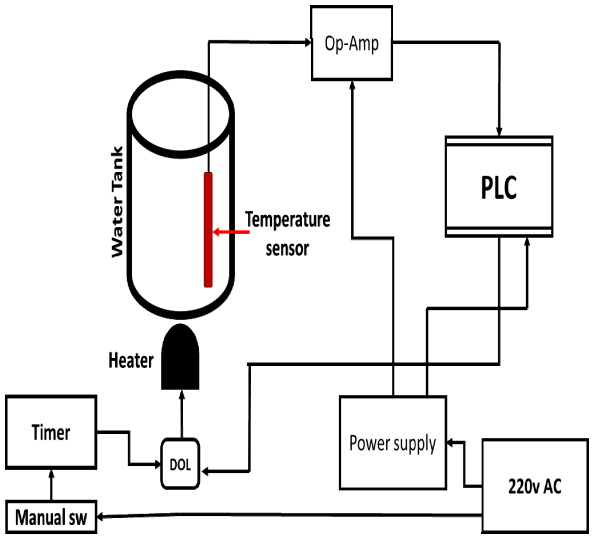


Fig. 1. Block Diagram of the System.

III. System Design

This section describes the methodology to develop the prototype of the system. In this project a LOGO PLC was employed to control the operation of the boiler. The development process includes a heater jug, NTC(Negative Temperature Coefficient, Regulator IC 7812, Timer and DC relay.

A. PLC(Programmable Logic Controller)

The functions for the system of PLC is simple and genuine which is shown in "Fig .2." It has three processes completes by CPU. first one is reads or scans from the input devices. Then it solve the program logic and at last updates a Program or Logic for the CPU to execute. Ladder Logic, Instruction List, Sequential Function Charts, or any of the IEC languages can be written for this logic. Then the program is downloaded to the PLC by the programmer which is done for short time to connecting the programmer to the PLC. If the program is installed once to the CPU then its unnecessary for the PC to remain connected [5].

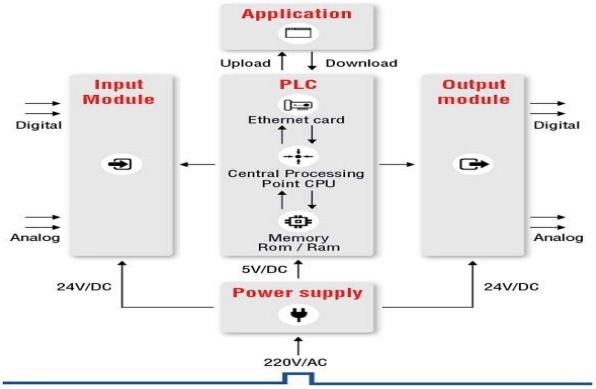


Fig .2. Operation of PLC [5].

B. NTC(Negative Temperature Coefficient) Thermistor

NTC thermistors shown in "Fig.3." are temperature sensors, basically are resistors have a negative temperature coefficient.It means the resistance of this sensors decreases with the increasing temperature.The temperature sensitivity coefficient of this sensors is about ten times greater and five times greater than that of resistance temperature detectors (RTDs) and silicon temperature sensors (silistors) respectively. NTC sensors are usually employed at −55°C to 200°C [5]. But some special families of NTC thermistors can be used at temperatures close to absolute zero (-273.15°C) and some are specifically designed for use above 150°C. The temperature sensitivity of NTC sensors is expressed as "percentage change per degree C" which depends on the materials used in it and the specifics of the production process. For change of every °C the temperature sensitivity range changes from -3% to -6%[6]. The characteristic curve of NTC thermistor is shown in "Fig.4."



Fig. 3. NTC thermistor [6].

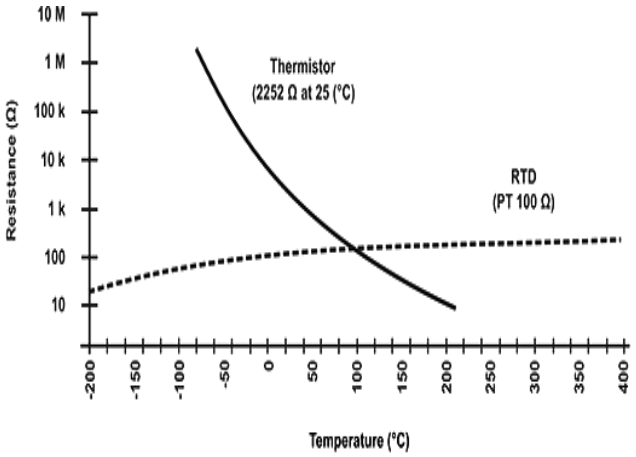


Fig .4. Characteristics of NTC Curve [6]

C. *Switched Mode Power Supply(SMPS) (AC to DC)*

SMPS is an AC to DC converter which is used to attain a fixed DC output voltage from unconditioned AC or DC input voltage.When it get an AC input SMPS convertes this AC input into DC by using rectifier and filter and by using a power MOSFET amplifier the switching is done.By this way very high gain can be attained. The switching action is controlled by a feedback through PWM oscillator and a switching frequency is chosen which must be inaudible to normal human beings (mostly above 20KHz). A feedback circuit is used which control the output voltage by comparing it with the reference voltage.If there is AC portion in the output side this AC voltage is again fed to the output transformer shown in the figure to step up or step down the voltage levels and by using the output rectifier and filter the output of this transformer is again rectified and smoothed [7]. The operation of AC to DC converter SMPS is shown in "Fig.5."

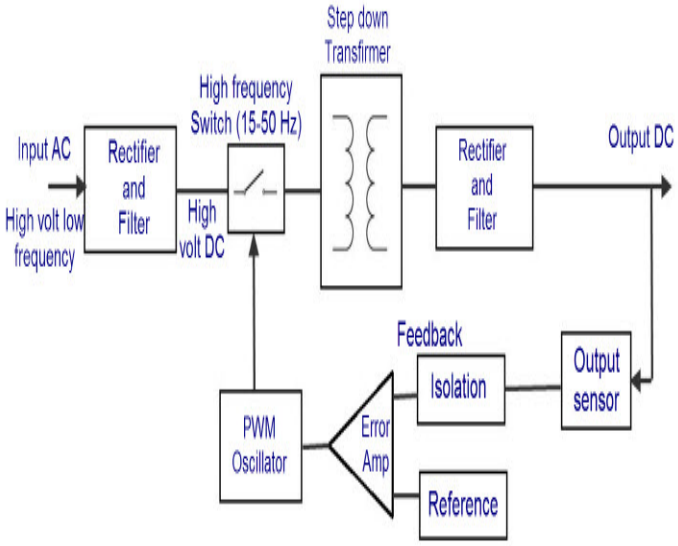


Fig.5. Operation of SMPS [7].

D. *Ladder diagram of system*

LOGO! Soft Comfort V6.1 software is used in this system. Ladder diagram method is used as a programming language here. The Logo PLC was enabled using this software. This software is used for drawing the ladder diagram of this system.

In the ladder diagram I have used make contact, break contact, timer, analog input, analog threshold trigger, lathing relay and relay coil blocks. There is two type of operation in this system. One is normal operation and another is timing operation. In normal operation, analog input (AI1), analog threshold triggers (SF001, SF003, SF005, SF007, SF009) lathing relays (SF006, SF008) and relay coils (Q1, Q3) are used. And in timing operation, break contact (I1), make contact (I2), Timer (T004), Timer break contact (T004) and relay coil (Q2) are used as shown in "Fig.6." and "Fig.7."

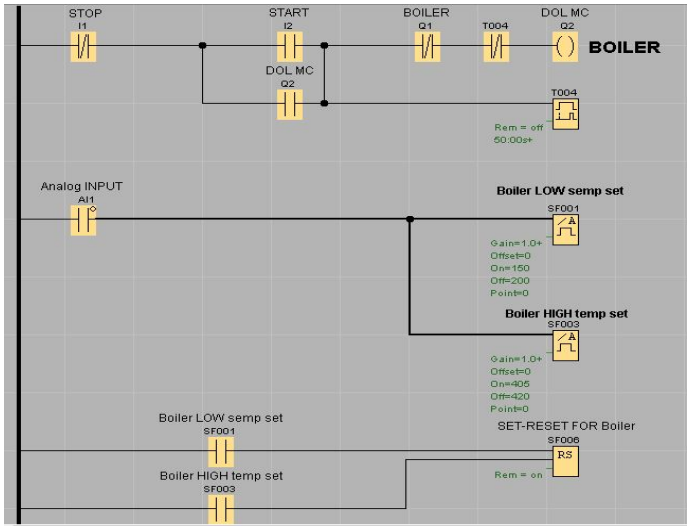


Fig. 6. Ladder diagram of the system.

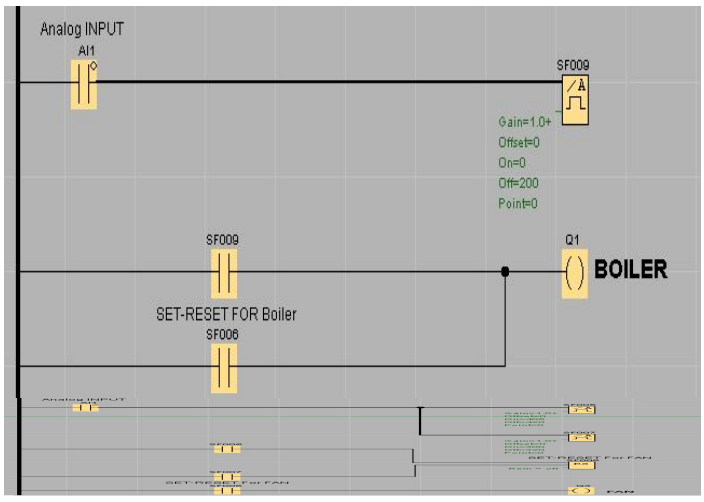


Fig.7. Ladder diagram of the system.

E. *Circuit diagram of this system*

After giving AC power and switching on the power button, the main power supply gets activated. NTC Thermistor is used to sense the temperature of the boiler. NTC Thermistor has internal resistance of 10K at 25°C that changes with the change of temperature. In this project, the voltage divider rule is applied and the divided voltage is given in analog input 1 of PLC. But due to change of temperature the internal resistance of NTC Thermistor changes and thus the voltage of analog input 1 changes. By calculation it is possible to get the voltage of analog input 1 of Logo PLC (OBA5) for every temperature changes. By using ladder diagram I have set the function of Logo PLC (OBA5) to work as follows:

At 30°C the voltage at analog input 1 is 1.5V due to change of internal resistance of NTC Thermistor with the change of

temperature. So, Logo PLC (OBA5) is given command through ladder diagram to start the boiler system when voltage in analog input 1 is 1.5V. Therefore, boiler system starts running when the voltage in analog input 1 is 1.5V i.e. the temperature is 30°C.At 96°C the voltage at analog input 1 is 4.05V due to change of internal resistance of NTC Thermistor with the change of temperature. So, Logo PLC (OBA5) is given command through ladder diagram to shut down the boiler system when voltage in analog input 1 is 4.05V. Therefore, boiler system shuts down when the voltage in analog input 1 is 4.05V i.e. the temperature is 96°C.

Moreover, the boiler system can be run for certain period of time using the ON push button. It will run for certain period of time and then stop. The OFF push button is used to stop theoperation of the system. And to stop the boiler system manually, a manual switch is used. Circuit diagram of the project is shown in "Fig.8".

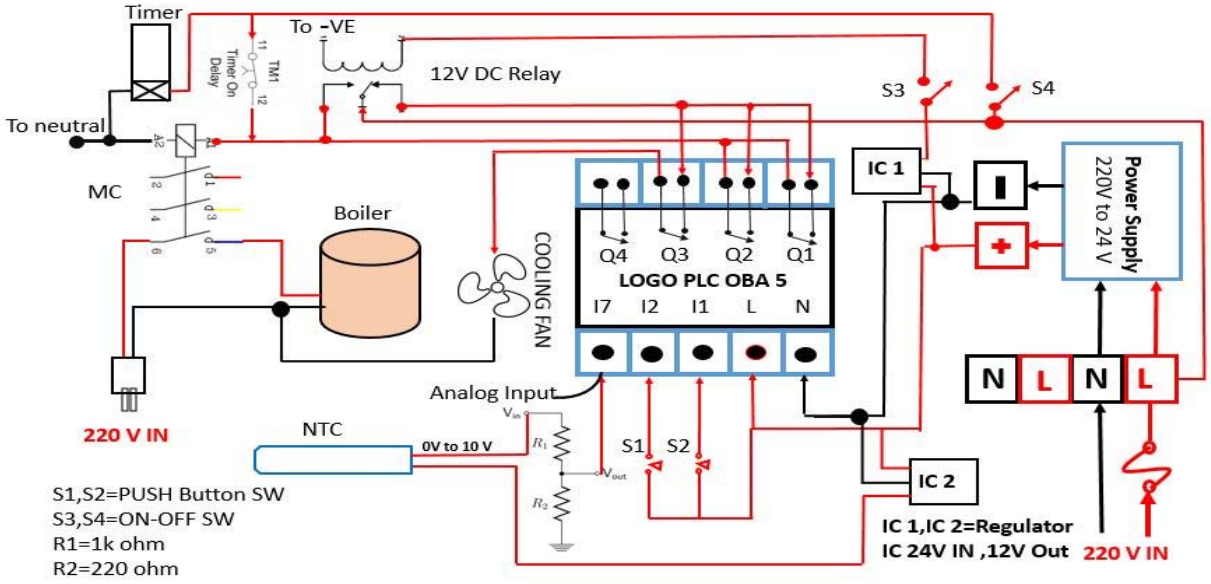


Fig. 8. Circuit diagram of the system.

IV. Programming and Implementation

This section illustrates the algorithm and practical implementation of our system.

A. Algorithm

"Fig.9." and "Fig.10." shows the flow chart of the system. At the beginning of the operation as soon as the main switch is turned on boiler system starts thus the temperature of the boiler starts increasing. And the NTC thermistor starts measuring the temperature of the boiler. The measured temperature is displayed on a digital thermostat temperature control module. When boiler temperature exceeds the set maximum value of temperature the boiler system shuts down. Then the cooling fan start running until the temperature is reduced to certain limit. The boiler system remains off till the temperature gets below the set minimum value of temperature. As soon as the temperature gets below the set minimum value of temperature the boiler system starts again. This operation is continued and maintained. And in case of timing operation we can use the ON push button to start the boiler system and run it for a certain period of time according to our need. After the certain period of time the boiler system will stop. The OFF push button can be used if we want to stop the boiler system during the timing operation and before the certain period of time. In both the ways the entire operation is controlled by Programmable Logic Controller. The programming of PLC is done using Ladder Diagram (LD).

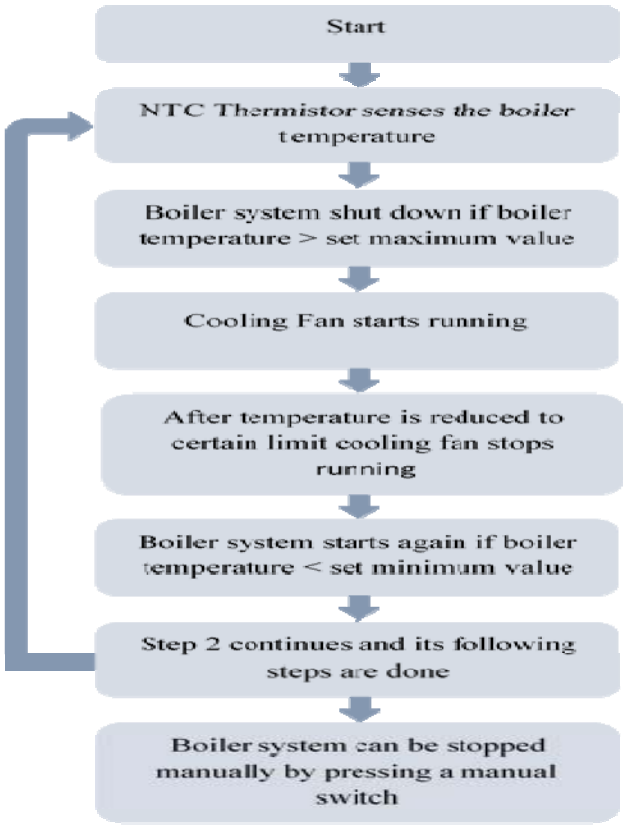


Fig. 9. Flow chart of the system (Normal operation)

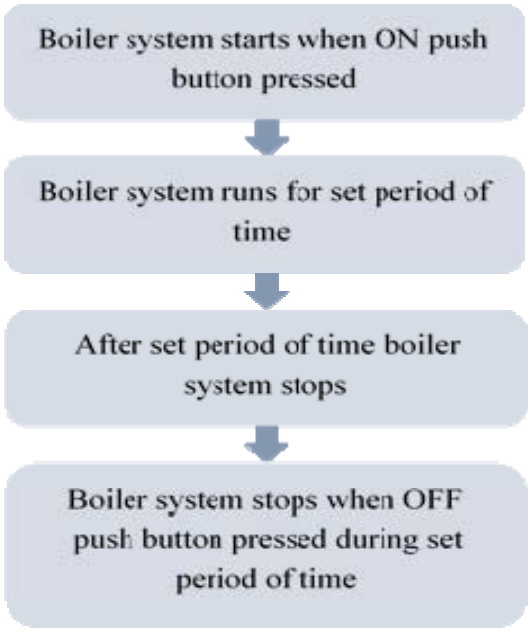


Fig. 10. Flow Chart of the system (Timer operation)

The flow chart of the sequential PLC operation for normal and timing both the operation is shown in "Fig.11." and "Fig.12." respectively.

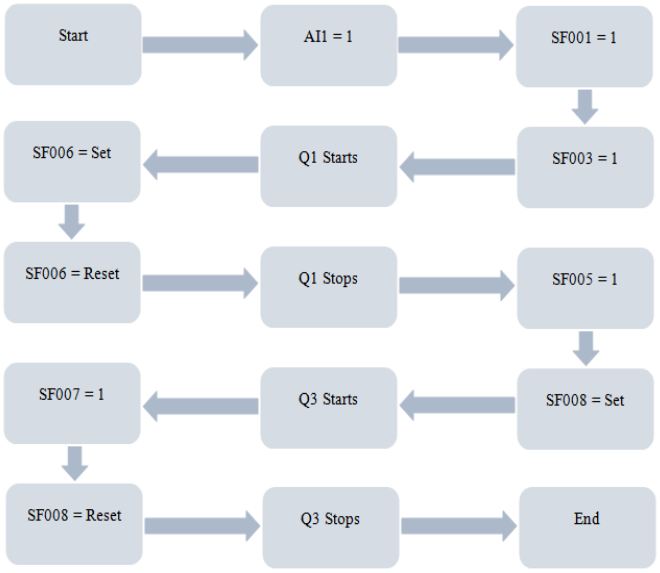


Fig. 11. Flow chart of PLC (Normal operation)

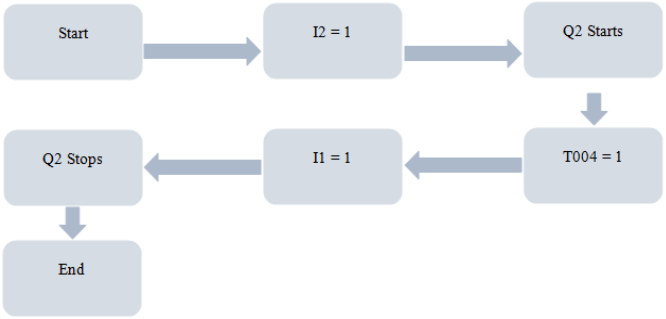


Fig. 12. Flow chart of PLC (Timer operation)

B. Hardware Implementation and Results

The overall performance of the boiler automation system was satisfactory but some improvement can be done on this system. The switching operation of this system was fast and accurate.The overall performance of this system is shown in below:

When the temperature is equal or above 30^0C as well as the voltage in LOGO PLC analog input (AI1) is 1.5V the boiler system starts running. The running condition of boiler system is shown in "Fig.13."

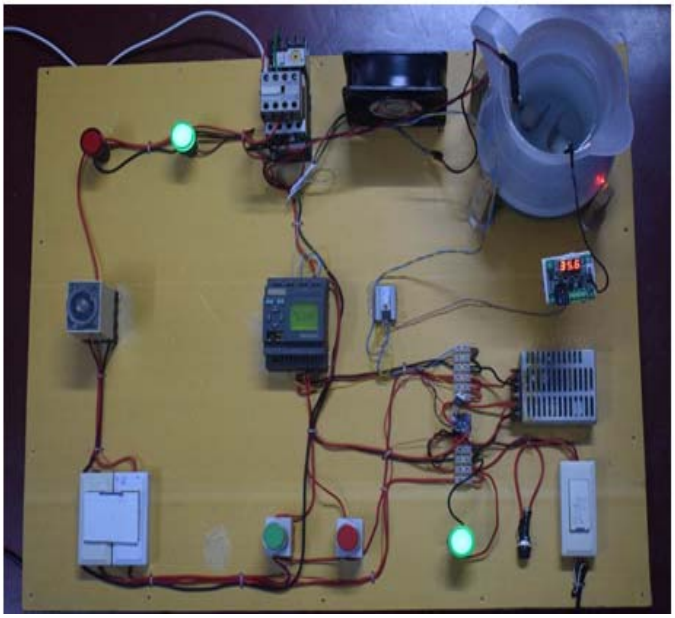


Fig. 13. Running condition of boiler system.

When the temperature is equal or above 96^0C as well as the voltage in LOGO PLC analog input (AI1) is 1.5V the boiler system will be shut down. The shut down system is shown in "Fig.14."



Fig.14. Shut down system of boiler system.

In order to stop the overall boiler system, a manual switch is used. When the manual switch is pressed the whole boiler system is shown in "Fig.15."

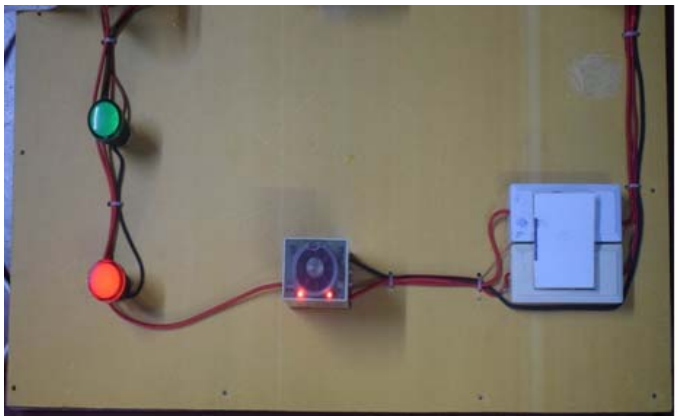


Fig.15. Boiler system switching off manually

C. Comparison with Real System

The system is developed primarily for research purpose as we design the prototype here. Real system may need some modification to be done. In this project heater jug is used as prototype of the boiler so during heating there is risk of hot water pouring out of the jug. But in practical industrial purpose real boiler is used which is usually closed vessel. So, there is less risk in the real system. In the prototype system, AC to DC and several voltage conversion is needed for operation purpose of the project. So, several extra devices are needed. However, in practical system these conversions may not be needed. Therefore, extra devices are not needed and thus the cost is reduced. In prototype system only the temperature of one boiler is controlled using single PLC. But in practical system temperature of several boilers can be controlled just by a single PLC.

V. CONCLUSION

To consider the necessities and demand of the industrial sector, there has been given significance attraction to boiler automation system which is done in this paper by using PLC. This system overcomes the lacking of the microcontroller.It also assures protection and safety to the human workers by automating the temperature control of boiler as it is risky due to high temperature during operation.Although the cost of this system is higher but the response rate of this system is faster and the operation of this system is more flexible compared to others.This system can be used to digitalize factories by making manually operated boilers fully automated. In future the system can be developed with considering the application-oriented implementation of remote monitoring of boiler automation by SCADA internet access.

References

[1] What is PLC, 1 August 2018, retrieved from https://www.engineersgarage.com/articles/plc-programmable-logic-controller.

[2] T. Karuppiah and Dr. Azha. Periasamy "Embedded system based industrial power plant boiler automation using microcontroller and GSM technology," International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 8, August 2013.

[3] Udaykumar S. Kulkarni, Venkat N. Ghodke, "An adaptive industrial boiler automation application using FPGA & GSM" *International Journal of Applied Research 2015; 1(8): 01-04*

[4] Automatic control of boiler system using distributed control system, 1 August 2018, retrieved from https://www.researchgate.net/publications.

[5] Working of Programmable logic controller, 3 August 2018, retrieved from https://www.engineersgarage.com/articles/plc-programmable-logic-controller.

[6] NTC Thermistor, 3 August 2018, retrieved from https://www.popularmechanics.com/home/interior-projects/how-to/a153/1275141/.

[7] Switched Mode Power Supply, 6 August 2018, retrieved from https://www.efxkits.us/switch-mode-power-supply-smps-works/.

[8] S.Kalaivani, M.Jagadeeswari" PLC & SCADA Based Effective Boiler Automation System for Thermal Power Plant" *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 4, April 2015,*

[9] S. S. Dhavale, T. H. Salunke, S. R. Ranaware, H. M. Sakhare, A. S. Bodhe," Effective Boiler Automation System" International Journal of Advanced Research in Electrical,Electronics and Instrumentation Engineering

[10] A Brief History of Automation, 1 August 2018, retrieved from https://www.phcfirst.com/words-in-motion/2016/9/7/a-brief-history-of-automation.

[11] Function of Magnetic Contactor, 6 August 2018, retrieved from https://sciencing.com/functions-magnetic-contactor-5598543.html.

[12] Introduction to Boiler, 3 August 2018, retrieved from http://www.boilers.guide/key-components-of-boilers/.

[13] Characteristics of NTC Thermistor, 3 August 2018, retrieved from http://www.resistorguide.com/ntc-thermistor/.

[14] Digital Thermostat Temperature Control Module, 5 August 2018, retrieved from http://www.communica.co.za/catalog/Details/P0402544492.

[15] Regulator IC 7812, 5 August, 2018, retrieved from http://qqtrading.com.my/7812-12v-voltage-regulator-linear-ic.

[16] DC Relay Switch, 5 August 2018, retrieved from http://www.theorycircuit.com/dc-relay-switch/.

[17] Working of Relays, 5 August 2018, retrieved from http://www.circuitstoday.com/working-of-relays.

Face Recognition System Based on Raspberry Pi Platform

Nafis Mustakim
Department of Electrical
and Electronic Engineering
Shahjalal University of
Science and Technology,
Sylhet-3114, Bangladesh
Email: nafismustakim@gmail.com

Noushad Hossain
Department of Electrical
and Electronic Engineering
Shahjalal University of
Science and Technology,
Sylhet-3114, Bangladesh
Email: sourav1023@gmail.com

Mohammad Mustafizur Rahman
Department of Electrical
and Electronic Engineering
Shahjalal University of
Science and Technology,
Sylhet-3114, Bangladesh
Email: mustafizmr07@gmail.com

Nadimul Islam
Department of Electrical
and Electronic Engineering
Shahjalal University of
Science and Technology,
Sylhet-3114, Bangladesh
Email: Hypernadim@gmail.com

Zayed Hossain Sayem
Department of Electrical
and Electronic Engineering
Shahjalal University of
Science and Technology,
Sylhet-3114, Bangladesh
Email: zhsayem66@gmail.com

Md. Asaduz Zaman Mamun
Department of Electrical
and Electronic Engineering
Shahjalal University of
Science and Technology,
Sylhet-3114, Bangladesh
Email: asadeee11@gmail.com

***Abstract*—Human face recognition plays an important role in video surveillance, human-computer interface, personalizing different applications. In this paper, we present an approach to detect and identify a human face from the real-time video that tracks a face and compares it with stored data of known individuals. Our approach recognizes an individual within a fraction of a second which completely ignores any background effect. It also shows additional information about that individual. Besides, this method also works on different lighting conditions which make it suitable to execute its purpose in a wide variety of environment without encountering any significant error. We have implemented our work using a Raspberry Pi. And this approach utilizes the benefits offered by different python libraries like OpenCV and numPy. SQLite database management system is used as the backbone of its memory. Through its faster and authentic performance, this system is worthy enough to be integrated with any applications to boost up automation and user-friendliness.**

***Index Terms*—— Real-time Face Detection, OpenCV, NumPY, SQlite, Gray-scaling, Raspberry pi**

I. INTRODUCTION

Human identification is a process to identify a person based on one or more unique features of that person. Many wide varieties of personal verification procedure are available in the commercial and law enforcing field. The most common personal verification method is Password identification number (PIN) or password system. But these methods are vulnerable to forgery, theft, and lapses in the user's memory. For these reasons, we are focusing on verifying a person through his/her unique biometric features [1]. These features include fingerprint, palm, iris, retina, and face detection if we want to identify the person using his unique biological structure. Human activity is a major concern in a wide variety of activities such as human surveillance, human-computer interface, [2], [3], and face recognition based database management [4]. Thirty years ago, face recognition was considered as one of the hardest in artificial intelligence technique. But through a series of successes in recent years, it is now not only technically feasible but also economically practical. In short, we can say that nowadays face recognition is a popular identification technology which has numerous commercial and law enforcement applications.

In recent years, application of image recognition in different parts of our life has increased many-fold. A foolproof security system solution which includes the sub-system such as surveillance CCTV, video management and wireless backbone has been developed by A. Mike *et al.* in their paper named "Face Recognition in Poor-Quality Video: Evidence From Security Surveillance" [5]. The System is capable of identifying an intruder within a restricted area covered by CCTV cameras. Usually, we need to see the video for a long period of time to inspect the intrusion. But the use of face recognition in the security system will show us the parts of the video where intrusion took place. ATM booths become more secured when face recognition is used. The face recognizing evaluation with exceptional occlusion handling (EOH) is feasible in most realistic environment [6]. Besides, Face recognition approach is also handy to count the number of people in a place by using the frame difference to detect the rough edges of moving people and then using the chromatic feature to locate the people's face [7]. The attendance system of educational institutes become more efficient and easy if face recognition is used to take attendance. Using personal component analysis

(PCA) in face recognition makes it much more feasible to record the attendance of a student. It also provides the facility to the faculty to access the information about the attendance of students easily by maintaining a log for clock-in and clock-out time [8]. Hospitals and medical centers can be widely benefited by using a face recognition system. In many cases hospitals receive patients who are unconscious. Facial recognition offers a faster way to access medical information of patients and the process of providing fast medical care can be accelerated. Even, to keep the track of newborn babies face recognition can be used [9]. Goods kept in the inventory of an industry having an image recognition inventory management system are safer than the conventional approach [10].

II. SYSTEM ARCHITECTURE

Raspberry Pi is the microprocessor used in this system. And we have used python for coding. Open source libraries of python like OpenCV and NumPY are used to process the image recognition algorithm. SQLite is used to store data of the desired people. The whole working process of the system can be best explained with three segments. These are:

Data segment: This segment of the python code is mainly dedicated to recording a person's identification data. Basically, it instructs the camera to take 20 sample pictures. And users need to provide information about their Name, age, gender, and department. This information is stored in SQLite database along with the 20 sample pictures. When the data of a person matches with any previous record, his/her sample pictures are updated. This feature makes our system adaptive to small changes in a person's facial pattern. The work flow diagram of data segment is shown in Fig.1.

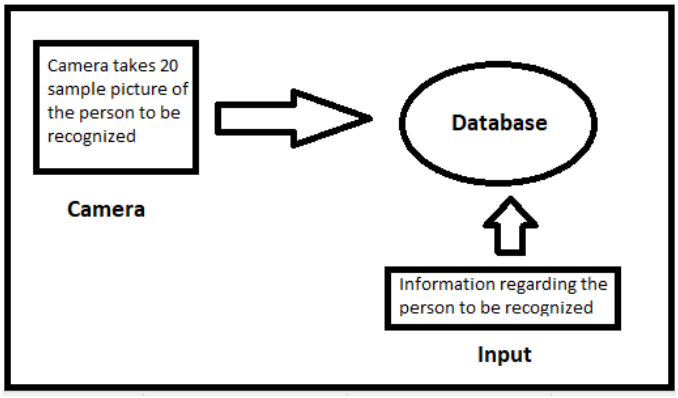


Fig. 1: Data Segment

No.	ID	Name	Gender	Age	Department
1	15	Noushad Hossain	Male	23	EEE
2	13	Mustafiz	Male	23	EEE
3	16	Saem	Male	22	EEE
4	12	Taskin	Male	23	EEE

Fig. 2: Database Table

The database of Fig.2 is made using SQL language. The ID and Age attributes of the database are integer type. Gender is basically text type and Department name is a string type attribute.

Trainer segment: This segment mainly trains the system to be able to identify specific persons. The sample pictures of a person are converted into the grayscale format. This system does not need the color explicitly. Besides, information regarding the color of an image sometimes makes it difficult to identify important edges. Therefore color in an image can be considered as noise in our study. Hence, grayscaling is done. So, Grayscaling makes the image processing algorithm more efficient and faster [11]. When an RGB image is converted into a grayscale format the 3D pixel values are converted into 1D values. As a result, we will need to deal with less data if grayscaling is used. Fig.3 shows the grayscale format of an RGB image.



Fig. 3: Conversion to grayscale

Detection segment: The camera feeds the microprocessor with real time video. The video is imported to the python code of the microprocessor as a two dimensional matrix. NumPy is a open source library of python which makes it easier to process the multidimensional matrix [12].

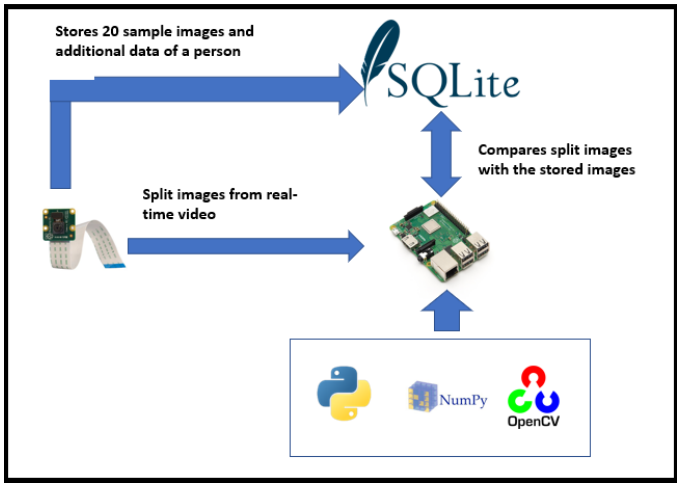


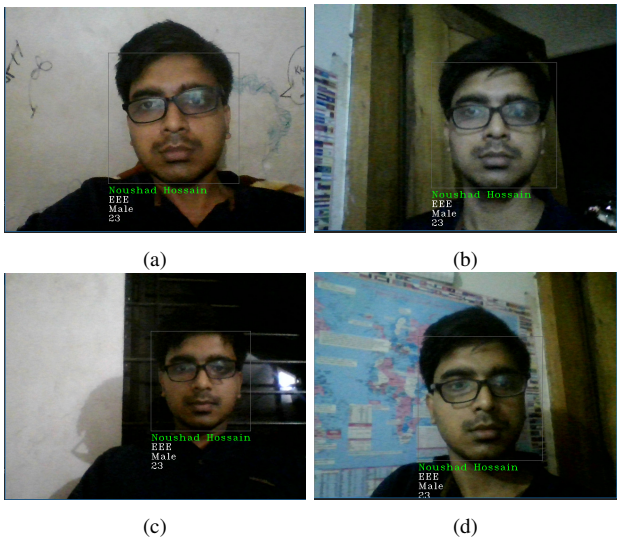
Fig. 4: System Architecture

First, the part containing the face is extracted from the full image to reduce the effect of background in image processing and stored in a matrix. Then the matrix is manipulated and compared with the sample photos for identifying any match. If any match is found, the real time video display shows the name and additional information of the identified person with a

green square mark tracking the recognized face of that person. The additional information is taken from SQLite database management system. The system architecture is illustrated in Fig.4.

III. System Performance

Zero background effect: In most cases, the face recognition approach is usually favorable for the place where previous data is taken. As a result, it does not function authentically when a test is run in a new environment. This happens when the total picture is compared with the stored data. But there is no effect of background on the performance of this system. Because the algorithm extracts the face of the person from the images stored in the database. Therefore, it works perfectly even though the background changes. In Fig.5 we see that an individual's face is being recognized though he is changing places.



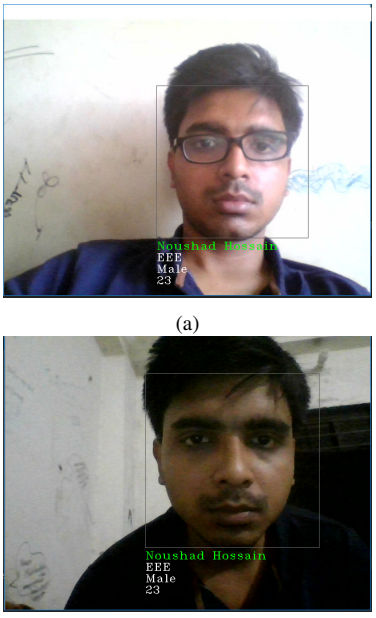
(a) (b)

(c) (d)

Fig. 5: Face recognized in four different backgrounds

Insensitive to small changes: Usually small beard, make-up, effect of light in a room has a small effect on a human face. So, gradually the performance of face recognition algorithm declines. It might work satisfactorily one day but will not be consistent for a longer period of time. But this system is smart enough to ignore these small changes and detect the human perfectly. The algorithm also checks the structure of an individual and compares the edge of different parts of a face. The system is capable of identifying a person in both natural and artificial lighting conditions. Again, if a sample picture of a person is taken where he had spectacles. But the person is recognized without any error when he had no spectacles. This is because the algorithm of the system checks the different structures of the face for finding a match. Fig.6 illustrates the insensitivity of the system to these small changes.

Fast recognition: Real-time face recognition system needs to have faster and efficient processing speed. This system is fast enough to complete the whole recognition



(a)

(b)

Fig. 6: Insensitivity of the system to small changes.A person is recognized a) With spectacles in a bright day light. b) Without spectacles in artificial lighting

process minimum three times in a second. At best the detection cycle is completed five times in a second. Fig.7 shows the time taken to complete each detection cycle. Our study shows, on average four detection cycle is completed in a single second. As a result, image recognition is not hampered even if the person is in motion. There are several reasons behind the faster recognition approach. A PNG (Portable Network Graphic) file is much more difficult and time-consuming to process than XML (Extensible Markup Language) file. For this reason, we have used an XML format to save the images in the database. The algorithm is designed to extract the face of a person only. So, other information of the images except the face is not processed and hence the image processing time is decreased. Grayscaling also plays a vital role to reduce the processing time. As a result, identifying a moving person is much easier. On the contrary, a slow image recognition algorithm fails to identify a moving person.

IV. CONCLUSION

In short, we can say that The system is built on Raspberry Pi which is the brain of the system. Python is used in this system for its open source library facilities. The open source librariy we used is OpenCV which mainly aim to process real-time computer vision. The still image from the video is stored in the python as a two-dimensional array. NumPy which is a python library for processing multidimensional array is used to process the still image faster. The system has a storage management system which is intended to store the data of the persons of interest. SQLite, a popular database management system is used for data management purpose. The database contains information regarding the person to be

Take	Time
1	Friday, January 11, 03:42:54,2019
2	Friday, January 11, 03:42:54,2019
3	Friday, January 11, 03:42:54,2019
4	Friday, January 11, 03:42:55,2019
5	Friday, January 11, 03:42:55,2019
6	Friday, January 11, 03:42:55,2019
7	Friday, January 11, 03:42:55,2019
8	Friday, January 11, 03:42:56,2019
9	Friday, January 11, 03:42:56,2019
10	Friday, January 11, 03:42:56,2019
11	Friday, January 11, 03:42:56,2019
12	Friday, January 11, 03:42:57,2019
13	Friday, January 11, 03:42:57,2019
14	Friday, January 11, 03:42:57,2019
15	Friday, January 11, 03:42:57,2019
16	Friday, January 11, 03:42:57,2019
17	Friday, January 11, 03:42:58,2019
18	Friday, January 11, 03:42:58,2019
19	Friday, January 11, 03:42:58,2019
20	Friday, January 11, 03:42:58,2019
21	Friday, January 11, 03:42:59,2019
22	Friday, January 11, 03:42:59,2019
23	Friday, January 11, 03:42:59,2019
24	Friday, January 11, 03:42:59,2019
25	Friday, January 11, 03:43:00,2019
26	Friday, January 11, 03:43:00,2019
27	Friday, January 11, 03:43:00,2019
28	Friday, January 11, 03:43:00,2019
29	Friday, January 11, 03:43:01,2019
30	Friday, January 11, 03:43:01,2019
31	Friday, January 11, 03:43:01,2019
32	Friday, January 11, 03:43:01,2019
33	Friday, January 11, 03:43:01,2019

Fig. 7: Demonstration of fast recognition process

detected. The system is very helpful and efficient because there is no background effect for recognizing the person. It can recognize a person even though the environment and background changes. And it works satisfactorily in different lighting conditions. Small changes in a person's face like make-up, the spectacles have little effect on the recognition process. Moreover, the system is fast enough to process four frames per second. So, it can be integrated with any system to enhance automation, effectiveness, and user-friendliness of that system.

REFERENCES

[1] A. Pentland and T. Choudhury, "Face recognition for smart environments," *Computer*, vol. 33, no. 2, pp. 50–55, 2000.

[2] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 705–741, 1995.

[3] C. Liu and H. Wechsler, "Evolutionary pursuit and its application to face recognition," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 6, pp. 570–582, 2000.

[4] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 12, pp. 1349–1380, 2000.

[5] A. M. Burton, S. Wilson, M. Cowan, and V. Bruce, "Face recognition in poor-quality video: Evidence from security surveillance," *Psychological Science*, vol. 10, no. 3, pp. 243–248, 1999.

[6] S. Eum, J. K. Suhr, and J. Kim, "Face recognizability evaluation for atm applications with exceptional occlusion handling," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*. IEEE, 2011, pp. 82–89.

[7] T.-Y. Chen, C.-H. Chen, D.-J. Wang, and Y.-L. Kuo, "A people counting system based on face-detection," in *Genetic and Evolutionary Computing (ICGEC), 2010 Fourth International Conference on*. IEEE, 2010, pp. 699–702.

[8] N. Kar, M. K. Debbarma, A. Saha, and D. R. Pal, "Study of implementing automated attendance system using face recognition technique," *International Journal of computer and communication engineering*, vol. 1, no. 2, p. 100, 2012.

[9] S. Tiwari, A. Singh, and S. K. Singh, "Can face and soft-biometric traits assist in recognition of newborn?" in *Recent Advances in Information Technology (RAIT), 2012 1st International Conference on*. IEEE, 2012, pp. 74–79.

[10] M. Capazario and J. Rubin, "Image recognition inventory management system," Aug. 14 2003, uS Patent App. 10/246,384.

[11] C. R. Dyer and A. Rosenfeld, "Thinning algorithms for gray-scale pictures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 1, pp. 88–89, 1979.

[12] S. v. d. Walt, S. C. Colbert, and G. Varoquaux, "The numpy array: a structure for efficient numerical computation," *Computing in Science & Engineering*, vol. 13, no. 2, pp. 22–30, 2011.

Automatically Controlled Energy Conservation System for Corporate Office Based on Microcontroller

Koushik Ahmed[1], Joydip Paul[2], Md. Mominur Rahman[2], Abu Shufian[2], Md Shahnewaz Tanvir[1], and Md. Mohaimanul Islam Sagor[1]
[1]Department of Electrical and Electronic Engineering, Ahsanullah University of Science and Technology, Dhaka-1208, Bangladesh
[2]Department of Electrical and Electronic Engineering, American International University-Bangladesh, Dhaka-1229, Bangladesh
koushik5699@gmail.com, dipupaul52@yahoo.com, mominur18aiub@gmail.com, shufian2aiub@gmail.com, mstanvir56@gmail.com, and sagor373@gmail.com

***Abstract*—Energy conservation is one of the major issues of recent days because all the natural resources will end up soon as the earth has a limited amount of such non-renewable resources. This paper ensures the conservation of energy by minimizing the wastage of electricity. Two microcontrollers are used for this setup. One of them is used to control the electrical appliances automatically depending on human presence in a particular room or area. When the passive infrared sensor (PIR) and proximity sensor find the motion and presence of human inside the area then it will turn on the electrical appliances and will count the number of persons. And the opposite action will be taken for the absence of human movement in a particular area using a relay circuit in the system. Although everything will be controlled using this system there is also a scope to control the appliances manually using separate switches. Once the switches are kept ON, the system will automatically shut down and turn ON when necessary. Another microcontroller is used to operate the safety features of the building and the entrance gate opening which will automatically open or close the door when someone enters or leaves the room and will show it on the LCD. With the help of the humidity sensor, it will show the temperature and humidity level in the LCD. A smoke detector sensor is used which will ring the fire alarm when smoke is detected and it will open the emergency exit door as well as the entrance door.**

Keywords— energy conservation, microcontroller, motion detection, PIR sensor, proximity sensor, smoke detection.

I. INTRODUCTION

"Energy cannot be created or destroyed; it can only be transformed from one form to another." quoted by one of the greatest minds of the 20th century [1]. Energy is a major concern for our generations nowadays because everyone is in need of it. Non-renewable resources are being used up to suffice the demands of the people's needs. Non-renewable resources are such resources which are from sources which will run out in the near future e.g. coal, petroleum, fossil fuel, and natural gas. The burning of such resources results in the production of energy which is converted to electricity to suffice the demands of the world. These resources produce a huge amount of energy and they are readily available to some countries and are cheap [2] [3].

Conservation of energy had been a major issue since the development of technology because as the earth is getting older its natural resources are decreasing from time to time. So long ago the research on conservation of energy began, some conservations are listed as Use of efficient lighting- light emitting diode (LED), Use of energy efficient appliances- Compact Fluorescent Bulb, use of efficient computers- Laptop, Turning off electrical appliances- Disconnection of appliances after use [4] [5].

Several works have been done in recent years regarding energy conservation for reducing the wastage of energy in both residential and industrial areas. In recent years, researchers adapted a method where the system can turn on or off the electrical appliance by sensing and counting the presence of a person through IR sensor and visitor counter in a particular area and also can control the speed of the fan by sensing the temperature of the monitored area through temperature sensor [6]. A microcontroller-based wireless remote control system is applied to a protection system which is used for wave energy conversion [7]. The smoke detection and alarm system are executed by using a microcontroller-based fire extinguisher [8]. A low-cost energy management system can be developed by using campus card system for controlling the classroom lighting [9].

The proposed system is to establish an energy conserving society with the help of some common components found every day in our electric component shops. The project is based on microcontroller and using PIR and proximity sensors energy will be conserved by saving electricity.

II. METHODOLOGY AND MODELING

The most basic and eventual requirement in an energy conservation system, a connection between some components such as sensors, relay, corporate appliances like fans, lights etc. with microcontroller are important. There are various way to link the devices, user and the system and the linking depends on the system, range, size of a corporate office, ease of users. If a user wants to control all the corporate appliances through the system, then all the appliances need to interface with this system. If the user wants to control a few of the appliances, then only those appliances need to interface with the system and other appliances can be controlled manually.

The basic concept of this project is to turn on or turn off the corporate appliances like fans, lights etc. automatically. Here, relay operates as a switch by passing information in between microcontroller and appliances. PIR sensor and proximity sensor is connected with a microcontroller. Control devices are also connected with the microcontroller. Microcontroller receives information from human detection PIR sensor and proximity sensor. According to information received by the microcontroller, it turns on or turns off the corporate appliances automatically.

In Fig. 1. microcontroller 1 is connected to PIR sensor, proximity sensor, loads through relay, power supply, and LCD display. The proximity and PIR sensor detect the presence or absence of a human movement in a particular area. Then this data is sent to the microcontroller. After getting this data microcontroller analysis this data and based on this data

microcontroller controls the loads (light, fan) with the help of a relay.

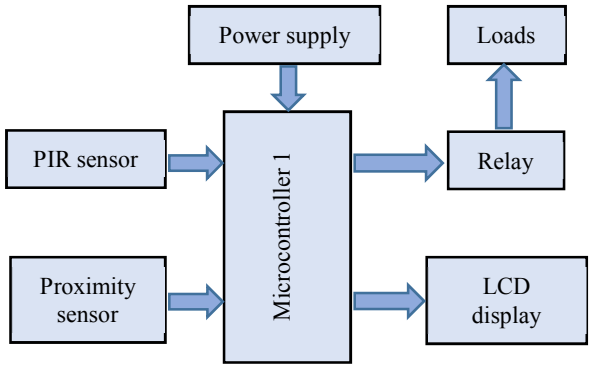


Fig. 1. Block diagram of the energy conservation system.

Microcontroller 2 in Fig. 2. is connected to the system which automatically controls the entrance gates, fire alarm, and humidity. There are three sensors connected to the microcontroller which are a smoke detector, IR LED and photodiode and humidity sensor.

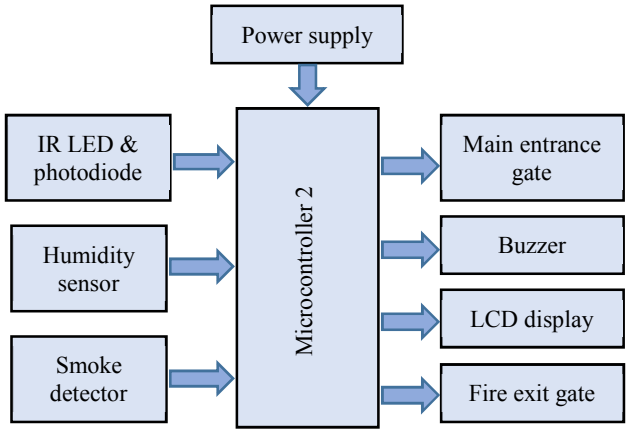


Fig. 2. Block diagram of the automated control system.

The smoke detector detects smoke and sends a signal to the microcontroller which triggers the buzzer and its ring indicates that there is smoke from the fire which is a safety measure required in a corporate building. Then the IR LEDs detect human figure and output is sent to the microcontroller which opens and closes the gates automatically receiving signals from the microcontroller. Lastly, the humidity sensor senses temperature level and humidity level and they are shown through an LCD which helps keep track of the temperature and humidity level of a room. All these sensors are controlled by the second microcontroller.

To design an energy conservation system for corporate appliances by using a microcontroller, PIR and proximity sensor some basics need to be considered. In the microcontroller, some programs are installed. It works with the programs are installed in it. Proximity sensor is used to detect any human entering or leaving the room. The PIR sensor is used to detect the movement of the human. If the sensor does not detect any movement more than 15 seconds then the loads will automatically turn OFF. Smoke detector is used to detect smoke and humidity detector measures the humidity.

III. Software Implementation

Simulation results of the total circuit are taken in different conditions using PROTEUS 8 Professional software in three different parts. The following components are used in simulation: microcontroller ATMEGA8, a battery, transistor BC547, capacitor, inductor 'ELJ-FA1R8KF', a diode, a motor, different LEDs, LCD 'LM016L', POT, a relay, resistors, AC input, MQ-135, HSM 20G, buzzer, and LOGICTOGGLE. As there is no way to use a PIR sensor in the simulation software, we can't vary this input in the simulation software. So, outputs are taken by varying only the count number of proximity sensor. These two conditions are described below.

A. Using only PIR sensor

1) When PIR sensor does not sense movement

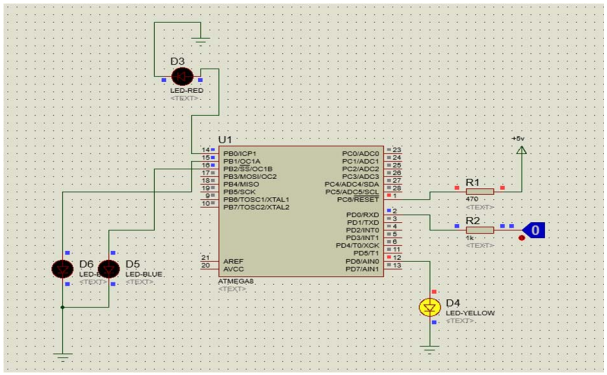


Fig. 3. Simulation result when PIR sensor does not sense any movement around it.

Here, PIR sensor input "0" indicates that it does not sense any movement up to 60 cm around it. As a result, the red LED is OFF and both the loads are also OFF. But for sensing the detection of movement, the yellow LED is blinking. Here input PIR sensor input is varied by toggling the input in the PIR sensor position.

2) When PIR sensor senses movement

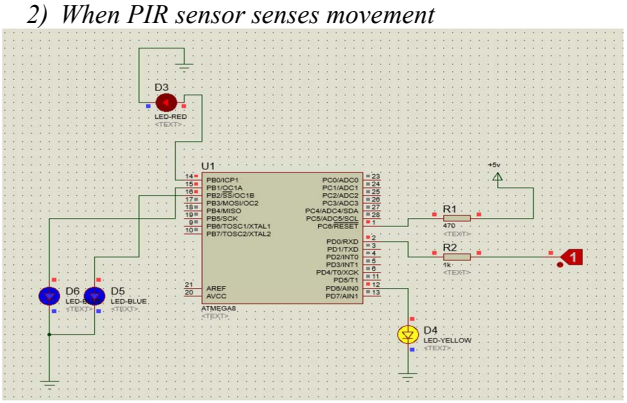


Fig. 4. Simulation result when PIR sensor senses movement around it.

Here, PIR sensor input "1" indicates that it senses movement around it. As a result, the red LED is ON and both the loads are also ON. The red LED at ON state indicates that there is movement around the PIR sensor. It will stay ON as long as there is movement. If the sensor suddenly does not sense movement, it will wait up to 15 seconds to detect the movement. During these 15 seconds, if it does not sense any movement, the loads will automatically turn OFF. But if it senses movement during these 15 seconds, it will keep the loads ON. By this method, PIR sensor will keep the loads ON and OFF in the room.

B. Using both PIR sensor and Proximity sensor

1) When the count number of proximity sensor is zero

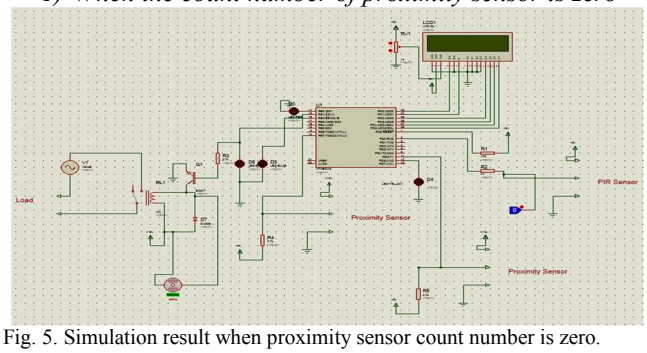


Fig. 5. Simulation result when proximity sensor count number is zero.

When proximity sensor input is zero, which indicates there is no person inside the room. As a result, both LEDs are off and the AC loads and red LED are also OFF. The fan also does not move. A relay is used in this case as AC loads are used. As no one is present inside the room, there is also no movement. For this purpose, PIR sensor input is kept at zero states.

2) When the count of proximity sensor is greater than zero but there is no movement

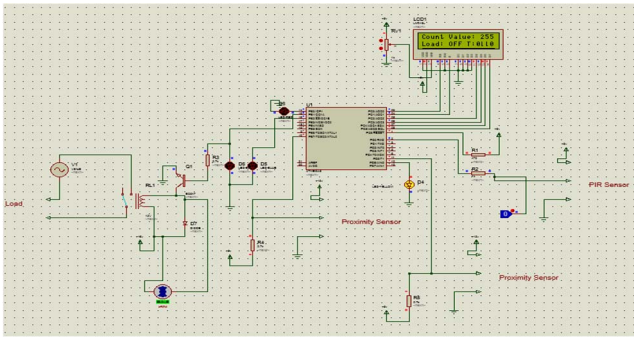


Fig. 6. Simulation result when proximity sensor count number is greater than zero without movement.

When proximity sensor input is greater than zero, which indicates there is at least one person in the room. But here, the PIR sensor input is zero. That means there is no movement inside the room. As a result, both the LEDs are OFF and the AC loads and red LED are also OFF. The fan is also not moving.

3) When the count of proximity sensor is greater than zero and there is movement

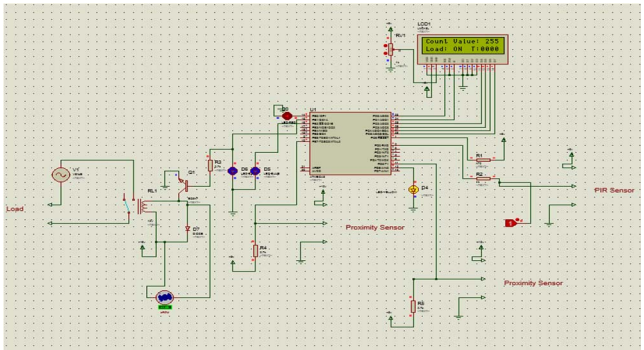


Fig. 7. Simulation result when proximity sensor count number is greater than zero with movement.

When proximity sensor input is more than zero and again there is movement inside the room that means there is at least one person in the room. Again, PIR sensor input "1" indicates there is movement inside the room. As a result, both the LEDs are ON and the AC loads and red LED are also ON. The fan is also moving.

C. With servo motor, fire detector, humidity sensor, IR LED and photodiode

1) The system before giving the power supply

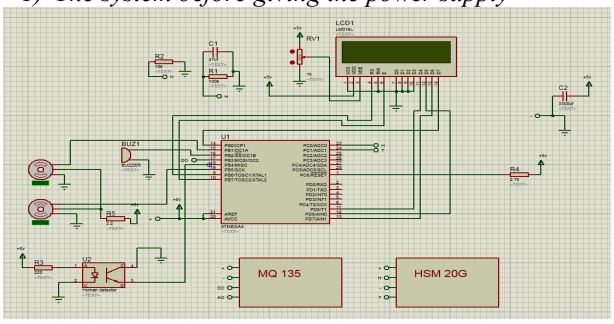


Fig. 8. Simulation result of the automated control system before giving the power supply.

Here, the second microcontroller is connected with different parts such as two servo motors, photodiode, IR LED, buzzer, smoke detector MQ-135, humidity sensor HSM-20G. The servo motors don't rotate as there is no human around IR LED and photodiode, which in this case is a human detector.

2) The system after fire smoke is sensed by the fire detector sensor

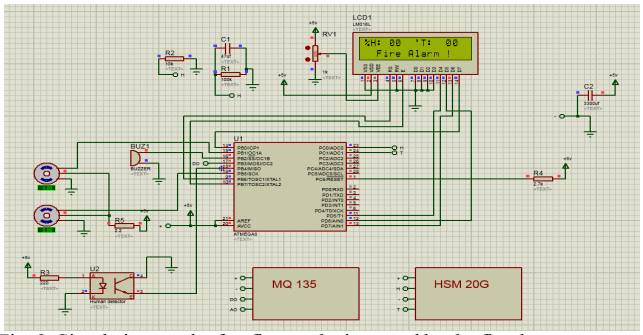


Fig. 9. Simulation result after fire smoke is sensed by the fire detector sensor.

When fire smoke is sensed by the fire detector sensor, the LCD shows the "FIRE ALARM" sign. As there is no scope to show humidity and temperature in the simulation, so it shows zero for both the cases. Here servo motor rotates when there is a presence of human around the human detector sensor which are actually IR LED and photodiode.

IV. Hardware Implementation and Result Analysis

In this section, hardware implementation, as well as result, have been discussed.

A. The initial setup of the system

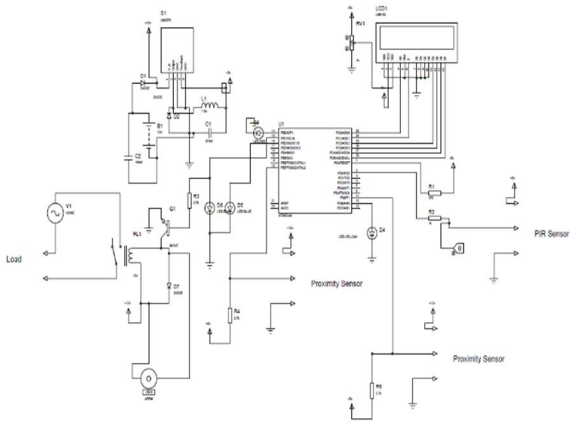


Fig. 10. The overall circuit diagram of Energy Conservation System.

This is the main circuit diagram of the power conservation system part. It includes PIR sensor, proximity sensor, LCD display, LEDs, loads and power supply unit.

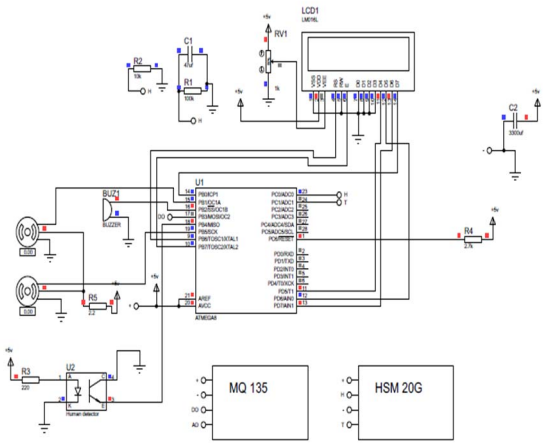


Fig. 11. The overall circuit diagram of the automated control system.

It is the circuit diagram of the automated including fire detector sensor, humidity sensor, human detector sensor, and LCD display.

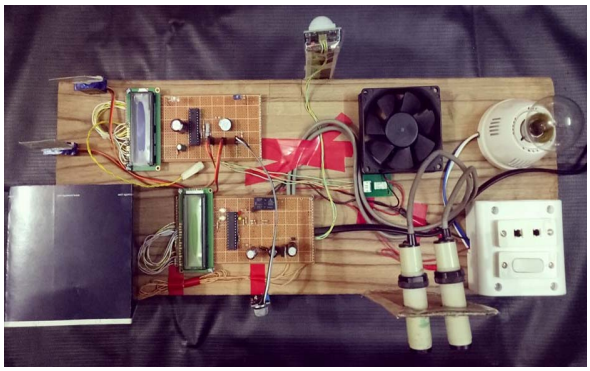


Fig. 12. Hardware implementation of Energy Conservation System.

This is the main hardware setup of our proposed system before turning ON the power supply. It consists of one PIR sensor, one proximity sensor, two microcontrollers, two LCD displays, transistors, capacitors, one motor which is used as a fan, different resistors, one relay, one power supply, different AC loads, two servo motors, IR LED, photodiode, buzzer, smoke detector MQ-135, humidity sensor HSM-206 and a switch to manually control a load.

B. The system after turning ON the power supply

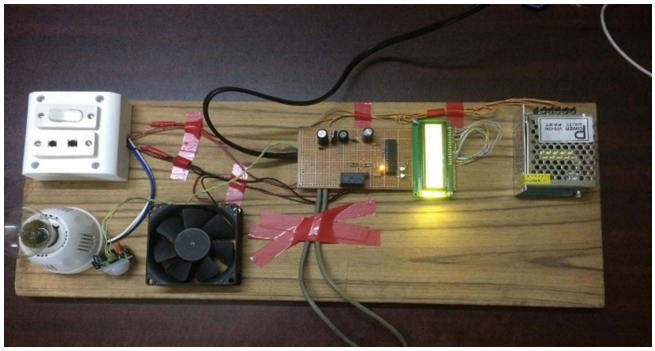


Fig. 13. The setup after turning ON the power supply.



Fig. 14. LCD display after turning ON the power supply.

This is the setup after turning ON the power supply. The LCD display is ON and yellow LED in blinking to detect any movement near the PIR sensor. The LCD display is shown in the second picture. The count value is zero. As a result, both the LEDs and AC loads are OFF.

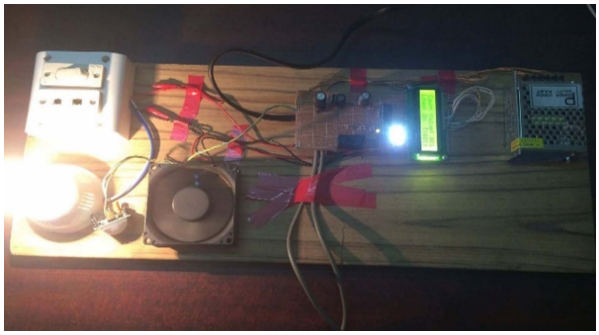


Fig. 15. The setup when the count for proximity sensor is greater than zero.



Fig. 16. The LCD display when the count for proximity sensor is greater than zero.

Here, the count value for proximity sensor is 3. As a result, both the LEDs are ON, the AC loads are also ON. This is applicable for any count value more than zero. But, when the count value is more than zero and still there is no movement around the 60 cm of PIR sensor for more than 15 seconds, the white LEDs and the AC loads will turn OFF. This is not shown here because it was not possible to capture the picture by not moving about 15 seconds. But the situation can be observed.

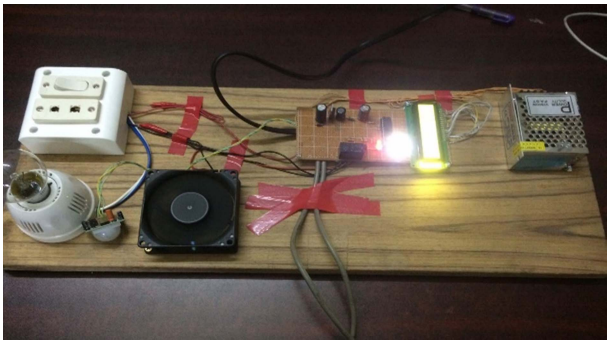


Fig. 17. The system after manually turning OFF the bulb.

The AC loads can also be controlled manually by using switches. In the above picture, the count value of proximity sensor is more than zero. As a result, both the white LEDs are ON. But among the AC loads, the only motor is ON, the bulb is OFF. A switch is used to control the bulb manually. So, in

spite of there is greater than zero counts in the proximity sensor and there is movement around the PIR sensor, the bulb is OFF. It will work according to the system only when the switch is turned ON. This manual setup can also be done for other loads but separate switches are needed for this case.

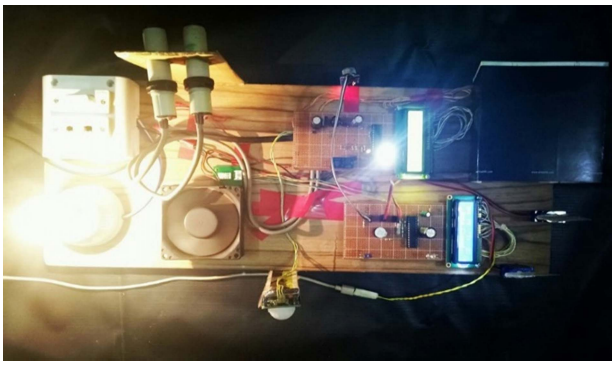


Fig. 18. The system after adding the automated control part.

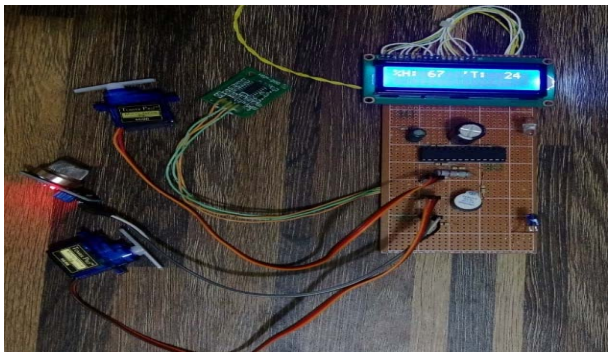


Fig. 19. The system when any person enters into the room.



Fig. 20. The LCD display when any person enters into the room.



Fig. 21. The LCD display after any person enters into the room.

There is a photodiode and IR LED in the entrance of the system. When someone passes these 2 devices the servo motor 1 with rotate 90 degrees that means the entrance gate will open instantaneously. The LCD display shows “Gate Open”. After that, the servo motor 1 will rotate to the opposite direction which indicates the door will be closed. The LCD display shows “Gate Closed”. But at that time servo motor 2 does not rotate.



Fig. 22. The system showing temperature and humidity.

The humidity sensor senses the humidity of the environment and shows the temperature and humidity in the LCD display.



Fig. 23. The system after there is smoke around it.

When the fire detector sensor senses smoke around it, both the servo motor rotates 90 degree that means both the door will open instantaneously. The LCD display shows “Fire Alarm!”.

V. Discussion

Such automated systems are encouraged nowadays to reduce human efforts and conserve energy but there are also many limitations. This system will work for a small space because the PIR sensors and proximity sensors that are used are not very strong. Therefore, for a large space, more than one system will be needed to ensure proper working rate. When an electrical appliance is manually turned off, it will not be turned on automatically until the switch is manually turned on. Several initiatives can be taken to improve the proposed system. The designed system may be implemented wirelessly. The circuit may be reorganized to get better output. Fewer components may be used to make the system more cost-effective.

VI. Conclusion

The system is built to conserve the energy which is being wasted by people while there are scopes to save the energy because the energy is produced by burning the natural resources which are vanishing with time. The system offers the loads to be controlled automatically and also manually which ensures energy conservation. Loads are connected to a relay system which is controlled by a microcontroller depending on inputs from the sensors that are used to detect motion. The system can be used to count the number of persons entered and left the room in an office. It consumes a very small amount of electricity to operate which is also cost effective. The loads connected can be operated automatically and if needed also manually depending on the operator's demand. The system is not expensive to build and consuming less energy proves it to be cost-effective and also eco-friendly. Although it has some drawbacks such as the PIR sensor is weak and will not be able to detect motion very far away from it. Secondly, if the loads are to be controlled individually then for every load there has to be a separate relay system connected to them individually. One relay system cannot control more than one load differently. By using more powerful sensors the system can be used to cover a large area which will increase the system cost by a little amount however it will not affect the demand of the system.

References

[1] "Energy Transfer Diagrams and Efficiency - Pass My Exams: GSCE Physics", Passmyexams.co.uk, 2019. [Online]. Available: http://www.passmyexams.co.uk/GCSE/physics/energy-transfer.html?fbclid=IwAR2YrXUft_8BsR-BrXna35Quuh1n90DxEpnol8m_dqH7X22ubrL6FULP8Oc. [Accessed: 02- Mar- 2019].

[2] "Coal and Air Pollution", Union of Concerned Scientists, 2019. [Online]. Available: https://www.ucsusa.org/clean-energy/coal-and-other-fossil-fuels/coal-air-pollution. [Accessed: 02- Mar- 2019].

[3] E. Institute, "Fossil Fuel Use Pushes Carbon Dioxide Emissions into Dangerous Territory", The Permaculture Research Institute, 2019. [Online]. Available: https://permaculturenews.org/2013/07/25/fossil-fuel-use-pushes-carbon-dioxide-emissions-into-dangerous-territory/. [Accessed: 02- Mar- 2019].

[4] B. Roisin, M. Bodart, A. Deneyer and P. D'Herdt, "Lighting energy savings in offices using different control systems and their real consumption", Energy and Buildings, vol. 40, no. 4, pp. 514-523, 2008. Available: 10.1016/j.enbuild.2007.04.006.

[5] "This course will provide participants with the skills and knowledge required to disconnect and reconnect appliances connected to low voltage installation wiring.", Plumber.com.au, 2019. [Online]. Available: https://plumber.com.au/training-development/disconnect-reconnect-composite-appliances-connected-to-a-low-voltage-installation-wiring. [Accessed: 02- Mar- 2019].

[6] M. Majumdar, H. Biswas, M. Shaim and K. Ahmmed, "Automated energy saving and safety system", 2014 International Conference on Electrical Engineering and Information & Communication Technology, 2014. Available: 10.1109/iceeict.2014.6919097.

[7] O. Farrok, M. R. Habib, M. R. Afefin and A. Habib "A new protection system of linear generators used for converting oceanic wave energy into electricity," IEEE Region 10 Humanitarian Technol. Conf., Dhaka, Bangladesh, Dec. 21–23, 2017, pp. 755–758.

[8] M. R. Habib, K. Ahmed, N. Khan, M. R. Kiran, A. K. M. Asif, M. I. Bhuiyan, and O. Farrok, "Quick fire sensing and extinguishing by using an arduino based fire protection device," Int. Conf. Comput., Commun., Chemical, Mater. Electron. Eng., Jul. 11-12, 2019 (Submitted).

[9] A. Jabeen, D. M. Kumar, "Automatic classroom lighting controller and energy saving based on microcontroller unit", International Advanced Research Journal in Science, Engineering and Technology, ISO 3297:2007, Vol. 3, Issue 7, July 2016.

Affordable Smart ECG Monitoring Using Arduino & Bluetooth Module

Mohammed Mahmudur Rahman
Department of Computer Science and Engineering
International Islamic University Chittagong
Chittagong-4318, Bangladesh
mmr@cse.iiuc.ac.bd

Md. Azizul Hoque Rimon
Department of Computer Science and Engineering
International Islamic University Chittagong
Chittagong-4318, Bangladesh
mdazizulhoquerimon@gmail.com

Muhammad Armanul Hoque
Department of Computer Science and Engineering
International Islamic University Chittagong
Chittagong-4318, Bangladesh
armnakl@gmail.com

Md. Redwan Sammir
Department of Electrical and Electronic Engineering
Chittagong University of Engineering and Technology
Chittagong-4330, Bangladesh
sammir.eee.cuet@gmail.com

***Abstract*— This project uses some sensors and Arduino to design and implement the Affordable Smart ECG Monitoring System. ECG monitoring is now becoming part of everyday life. Through ECG characteristics such as patient's heartbeats, heart conditions, and heart disease can be analyzed. The proposed system sensed the patient's ECG through 3 lead electrode system via AD8232 which amplifies minor and small bio-signals to the Arduino which processes them. This sensor was used to measure the heart's electrical activity called ECG. In this project, Arduino and Bluetooth module are used to collect and transfer the heart activity and send it to the mobile application. The current portable ECG technologies are costly so that anyone can't use that on their home for an emergency if they want and this new technology is less expensive.**

Keywords*— *ECG Monitoring system; Portable; Arduino; Bluetooth;

I. INTRODUCTION

In many cases, the patient requires continuous monitoring and it needs a long time to stay in hospitals which is expensive now a days. Traditional monitoring system enables monitoring of vital parameters requiring the sensors to be connected to bedside machines and the patient is essentially confined to bed. However, busy in today's world and the increase in sudden death events motivate a monitoring system that monitors patients remotely located on a continuous basis. With the technological advancement, it has become possible to design low-cost, home-based healthcare monitoring system in which the system records and displays the signals from the human body and transmits to any other location. Different studies have been carried out in connection with the development of remote healthcare systems, particularly heart rate monitoring systems.

II. LITERATURE REVIEW

UMHMSE introduced a real-time mobile healthcare system for monitoring elderly patients indoors or outdoors [6]. The main components of the system are a bio - signal sensor and a smartphone. The data that the bio - signal sensor collects is transmitted via the GPRS / UMTS network to an intelligent server. The system can monitor the elderly patient's mobility, location, and vital signs from a remote location.

In the Wireless Sensor Network Wearable Smart Shirt[7], a smart shirt has been designed. In order to monitor the continuous and real - time health of a patient, this shirt can measure electrocardiogram (ECG) and acceleration signals. The shirt consists mainly of sensors and conductive fabrics for the body signal to be received. The measured body signals are transmitted via IEEE 802.15.4 network to a base station and server PC. The wearable devices have low power consumption and are small enough to fit in a shirt. An adaptive filtering method was also proposed in this work to reduce the noise associated with the ECG signal.

While mobile devices are always considered a promising tool for monitoring and managing patients' own state of health, computational or data-intensive activities have some inherent limitations on these appliances. In Enabling Smart Personalized Healthcare[12], a new hybrid mobile computing solution has been proposed to overcome these limitations. A mobile cloud-based electrocardiograph monitoring system was implemented by the authors. The experimental results show that the proposed system can significantly improve the diagnostic accuracy, performance efficiency, and energy efficiency of the conventional mobile-based medical monitoring system.

In the ECG Remote Monitoring System and Human Body Temperature Signals[14], a remote healthcare system for monitoring electro-cardio graphic and temperature data has been described. The system consists of three modules namely (i) a hardware module, (ii) Bluetooth module, and (iii) display module. For data acquisition, the hardware module is used. As well for data transmission, the Bluetooth module is used. Finally, the display module displays the data. Using GPRS or Wi-Fi, the acquired clinical data is sent to a database server. The system performance has been tested on various patients

and the proposed change has been found to be very helpful to the physicians.

III. Methodology

A simple architecture has built here to monitor the Electrocardiogram. Here, Microcontroller is working as the core part of the whole system. Here the AD8232 sensor getting the analogue signal and send it to the microcontroller. Microcontroller converts it into digital data. Then holding the digital data, the microcontroller sends it to the android device via Bluetooth module.

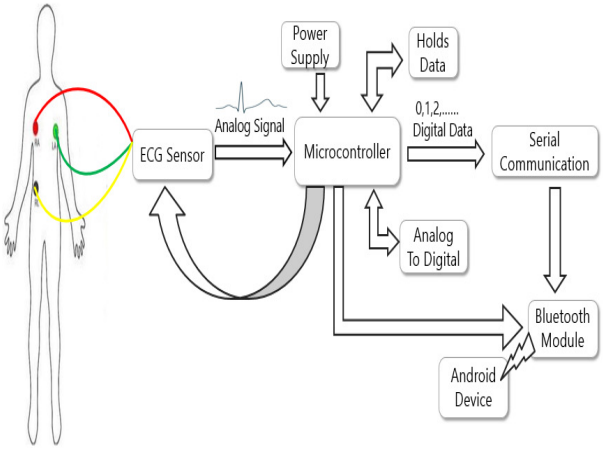


Fig. 1. Block Diagram of the Proposed System

A. Circuit Diagram

Here A8232 chip has 6 pins, we have used only 3 pins above them, which are output, voltage, ground. The Bluetooth module has 4 pins which are tx, rx, vcc, gnd. Arduino's TX1, RX0, 5V,GND pin are connected with the Bluetooth module through tx, rx,vcc and gnd respectively. The output, voltage, ground pin of AD8232 are connected with the Arduino's analog pin A0, 5V, GND. Arduino's VIN and GND is connected with the batteries positive and negative respectively.

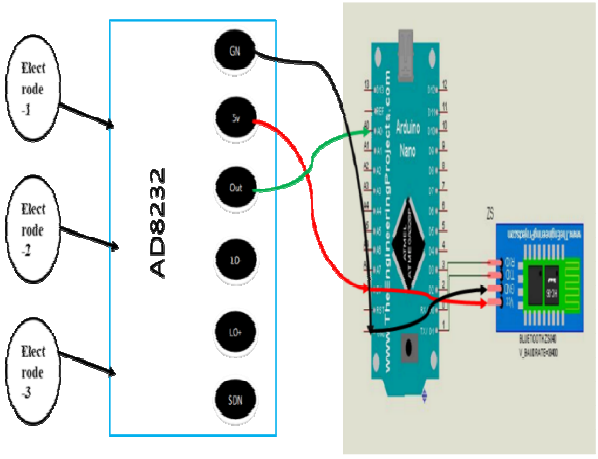


Fig. 2. Circuit Diagram of the Proposed System

IV. Designing Process of the Layout

After all the schematic circuit design process, the layout design process was done where several steps are required until the layout can be done properly. The layout design procedures are stated as follows:-

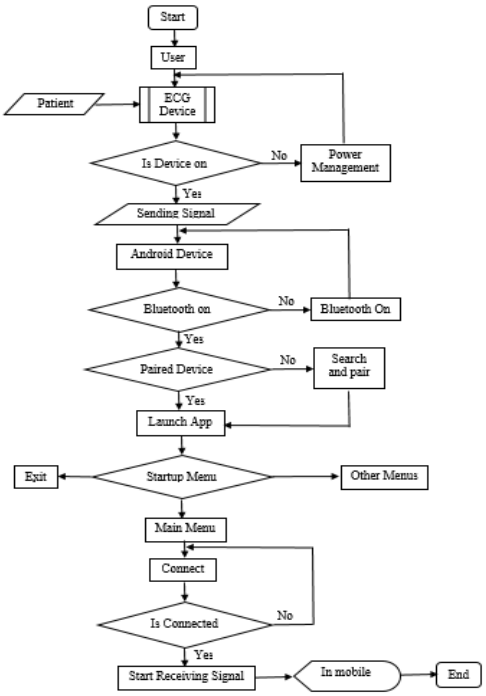


Fig. 3. Flow Chart

- Annotate the circuit of Portable ECG Monitoring Device.
- Run the Design Rule Check (DRC).
- Netlist process.
- Creating the circuit layout process.
- Rearrangement of the circuit.
- After all the process had been done, then the circuit will be route in automatically and so on, then it will be checked again using the design rule check.

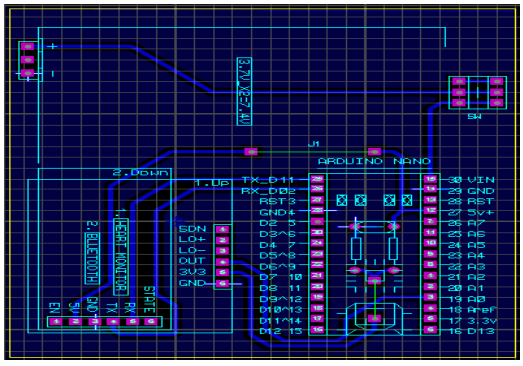


Fig. 4. Proteus Final Design

A. *User Interface*

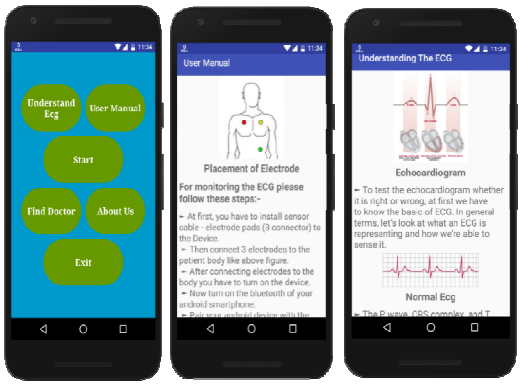
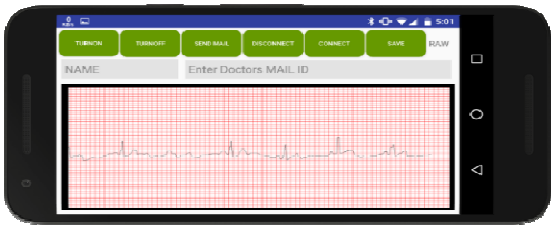
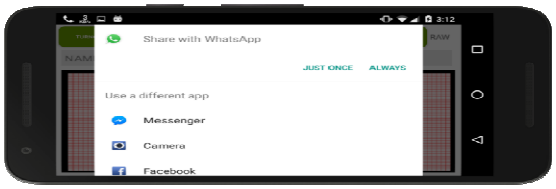


Fig. 5. (a) Startup Menu (b) User Manual (c) Understand ECG.



(d)



(e)

Fig. 6. (d) Showing ECG (e) Sharing Option

V. IMPLEMENTATION

Implementation takes a week to produce the printed PCB, completing each component located with the mark and drilled holes for mounting components. Firstly, this work printed the proteus design image on Photo Paper on a printer and make sure the output quality is the best.

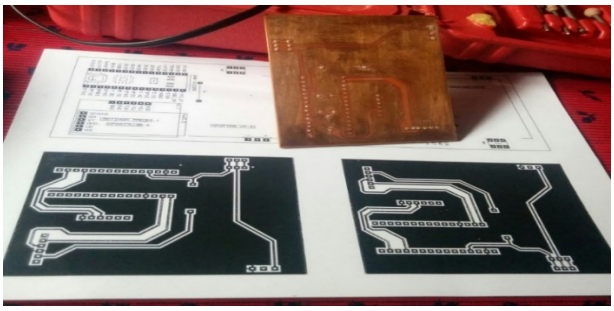


Fig. 7. Photo paper print

In this work placed the PCB transfer face down on the copper clad and set the household iron.



Fig. 8. Iron heat for making PCB

After 15 minutes, all the visible copper is gone, and then removed the PCB from the etchant and gave it a wash in cool water and dry with a towel. Then the board became ready to populate with parts.

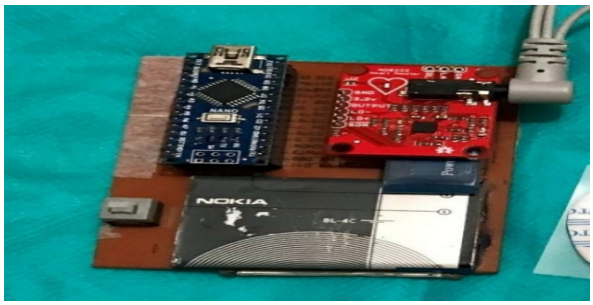


Fig. 9. Complete PCB

VI. OPERATING PRINCIPLE

In this system, Arduino is the main junction of all these peripheral devices and sensors. At first, we have to install sensor cable - electrode pads (3 connector) to the AD8232 sensor. Then connect 3 electrodes to the patient body like figure 10.

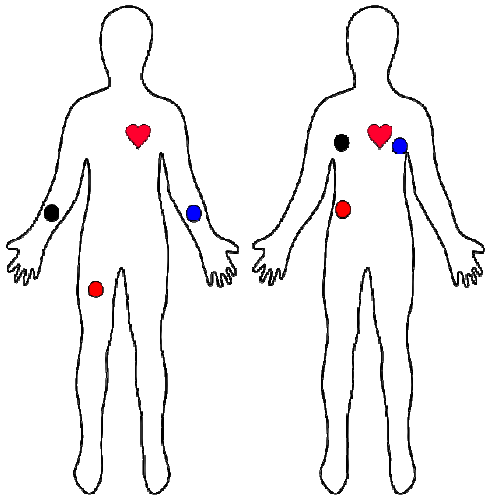


Fig. 10. Connecting Electrodes

After connecting electrodes to the body, have to turn on the proposed device and Bluetooth of the android device. Then this system have to pair the device with android device through via Bluetooth. Now we have to open the android app. Then press the start button. After opening the app, have to

press the connect button to connect with the ECG device. Then the device will show the electrocardiogram of the heart.

A. *Understanding the ECG*

To test the echocardiogram whether it is right or wrong, at first we have to know the basic of ECG. Generally speaking, let's look at what an ECG stands for and how we can feel it. The ECG is divided into two basic intervals, the interval of PR and the interval of QT as described below.

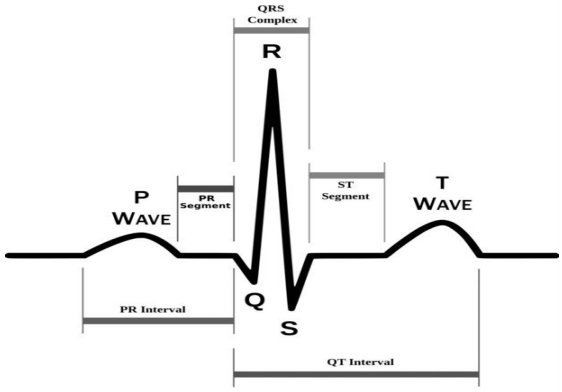


Fig. 11. .Basic ECG

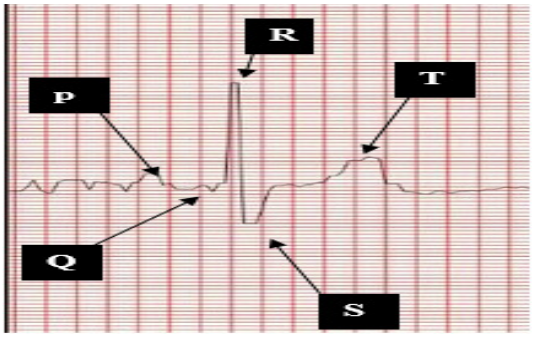


Fig. 12. Resultant ECG

In the above figure we compared our ECG with a normal ECG, which almost looks like. The P, Q, R, S, T segments are almost same in both pictures. Therefore, we can say that our system is showing normal ECG.

B. *Result*

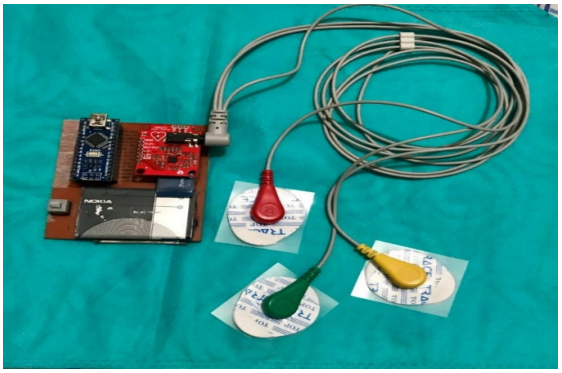


Fig. 13. Device Main view

Finally, in this season we are going to discuss about ECG monitoring and testing result. In this system, we will measure ECG through AD8232, Arduino and Bluetooth module and transfer it to the connected Android device. This project worked perfectly but sometimes it cannot give us the perfect reading due to ac noise.

VII. Conclusion

The project developed and implemented an affordable Arduino - based ECG monitoring system. The system aspects are developing a low - cost portable ECG monitoring system to analyze the heart conditions of the patient. This project proposes an effective heart attack detection system that helps to reduce deaths caused by heart attacks as the main cause of deaths from heart attacks is due to delay in proper treatment. This can be avoided as with the ECG report, the system will notify the doctor. In this work, an android application has been proposed that can receive ECG signals from the acquisition device wirelessly, detect QRS complexes, calculate heart rate and plot the real-time ECG signal for the cell phone display. It can also send this information to the physician concerned for medical decision by means of a server. Experiments show that the system proposed is unobtrusive and that the user can use it comfortably during daily activities.

Reference

[1] mHealth App Developer Economics(2014) available at http://mhealtheconomics.com/mhealthdeveloper-economics-report/

[2] Bourouis, A., Feham, M., and Bouchachia, A.(2011), " Ubiquitous Mobile Health Monitoring System for Elderly (UMHMSE)", International Journal of Computer Science and Information Technology, Vol.2, No. 3, June, pp. 74-82

[3] Lee, Y.D. and Chung, W.Y. (2009) "Wireless Sensor Network Based Wearable Smart Shirt for Ubiquitous Health and Activity Monitoring", Sensors and Actuators B: Chameical, Vol. 140, No. 2, July, pp. 390-395

[4] Orlando R. E. P., Caldeira, M. L. P. Lei S., and Rodrigues, J.P.C (2014), "An Efficient and Low Cost Windows Mobile BSN Monitoring SystemBased on TinyOS", Journal of Telecommunication Systems, Vol. 54, No. 1, pp. 1-9

[5] Yuce, M. R.(2010)" Implementation of wireless body area networks for healthcare systems", Sensor and Actuators A:Physical, Vol. 162, No. 1, July, pp. 116-129

[6] Lei Clifton, David A. Clifton, Marco A. F. Pimentel, Peter J. Watkinson, and Lionel Tarassenko (2014)," Predictive Monitoring of Mobile Patients by Combining Clinical Observations with Data From Wearable Sensors", IEEE Journal of Biomedical and Health Informatics, Vol. 18, No. 3, May , pp. 722-730

[7] Parane, K.A., Patil, N.C. ; Poojara, S.R. ; Kamble, T.S(2014) "Cloud based Intelligent Healthcare Monitoring System", In the proceedings of International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), February 7-8, Ghaziabad, Indian, pp. 697-701

[8] Xiaoliang Wang ; Qiong Gui ; Bingwei Liu ; Zhanpeng Jin et al (2014), "Enabling Smart Personalized Healthcare: A Hybrid Mobile-Cloud Approach for ECG Telemonitoring", IEEE Journal of Biomedical and Health Informatics, Vol. 18, No. 3, May, pp. 739 – 745

[9] Dunsmuir, D., Payne, B. ; Cloete, G. ; Petersen, C.(2014), "Development of m-Health Applications for Pre-eclampsia Triage", IEEE Journal of Biomedical and Health Informatics, Vol. PP, No. 99, January , pp. 2168-2194

[10] Tello, J.P. ; Manjarres, O. ; Quijano, M. ; Blanco, A. et al(2013) , " Remote Monitoring System of ECG and Human Body Temperature Signals", IEEE Latin American Transaction, Vol. 11, No. 1, February, pp. 314-318

Six Tier Multipurpose Security Locker System Based on Arduino

A. Z. M. Tahmidul Kabir
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
tahmidulkabir@gmail.com

Nirmol Deb Nath
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
debnathnirmol@gmail.com

Rafin Akther Utshaw
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
rafin.utshaw@gmail.com

Fukrul Hasan
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
dipta.hasan555@gmail.com

Tawsif Ibne Alam
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
tawsif.alam@aiub.edu

***Abstract*— This study presents a security system for transporting or storing valuable while restricting unauthorized access. The system consists of Memory Module, PIR sensor, fingerprint security, Encoder-Decoder, RF module, GPS and GSM module etc. to provide the maximum level of security. The proposed security system can be a major uplift in transporting important documents, money or ornaments from one place to another especially for banks in transporting valuables as the proposed security system is designed in such a way that the vault can only be opened by an authorized person, in specific places, using proper credentials, hence providing maximum security.**

***Index Terms*— DS3231 RTC module, PIR sensor, RF module, Encoder Decoder, Fingerprint sensor, Password keypad, Global positioning system (GPS) and Global System for Mobile communications (GSM).**

I. INTRODUCTION

The job of a security system is to keep valuable objects safe and secured for its rightful owner. Day by day security for valuable objects is becoming a major issue with increased crime rate. The conventional method of lock and key system is no longer appropriate in today's modern world. In recent times, use of alarm systems with dual key lockers, digital passwords, RFID, etc. is common. Variety of locker or vault type systems are available but each has its room for improvement, hence having its own flaws in the system. Theft of gold ornaments from Amin Jewelers is still a thing of commotion in Dhaka, Bangladesh [1]. Also, in terms of documents, leaking of classified information is nothing new. For example, over the years SSC and HSC examination question leaks was a threateningly auspicious debatable issue for Bangladesh. [2][3].

As a solution to these addressed issues and more, the study of a six-tier locker system has been proposed keeping in mind the drawbacks of the available systems. It has been prototyped, focusing on transporting valuables from one place to another in the most secured form. These valuables can be jewelry, bank drafts, antiques, gemstones, stamps, question papers and many more. The six-tier locker system is made using DS3231 RTC module, PIR sensor, RF module, Encoder Decoder, Fingerprint sensor, Password keypad, GPS and GSM.

II. BACKGROUND STUDY

One widely used security system is a locker having two keys and both are needed to open the lock, where one key generally is for the owner of the contents in the locker and another key is for the agent (a person who ensures security of the locker). But in this case, someone can easily pick the lock or make duplicates of the keys to open the lock causing a security issue.

There are also digital locking systems available but the problem with this is that by knowing the password or hacking into the system one can easily gain access.
Also, Srivatsan Sridharan in a paper [4] proposed a method which uses two levels biometric authentication meaning 2 people's biometric identities are needed to open the vault, ensuring only certain individuals to be able to open the vault. But it has no security check on time, date and location the vault is being opened at.
Authors in paper [5] showed a project which uses RFID reader module along with fingerprint and iris to secure the locker. Similarly, in the paper [6] they proposed a GSM-based locker system using RFID with fingerprint and a password.
In paper [7], they used GPS, fingerprint module and password protection to make a secured vault hence denying any access inside the locker while in transit.
Ch. Sumalatha et al in a paper [8] in 2016, proposed a bank locker security system also on RFID and GSM technology. The system was developed using GSM modem, microcontroller AT89C52, RFID reader, keyboard, and LCD. The id number is read via the RFID

reader and conveyed to the microcontroller. After that, the microcontroller checks if the id is authorized and if authorized, the locker is opened.
In paper [9], U. J. Ogri, D. E. B. Okwong and A. Etim created a model to use in security doors. Here one GSM device is used as a transmitter and another GSM device with the DTMF associated with the motor attached to a door with the use of DTMF decoder, a stepper motor and microcontroller unit.

All these existing systems are promising but all have their own drawbacks or faults in the system.

III. System Description

This system is developed for smart and efficient performance. Below are the main components of the Six Tier Security Locker System:

TABLE I. MAIN COMPONENTS AND THEIR FEATURES

Component name	No. of unit	Use
Arduino Mega	1	• To control all of the sensor and module. • Input voltage 6-20V
DC3231 RTC Module	1	• Operating voltage 3.3-5.5V • Used as a real-time clock • Storage capacity is 32K
PIR Sensor	1	• To detect any suspicious movement inside the box. • Operating voltage 4.5V
Fingerprint sensor	1	• Operating voltage is 3.6-6V. • For inserting a biometric password.
Password keypad	1	• For inserting a keypad password.
Encoder Decoder	1	• Wide operating voltage and high noise immunity • For encoding and decoding operation.
GPS	1	• The global positing system is used for tracking location. • Operating voltage 5V.
GSM	1	• GSM is used to send the updated status of this box to the user through SMS or phone call.
RF module	1	• This module has two part one is receiving and another one is transmitting. • RF module and Encoder Decoder is used to make a wireless password system.

[a] Components feature.

IV. System Design

A. Simulation:

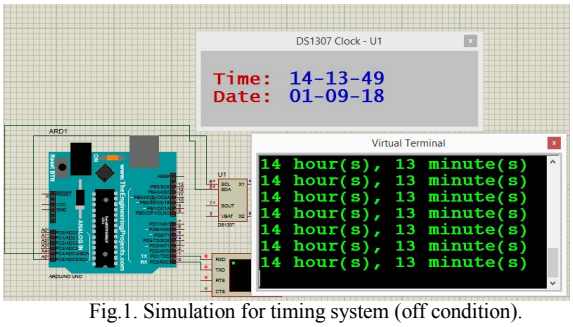


Fig.1. Simulation for timing system (off condition).

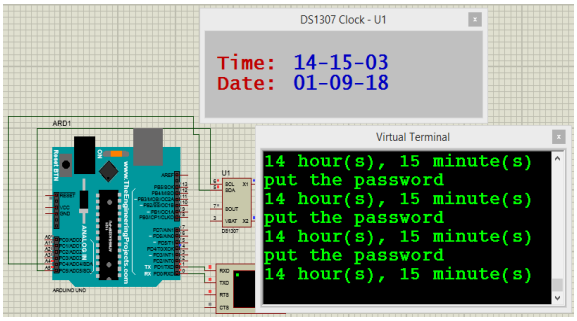


Fig.2. Simulation for timing system (on condition).

In Fig 1 and 2, it shows the simulation of the timing part. Where the Arduino and DS1307 RTC module has been used instead of DS3231 RTC module. Here DS1307 RTC module is working as a real-time clock. The whole system has been programmed in such a way that it will ask for the password over and over again after certain periods of time. For the rest of the time it will be inactive.

B. System overview:

The system will only unlock at specified destinations and at specified time. Upon arrival at the correct geographic location, when the clock approves of the correct time, to unlock the system user will have to enter the correct password as shown in fig. 4. Once the password has been accepted, the authorized person's fingerprint needs to be provided as shown in fig 5. An extra feature of grand access has been also added to the system. If a person has grand access, he/she shall be able to open the locker using his fingerprint alone, regardless of time and location.



Fig.3. Prototype box.



Fig.4. Giving password.

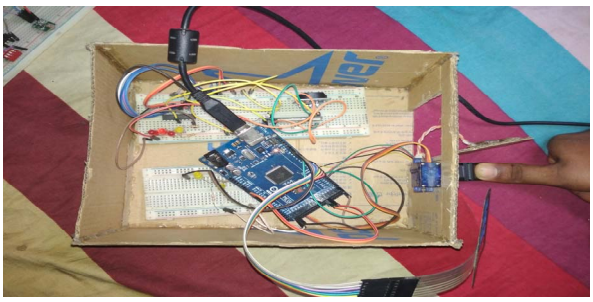


Fig.5. Box unlock by giving fingerprint.

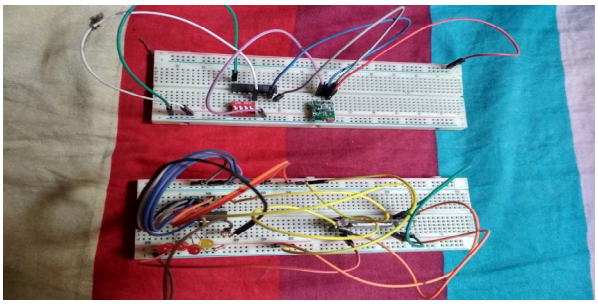


Fig.6. Transmitter and receiver circuit.

V. FEATURES OF THIS STUDY

- Six layers of high-security level.
- All electronic component used has low power consumption.
- Only one grand access using fingerprint is provided, making the system very secure.
- Suspicious movement inside the box before opening of the locker door will be notified to the owner using GSM module.

VI. WORKING PRINCIPLE

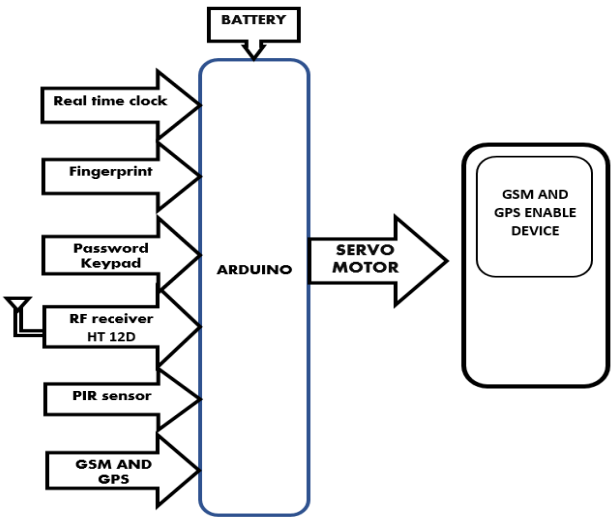


Fig.7. Working block input and output.

Fig '7' shows the Inputs and Output block diagrams. The battery is working as the power source of the system. On the left side of the diagram, the inputs are shown and the output is shown on the right side. All the inputs on the left side are connected to the Arduino and all the inputs are taken from the users of the system, except the GSM and GPS. The Arduino analyses all the inputs according to the instruction given and provides the desired result on the output end.

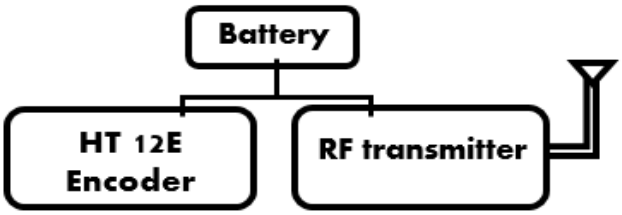


Fig.8. Working block of wireless password system.

Fig '8' shows the wireless password generation. It includes a battery, encoder and an RF transmitter. Combining the encoder and the RF transmitter a wireless password generator is created that generates a password which will be counted as input at the receiver end of fig '7'. And if the input password matches with the preset password then Arduino will get a positive feedback.

A. Flow diagram:

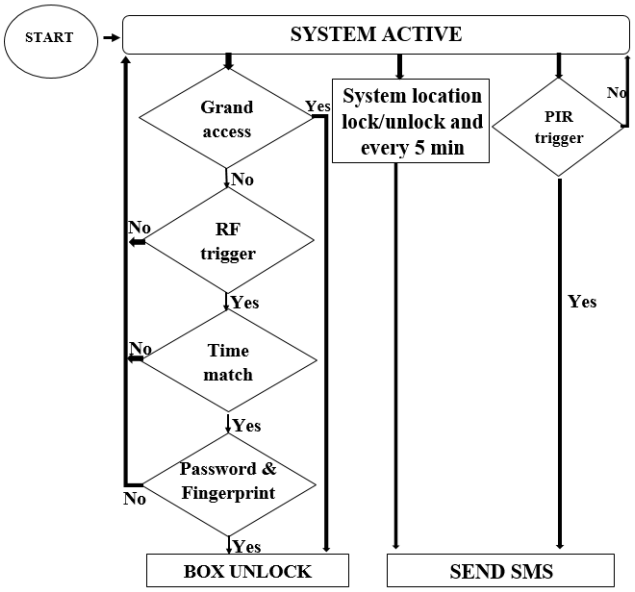


Fig.9. Flow diagram of six tier security locker system.

The Six Tier Security Locker System starts initiating at the very moment it has been locked. Locking gives the Six Tier Security Locker System power to initiate and it starts all the components inside the box. Arduino Mega (ATmega2560), RF module, DS3231 RTC module, Fingerprint sensor etc. This system's flow-diagram works in two stages. At the first, stage, if the fingerprint of the grand access is matched with the pre-set fingerprint it doesn't need to go towards the 2nd stage. But if somehow, it's unmatched then it will definitely go on to the 2nd stage, for unlocking and other security purposes.

After the initiation of the Six Tier Security Locker System, firstly the RF module will be triggered. The RF

module has a receiving part and a transmitting part and combining with an Encoder-Decoder module, such a system has been made that will transmit value from the transmitting part and that value will be received at the receiver part. If the value matches at both the transmitting and receiving parts it will be analyzed at the Arduino Mega (ATmega2560) as positive and the DS3231 RTC module will then be triggered otherwise, it won't. The memory module has been programmed in such a way that it has been clocked at a fixed time to initiate the password keypad. When the Memory module has been initiated, it will be analyzed at the Arduino Mega (ATmega2560), if it has reached at the preprogrammed fixed time or not. If the analyzed time meets the requirement of the DS3231 RTC module, only then the password keypad will be triggered and will be ready to take input otherwise it won't. If the input password matches with the stored preset password, the fingerprint sensor will be triggered. The Fingerprint sensor has been on from the very beginning of the initiation of the Six Tier Security Locker System and now that it has been triggered, it can take input and given the correct input, finally, it will unlock the Six Tier Security Locker System otherwise it won't unlock.

To provide a better security level, a GPS module has been implemented that will detect the latitude and longitude of the Six Tier Security Locker System and through a GSM module, the latitude and longitude along with the current status of the locker weather it has been opened or not, this information can be sent through SMS to a specific cellular device. And this information will be updated at 5-minute intervals. That way the location of the Six Tier Security Locker System will be traceable. The user can choose not to share the location of the box as well.

Another feature that has been integrated is, to trace any kind of movement inside the Six-tier Security Locker System through a PIR sensor and send the location of the box to the tracing devices, before opening of the locker door.

B. Power calculation

The power consumption of the whole system has been shown in the table II. We have calculated the power consumption and predicted the time for sending the system to its proper destination. We have used 5V, 20 mAh battery, that gives the system approximately 11 to 12 hours backup which is enough for the system to reach the destination.

TABLE II. POWER CONSUMPTION OF THE SYSTEM

Component	Power Consumption per day (mA)
Servo motor (6sec ON)	2 A/s *6s (max 6s ON) =12A
Arduino Mega (30 min ON)	170mA*60s*30min =306000mA =306A
GPS module (24 hour ON)	0.006(mA/h) *24h =0.144 =1.44*10^-4A
GSM module(24h)	2(mA/s) * 60s*60min*24h =172.8A
Fingerprint module(30s)	150(mA/s) *30s =4500mA =4.5A
Memory module (24h)	40(mA/h) *24h =0.960A
Matrix keypad(60s)	30(mA/s) *60s =1.8A
RF module (24h)	5(mA/s) *60s*60m*24h =432A
PIR sensor(24h)	504mA
Total	929604.144mA/d
Total	38733.506mA/h
Total Service Hour	~ (11 or 12) h We Provide (5V 20000mAh battery) All Components Input 5V

[b] total power consumption.

VII. COST CALCULATION

Each component with their corresponding price is given in Table III.

TABLE III. COST CHART

SL NO.	Component List	Price Per Unit (*Dollar*)	Quantity (PCS)	*Total Cost* (Dollar)
01.	Arduino mega2560	$10.10	1	$10.10
02.	DS3231 RTC module	$1.92	1	$1.92
03.	PIR Sensor	$0.95	1	$0.95
04.	Fingerprint sensor	$19	1	$19
05.	Password keypad	$2.5	1	$2.5
06.	RF module	$3.45	1	$3.45
07.	Encoder Decoder	$2.38	1	$2.38
08.	GPS and GSM	$23.78	1	$23.78
09.	Battery	$23.78	1	$23.78
10.	**Total Cost**			**$87.86**

[c] Device cost.

VIII. NOVELTY OF THE WORK

The tradeoff between cost and quality of work is something to point out in this study. The components used to build this system are very cheap, although the

output and the end results are very satisfactory. With the use of PIR sensor we can easily detect any intrusions that may appear very effectively and easily. Also, another contributing factor when it comes to novelty is the use of RF module instead of using RFID. Whereas RFID needs its own set of components and Arduino system to work, RF module has no such hassle. An RF Module consists two major components such as a transmitter and a receiver. One is placed inside the locker system and another at its destination, ensuring the locker to be only enabled for opening at desired destination(s). Adding to this, with the use of GPS and GSM the location and status of the locker is sent to a specific cellular device at 5-minute intervals making it traceable and ensuring the locker doesn't go off route. The user can also choose to opt in and out of the location sharing option.

IX. Future Aspects

Future improvements include three major upgrades. Firstly, using iris recognition along with fingerprint to add an extra layer of security. Secondly, use of a CMOS camera inside the box for monitoring, which will act as a complement to PIR Sensor. This will provide a visual view of any disturbance the PIR sensor detects. Lastly, syncing and preserving all kind of information sent from or received to the locker in a server. This may act as a log of all kinds of alarms, warnings or actions.

X. Conclusion

In Six Tier Security Locker System, all the features that help in the security of the device are brought together in a single module. Here Memory Module (real time clock) has been used to determine the preset time to unlock the Six Tier Security Locker System. PIR sensor has been used to detect any kind of action or movement inside the box. The fingerprint sensor has been used to provide fingerprint unlocking security, only when the preset time on the timer module is reached. Only one single fingerprint has been used as the grand access to unlock the box anytime. An RF module with encoder-decoder is used to create a wireless password system to generate passwords for the Six Tier Security Locker System to unlock. GSM and GPS module has been used to trace the current location of the Six Tier Security Locker System and sending the location details and other information via SMS service as an additional security feature. Overall the system is also very much cost effective. It took only $87.86 to build the whole system, without the structure of the safe, which may vary depending on the size and the material of the locker, required by the consumer. The device can be considered affordable to anyone or any organization at an affordable cost, while ensuring safely.

Reference

[1] "Dhaka Tribune," Dhaka Tribune, 18 April 2018. [Online]. Available:https://www.dhakatribune.com/bangladesh/crime/2018/04/18/stolen-gold-of-amin-jewelers-recovered-four-arrested?fbclid=IwAR3rpLzwmt-b4_pIL1uDFlMPDzIAfXDB6ctA4L1TUDu3hJRhbSopRTXwYec. [Accessed 13 April 2019].

[2] "The Daily Star," The Daily Star, 1 February 2018. [Online]. Available:https://www.thedailystar.net/frontpage/ssc-exams-question-leak-allegations-surface-again-1528792. [Accessed 13 April 2019].

[3] "Dhaka Tribune," Dhaka Tribune, 23 February 2018. [Online]. Available: https://www.dhakatribune.com/opinion/special/2018/02/24/question-paper-leak-ominous-sign-nation?fbclid=IwAR2EN9XvGk01gRSPRxr6dUNVqnJiEOPqKXJxG9aHtxkIGhQvnn6lSR8oAVU. [Accessed 13 April 2019].

[4] S. Sridharan, "Authenticated secure bio-metric based access to the bank safety lockers," *International Conference on Information Communication and Embedded Systems (ICICES2014)*, Chennai, 2014, pp.1-7. doi: 10.1109/ICICES.2014.7034063

[5] D. Goud, I. Md and PJ.Saritha, "A Secured Approach for Authentication System using Fingerprint and Iris ISSN : 2277-6370," *Global Journal of Advanced Engineering Technologies,* vol. 1, no. 3, pp. 170-175, 2012.

[6] H. S. Detroja, P. J. Vasoya, D. D. Kotadiya and P. C. B. Bambhroliya, "GSM Based Bank Locker Security System using RFID, Password and Fingerprint Technology," *IJIRST –International Journal for Innovative Research in Science & Technology,* vol. 2, no. 11, pp. 110-115, 2016.

[7] Sowndhaya.A, Jwalalakshmi.P and Abrahamsudharsonponraj, "Location Based Bank Locker Security System," *National Conference on Science, Engineering and Technology (NCSET – 2016),* vol. 4, no. 6, pp. 50 - 52, 2016.

[8] Ch.Sumalatha, A.Viyayamanasa, K.Ramasrujana, I.Meghamala, K. Prasanna and K. Rani, "International Journal for Research in Applied Science & Engineering Technology (IJRASET)," vol. 4, no. 4, pp. 989-993, 2016.

[9] U. J. Ogri, D. E. B. Okwong and A. Etim, "DESIGN AND CONSTRUCTION OF DOORLOCKING SECURITY SYSTEM USING GSM," *International Journal Of Engineering And Computer Science ISSN:2319-7242,* vol. 2, no. 7, pp. 2235-2257, 2013.

Design and Implementation of a Vault Security System

Md. Mahadi Hasan Moon
Dept. of Electrical and Electronics Engineering
American International University-Bangladesh
Dhaka, Bangladesh
mahadihasanmoon95@gmail.com

Debashish Kumar Ghosh
Dept. of Electrical and Electronics Engineering
American International University-Bangladesh
Dhaka, Bangladesh
debashishkumarghosh786@gmail.com

Nafiz Ahmed Chisty
Faculty of Engineering
American International University-Bangladesh
Dhaka, Bangladesh
chisty@aiub.edu

Md. Mahidul Islam
Dept. of Electrical and Electronics Engineering
American International University-Bangladesh
Dhaka, Bangladesh
mahidulislam2021@gmail.com

Md. Aurongo Jeb
Dept. of Electrical and Electronics Engineering
American International University-Bangladesh
Dhaka, Bangladesh
Aurongoj24@gmail.com

***Abstract*—In the era of technology, humankind is now looking for a rapid and prominent solution for day to day problems which can meet the demand and makes life easier. As security is an essential perspective with the end goal to protect our secret belongings, people are looking for an advance, unbreakable and user-friendly solution. However, considering a vault security system, we have to think a step further, as vaults usually used for higher security purposes. Security system using facial recognition, fingerprint scanner, password lock and RFID reader is going popular and common these days because of the availability and trustable security. The main goal is to implement a security system including all of these four sensors in single hardware which is able to provide much stronger security. The system is able to detect multiple human faces and also able to monitor the area continuously using one single camera sensor with the help of real-time video processing using open CV method. The system implemented in such a way that anyone who wants to enter the vault has to pass through all of these four (facial recognition, fingerprint, password and RFID scanner) steps. Neither one can be skipped. All the sensors are connected with Raspberry Pi to achieve the goal.**

Keywords— Security, Facial recognition, Open CV, Raspberry Pi, Fingerprint scanner.

I. INTRODUCTION

At present human work diminished by technology, thus technology plays a major role to ensure possession of individuals belongings. The automation of electronic equipment made it possible rather than wasting human resources and money for security [1]. Relying on technology this project develops a system that restrains human substances in a comprehensive way. This system can be served any individuals or any institutions to defend their confidential such as bank vault, money locker etc. Consider a breach of security the system is a combination of biometric and prior tech to ensure security escalation of this system. The objective is delivering security progression over extremely concealed premises effortlessly at an affordable cost [1]. Surveillance and facial recognition being operated by a camera for premium security purpose as it stores surveillance data on a memory chip for a definite period. The camera is continuously searching for a human face, and an authorizer is able to see the situation through the camera in real time. To achieve this goal, wireless communication established between the security system and the authorizer via VNC viewer. When a person places himself in front of the camera, the camera immediately scans his face with the image database stored in Raspberry Pi. The authorizer is able to see the name of that person at the top of his face in real time video. If the face doesn't match with the database, the name will be replaced by an unknown, and the system generates a sound which indicates the sound of an unknown person. After facial recognition, the system will ask for fingerprint, password and RFID card scan.

II. PRIOR AND RELATED WORK

Almost every smart device uses facial recognition to secure their devices from unauthorized people. It is also massively used for smartphone, network security and Deep learning-based video analytics. Other sensors like Fingerprint sensor, RFID reader are also used in security system and various application. A smart security system based on face recognition was proposed by Dwi Ana Ratna Wati and Dika Abadianto in 2017 [2]. The security system was based on MyRIO 1900 and programmed using LabVIEW [2]. The system can detect a person's face when the distance is less than or equal to 240 centimeters or even if the person uses various accessories like glasses or cap [2]. But it cannot identify the face if there is too much change or if the person is a little far from the camera. Another face recognition security system was implemented in 2018 by Ibrahim Mohammad Sayem and Mohammad Sanaullah Chowdhury using Raspberry Pi based on Python language with the help of OpenCV Library which is very similar to the face detection process used in this paper [3]. The device also sends email to the authorizer if there is an unknown person in front of the camera [3]. But as the system is only based on face detection, it is easy to break or bypass. An improved version of the security system was implemented by Teddy Mantoro and

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Suhendi which can detect multiple faces at a single time using a hybrid method of Haar Cascade Classifier [4]. This system is more efficient and accurate compared to others. An advanced bank security system was designed by Raj Gusain, Hemant Jain and Shivendra Pratap which use face recognition, iris scanner and palm vein for authorization [5]. But the system is hard to maintain and very costly. Whether there is also a cost-effective solution, which is based on Arduino designed and implemented by Kanza Gulzar, Jun Sang and Omar Tariq [6]. But the time consumption is higher in that device and the device is not able to recognize a human face in real time video [6].

III. System Overview

The following "fig. 1" represents the block diagram of this system. To assure real-time broadcasting without interrupting the whole system, two of the separate virtual environments introduced. One is for live image processing (Open CV), and another is for other sensors (fingerprint, password and RFID). i2c communication is established to ensure a continuous connection between these two virtual environments. The camera sensor is continuously monitoring the area send signals to another environment if there is a person (e.g. alpha for an unknown, beta for David, gamma for Raj etc.). Thus, the Open CV environment is only focusing on video broadcasting and transferring signal. As a result, an authorizer is able to monitor the area continuously without any interruption. The second environment analyzes the signal received from the Open CV and commands the system for further instruction. In this case, if the signal represents an unknown person (e.g. the signal alpha which represents an unknown person) the system provides a buzz. But if the signal represents a known person (e.g. the signal beta which represents David) the system will command for fingerprint scan with the help of a 20x4 LCD display which placed at the system. When the person/visitor placed his/her finger on the fingerprint scanner, the system matched the fingerprint template with stored fingerprint of that person. If the fingerprint is matched, the system will give the confirmation and ask for a password to the person/visitor via a 20x4 LCD display. At this time, the visitor will place his password to the system via a 4x4 matrix keyboard and can also able to see whether it matched or not. If the password is matched, the LCD display confirmed it to the visitor and asked for RFID card scan. In this stage, the person/visitor will place the RFID card in front of the RFID scanner. And if the RFID card is matched, the system will activate the electronic door lock and allow that visitor/person to enter the vault.

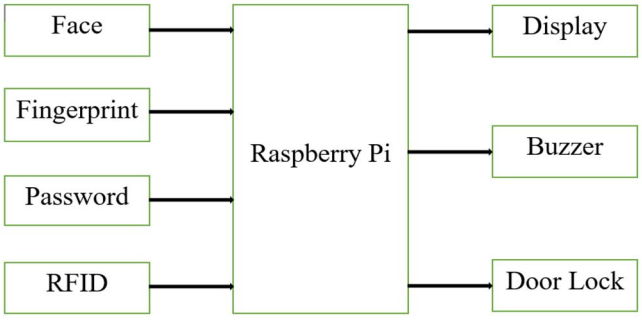


Fig. 1. System block diagram

If one of these steps is failed or not matched, the system will start from the beginning. Keep in mind that, password, fingerprint and RFID card vary from person to person (for example, if the face matches with David, the system require Divides fingerprint, password and RFID card). Neither of them can be entered changeable even if both of them are authorized.

IV. System Methodology

This section demonstrates the methodology of how this system work and interfacing with sensors and Raspberry Pi. Raspberry Pi is the main brainer in this system which interprets all the data to establish a decision.

A. Facial Recognition using Open CV

The facial recognition of a human is made possible with the help of Open CV library in python language. Open Source Computer Version Library which is known as Open CV, provide a common infrastructure for computer vision application [7]. Open CV library contains over 2500 optimized algorithms including a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms which can be used to detect and recognize a face, identify objects, classify human actions in videos, track camera movement, track moving object and many more [7].

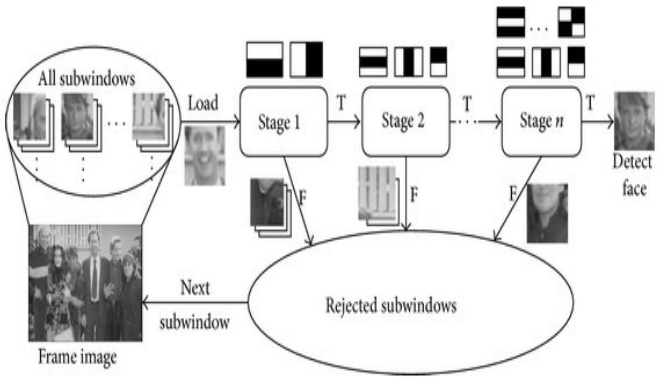


Fig. 2. Cascade structure for Haar classifiers [8].

Here, the goal is to detect/identify a human face in real time video. The algorithm used to achieve the goal is Haar Cascade Classifier. The hybrid method of Haar Cascade Classifier is used in this system so that the system can detect multiple faces at a single time. The cascade function is trained from a lot of positive and negative images to detect an object in other images [7]. Three types of features (edge features, line features and four-rectangle features) are used to train the classifier obtaining single value by subtracting the sum of pixels under the white rectangle from a sum of pixels under the black rectangle [7]. The classifier uses a fixed scale, for example, 50x50 pixels. Applying each and every feature on all the training images, the classifier can find the best threshold which will classify the faces to positive and negative [7]. As most of the image is non-face region, the features are grouped into different stages of classifiers and applied one-by-one [7]. The window, which passes all the stages is a face region [7]. Open CV contains two parts (Trainer and detector). Trainer is used to train the classifier and the detector is used to detect the face. To train the

classifier a dataset of images is needed. Every image contains cropped face of a human from a different angle in grey scale.

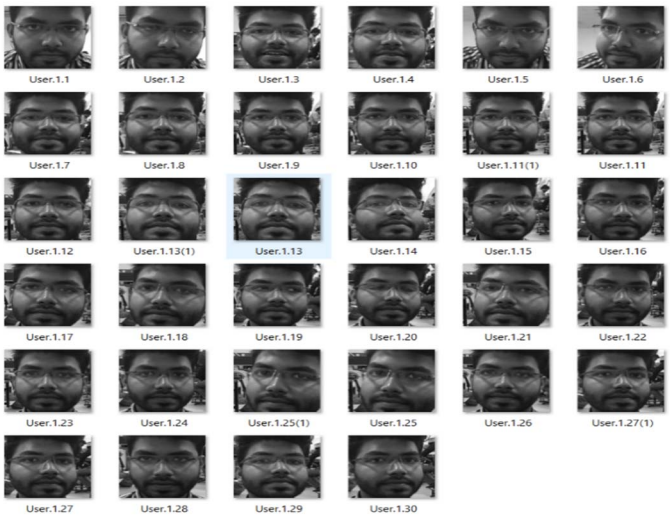


Fig. 3. Dataset of images in grey scale.

With the help of python language, the detector continuously matches the real-time image with those images stored in the dataset.



Fig. 4. Face recognition using Open CV.

B. *Fingerprint Scanner*

The fingerprint matching algorithm is responsible for producing a similarity score for the input and template prints. After analyzing similarity, the resulting match score will be compared with a certain threshold if both original fingerprints are formed (or not) from the same finger. Many methods are available for fingerprint matching among them minutia-based matching technique commonly used due to his less computational cost and good performance [9]. Minutiae-based algorithm established on local structure (point pattern matching), and global structure (finding the alignment of the corresponding pair of minutiae sets) [9].

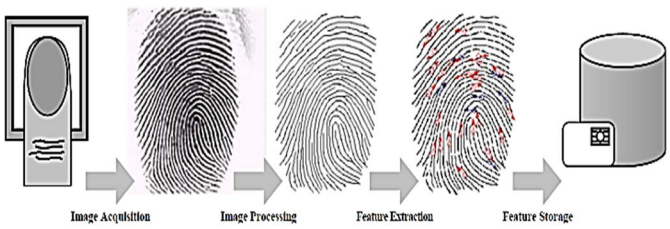


Fig. 5. Finger-print enrollment process [9].

First, the local structure of each minutia point is obtained in both minutiae sets to determine the similarity of fingerprints. The local structure describes minutiae spatial characteristics taking into account its meticulous neighborhood. This local descriptor (ridge ending, ridge bifurcation) is an invariant feature of rotation and translation that is inherent in minutiae [9]. Comparing both the template and the input minutiae a similarity matrix formulated in order to analyze the score of similarities between any combination of minutiae pairs. Local structure similarity scoring allows the best-matched minutia pair to be identified and exercised as a reference for global structures.

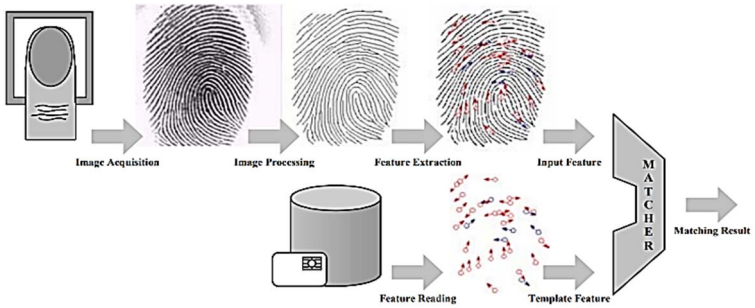


Fig. 6. Finger-print authentication process [9].

Global structures defined by minutiae sets where it describes the spatial characteristics from a reference point. After accomplished with the global structures the system decides the results obtain from global structure compared with the certain threshold.

C. *4x4 Matrix Keypad*

The 4x4 matrix keypad is used to input a password into the Raspberry Pi. The number of pin requirement for digital input can be reduced, by using the matrix keypad. The matrix keypad contains 16 pins in 4 rows and 4 columns. The keypad has 8 pins. 4 is for columns and 4 is for rows. By making 4 rows as output and considering them logic low, and 4 columns as input, the possible key pressed can be detect when correspondent columns gets o or low [10]. When the switches are not pressed it needs to make the column pins as high or 1, otherwise it is not possible to detect the logic changes when the switch is being pressed [10]. But there is a chance of spikes or noise generation if multiple switches are pressed at a time [10]. To avoid the noise, a few milliseconds delay is required to recheck the switch conditions [10].

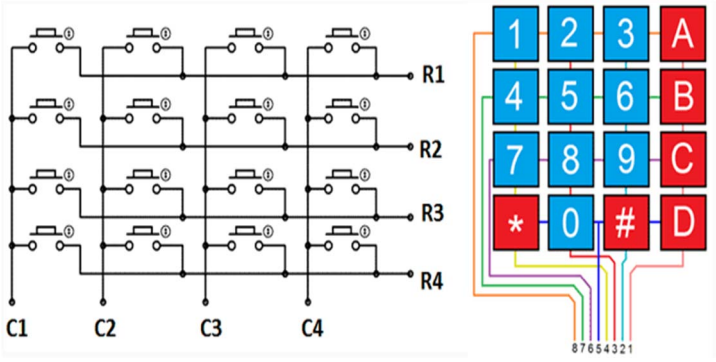


Fig. 7. Connections of a 4x4 matrix keypad [10].

D. RFID Card Scanner

The Radio frequency identification namely called as an RFID. It has two main components a transponder, associated with the object which is supposed to identify. The other one is Transceiver which consists of a radio frequency module, a control unit and antenna coil, which generates a high-frequency electromagnetic field [11]. When the transponder approaches towards the transceiver, a voltage induced due to induction and a microchip inside the transponder gets power to operate. Load manipulation technique is used to transfer data between transponder and transceiver. Switching of tag antenna using control unit causes a voltage drop that resultant in ones and zeroes [11]. This data used to transmit the signal between the transponder and transceiver.

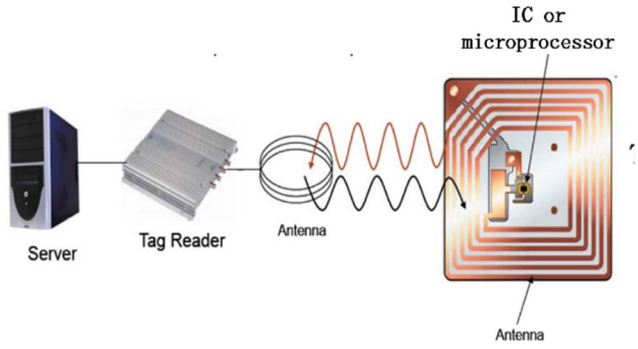


Fig. 8. RFID system [11].

V. Experimental Setup & Programming Flow Chart

The Experimental Setup with all the interfaced components are shown in "Fig. 9".

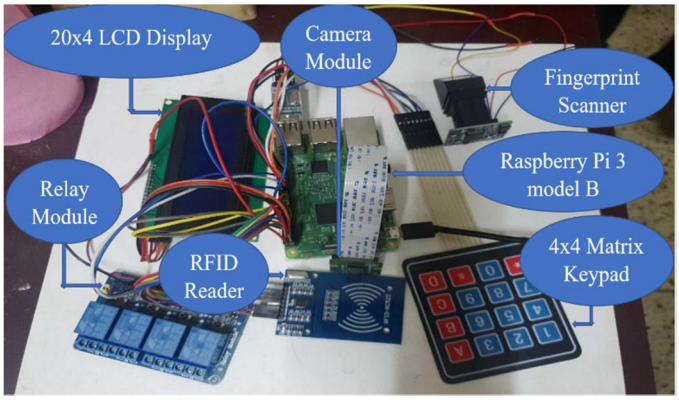


Fig. 9. Experimental setup of the system.

The programming flow chart of the system is shown in "Fig. 10".

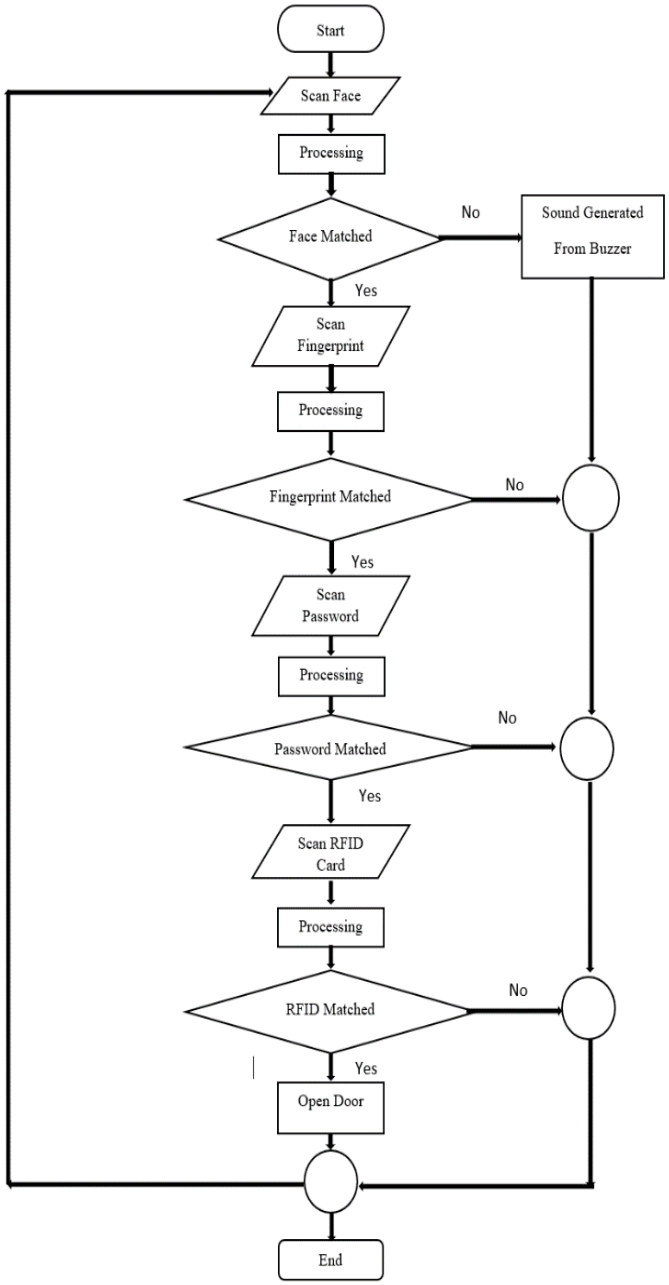


Fig. 10. Flow chart of the program.

VI. Implementation & Results

The system is developed in a single Raspberry Pi board using python language. Raspberry Pi is a mini portable onboard computer which has 4 USB ports, an Ethernet port, a HDMI port, a 3.5mm audio jack, onboard camera and display port, a MicroSDHC slot, 40 GPIO pins including power and ground and a micro USB port to provide 5 volts dc power. The system is powered by a Broadcom BCM2837 chip with ARMv8-A instruction set architecture (ISA) using 4 core Cortex A53 based CPU clocked at 900 MHz. It has 1GB RAM shared with Broadcom VideoCore IV GPU. The Raspberry Pi 3 model B also comes with on-board network (2.4 GHz wireless and Bluetooth 4.2). The system consists of a controlling device (e.g. laptop, smartphone etc.) over same Wi-Fi connection, a camera module, fingerprint scanner, matrix keypad, RFID reader/scanner, relay module, electronic door lock, buzzer and some LEDs. "Fig. 11" shows the real time monitoring over wireless communication via VNC viewer.

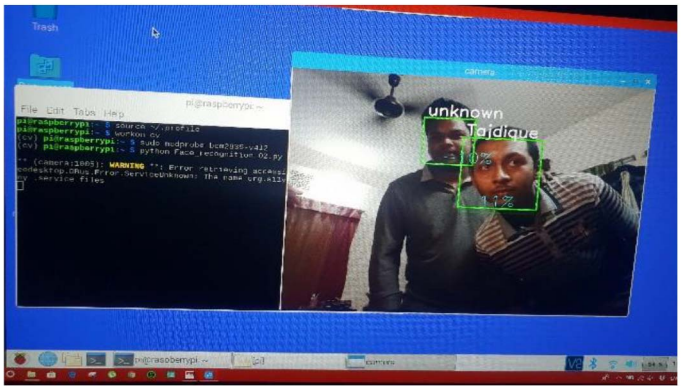


Fig. 11. *Monitoring real time video streaming via VNC viewer.*

A 20x4 Character LCD Display is used in this system to communicate with the user. The display is connected with the GPIO pins of the Raspberry Pi. An I2C LCD Gadgeteer module is used to reduce the connecting wire. The display provides instructions to the user. “Fig. 12” shows the initial state of the LCD display.

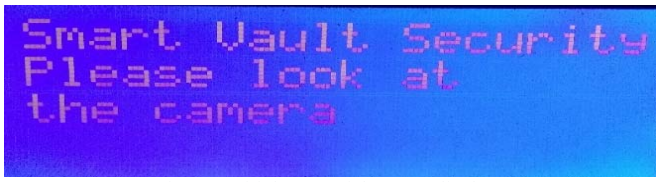


Fig. 12. Initial stage of the LCD display.

The Raspberry Pi Camera Module V2 is connected with the Raspberry Pi using onboard camera port for the purpose of real-time video broadcasting and facial recognition. The Camera Module has a 8 megapixel Sony IMX219 sensor which can be used to take high-definition video and photo. The camera identifies the person/visitor and the LCD display sends the confirmation and farther steps to the person/visitor which are shown in “Fig. 13” and “Fig. 14”.

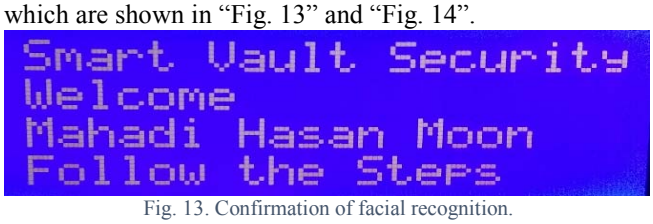


Fig. 13. Confirmation of facial recognition.



Fig. 14. Instruction for finger placement.

R30X Fingerprint Sensor is used to scan and identify the fingerprint of the visitor. It requires a UART interface over USB port. To connect the Fingerprint Sensor with Raspberry Pi a PL2303 USB to TTL Converter is needed. The voltage requirement is in between 3.6 to 6.0 volts. The Fingerprint Sensor will activate after facial recognition. The visitor is able to see whether his/her fingerprint is matched or not. All the results and steps are being placed on LCD display which will shown in “Fig. 15” and “Fig. 16”.

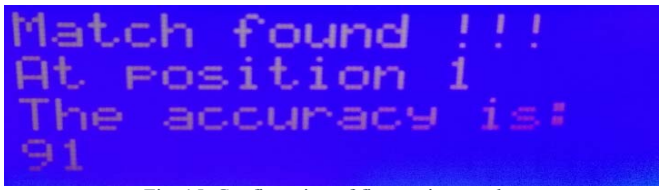


Fig. 15. Confirmation of fingerprint match.

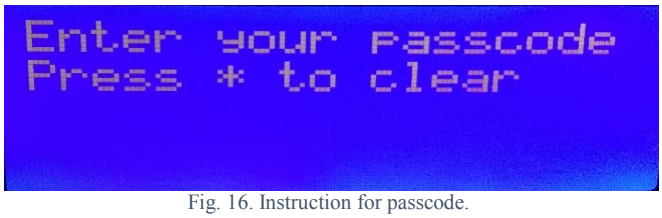


Fig. 16. Instruction for passcode.

A 4x4 matrix keyboard is used to enter the password to the system which is directly connected with Raspberry Pi GPIO pins. From total 8 connectors, 4 of them are input and 4 of them are output. The process wise results are shown in following figures from “Fig. 17” to “Fig. 19”.

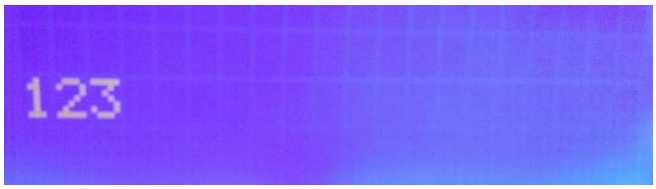


Fig. 17. Input Password.

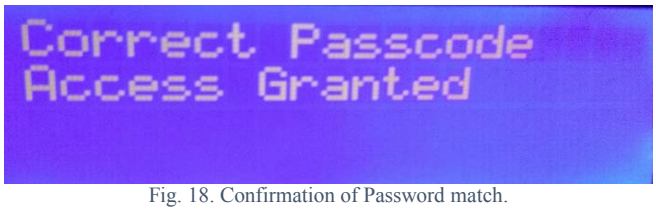


Fig. 18. Confirmation of Password match.



Fig. 19. Instruction for RFID Card placement.

To read and write RFID cards, MFRC522 RFID sensor is used. It is a highly integrated RF frequency reader/writer IC for contact less communication at 13.56 MHz and also supports ISO/IEC 14443 A/MIFARE mode. The voltage requirement for the sensor is 3.3 volts. It is the last and final stage of this security system. After placing the RFID card on the RFID sensor, the system matched the card number and gives the confirmation which shown in "Fig. 20".



Fig. 20. Confirmation of RFID match.

After completing all the steps, the system turned on the 12 volts electronic door lock which is connected with the raspberry pi via relay module. A 1100mAh 12 volts Lithium Polymer Battery is used to power the electronic door lock. “Fig. 21” shows the complete setup of the system.



Fig. 21. Complete setup of the system.

VII. LIMITATIONS

Each project has some limitation that can be eliminated by conducting future research or development. In this project, the calculation of entering people in a vault or confidential room is not ensured. Face detection of the night mode is not available. In the limitation side, the real-time delay in broadcasting the surveillance has been counted.

VIII. CONCLUTION

The ultimate goal is to provide society with a reliable and uncompromising security system that people can afford to obtain better security assurance. The system improvises the multi-step security authentication to obtain a result. It strengthens security by effectively limiting trespassing, heist, or security threats of any kind. This project has developed with the raspberry pi embedded with the identification sensors. The OpenCV environment has used for face detection, which bases on viola jones algorithm using Haar cascade classifier which superimposes the positive image over a set of negative images known as favored face detection classifier in OpenCV. The entire project framework has programmed with the python language. Future development can be developed adding gas sensors, humidity sensors for measuring their concern-parameters. An emergency alarm can be generated in the event of Emergency. Although the system is designed for vault security, it can be used for any security purpose like home security, office security etc. Considering the social impact of this project, it has done with a lot of dedication and integrity.

REFERENCES

[1] K. H. S. Murugan, V. Jacintha and S. A. Shifani, "Security system using raspberry Pi," 2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM), Chennai, 2017, pp. 863-864.

[2] D. A. R. Wati and D. Abadianto, "Design of face detection and recognition system for smart home security application," *2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, Yogyakarta, 2017, pp. 342-347.

[3] I. M. Sayem and M. S. Chowdhury, "Integrating Face Recognition Security System with the Internet of Things," *2018 International Conference on Machine Learning and Data Engineering (iCMLDE)*, Sydney, Australia, 2018, pp. 14-18.

[4] T. Mantoro, M. A. Ayu and Suhendi, "Multi-Faces Recognition Process Using Haar Cascades and Eigenface Methods," *2018 6th International Conference on Multimedia Computing and Systems (ICMCS)*, Rabat, 2018, pp. 1-5.

[5] R. Gusain, H. Jain and S. Pratap, "Enhancing bank security system using Face Recognition, Iris Scanner and Palm Vein Technology," *2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU)*, Bhimtal, 2018, pp. 1-5.

[6] K. Gulzar, Jun Sang and O. Tariq, "A cost effective method for automobile security based on detection and recognition of human face," *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, Chengdu, 2017, pp. 259-263.

[7] J. J. Patoliya and M. M. Desai, "Face detection based ATM security system using embedded Linux platform," *2017 2nd International Conference for Convergence in Technology (I2CT)*, Mumbai, 2017, pp. 74-78.

[8] Kim, Mooseop & Lee, Deok Gyu & Kim, Ki-Young,” System Architecture for Real-Time Face Detection on Analog Video Camera,” International Journal of Distributed Sensor Networks. 2015. pp.1-11.

[9] M. Fons, F. Fons and E. Canto, "Design of an Embedded Fingerprint Matcher System," 2006 IEEE International Symposium on Consumer Electronics, St. Petersburg, 2006, pp. 1-6.

[10] *Circuit Digest-4x4 Matrix Keypad Interfacing with PIC Microcontroller*[Online]. Available: https://circuitdigest.com/microcontroller-projects/4x4-keypad-interfacing-with-pic16f877a . [Accessed: 3-March-2019]

[11] Y. Chen, Z. Chen and L. Xu, "RFID System Security Using Identity-Based Cryptography," *2010 Third International Symposium on Intelligent Information Technology and Security Informatics*, Jinggangshan, 2010, pp. 460-464.

Design and Implementation of Remote Controlling and Monitoring System for Automatic PLC Based Packaging Industry

Raihan Bin Mofidul
Department of Electrical and Electronic Engineering,
Khulna University of Engineering & Technology,
Khulna-9203, Bangladesh
rbm.eee.kuet@gmail.com

Md. Shahadath Hossain Sabbir
Department of Electrical and Electronic Engineering,
Khulna University of Engineering & Technology,
Khulna-9203, Bangladesh
sabbir7896@gmail.com

Amit Kumer Podder
Department of Electrical and Electronic Engineering,
Khulna University of Engineering & Technology,
Khulna-9203, Bangladesh
amitpodder09@yahoo.com

Mohammad Shaifur Rahman
Department of Electrical and Electronic Engineering,
Khulna University of Engineering & Technology,
Khulna-9203, Bangladesh
msrbitk@yahoo.com

Abstract— **This paper presents an implementation of a human machine interface (HMI) for a programmable logic controller (PLC) based automated packaging process that can be remotely controlled and monitored using mobile application. A conveyor belt does the moving of products to a destined place for packaging, is controlled by Siemens LOGO! 230 RCE PLC. Microcontroller (arduino uno) operates the relays according to the output of IR sensors, which are used to sense the products. States of relays are taken as the input of the PLC. An HMI is designed with an android application and the microcontroller (as the processor) that establish a communication path between the operators and PLC based plant. The interface provides remote controlling and monitoring feature. The proposed design minimized the overall cost by replacing the widely used supervisory control and data acquisition (SCADA) system with normal sensors and mobile application based control. The performance of the implemented prototype is evaluated through several real-time operations, which shows satisfactory results.**

***Keywords*— PLC, Arduino Uno, Android-based application, SCADA, Conveyor belt, Packaging process.**

I. INTRODUCTION

Industrial automation is one of the most global trends that enhances the quality of production, safety and proper utilization of plant and resources. From the economic view, mass production is the main goal of any industry to stay ahead in the recent competitive market [1]. The mass production can be eased through the process automation. In any kind of industry, packaging is a widely used process. Packaging is done for providing protection and safety of products from dust, moistures and other objects that are harmful to the products. The packaging also provides better handling, transporting and, stable storage capability, company branding and many more [2]. The manual packaging process is toilsome work and costly, as a large number of workers is required for it. Therefore, many companies switching to industrial automation for cost-effective and uninterrupted production. Automation of the packaging process with remote control and monitoring system will provide more reliable and developed packaging with less error and less probability of hazards [3]. In industry, an essential equipment of product processing is the conveyor belt. Conveyor belt performs the functions include rotation, accumulation, flipping, stacking, and diversion of products depending upon the application. There are several conveyor belt systems are used in industries. In Table I, numerous conveyor belt systems are enlisted with their working domain [4].

TABLE I
CONVEYOR BELTS AND WORKING DOMAIN

Types	Working Domain
Roller	Belt, rings or chains powered rollers
Sortation	Diverts product one region to another or one belt to another. Mainly used for sorting.
Accumulation	Holds products for a fixed time or for a fixed quantity gathering.
Screw	Consist of helix pushes loose products forward such as powdered or mixed products.
Carousel Indexing	Conveyed products to upward.

In the area of industrial automation, programmable logic controller (PLC) found greater interest for its robustness and low-cost implementation. PLC is a kind of small computer specially designed for industrial purpose, performs several functions including logical operations, counting and arithmetic functions to control various process or machinery through digital or analog modules [5]. Various automation of packaging using PLC and supervisory control has been proposed and implemented by numerous researchers around the world. Each automation has its own pros and cons. A brief discussion has been made on the various automation of packaging processes in Section II. It is found that most of the automation is applicable for large industries, complex in nature, expensive and not affordable for small industries. To overcome the aforementioned limitations, an affordable, less-costly and simple packaging process is proposed. The main contributions of this paper are:

i. For lower capital industries, SCADA and other supervisory devices are used for feedback and

monitoring, which are costly to implement. In this paper, an HMI is implemented with low-cost equipment and easier accessing technology.

ii. Implemented a mobile-based user interface i.e. an android application that is user-friendly and easy to operate and learn.

iii. Developed a platform for improvising and innovation in mobile-based industrial automation with low cost, robustness, and reliability.

In this paper, a prototype of two conveyor belts is considered, one for products carrying to the destined box and another for carrying filled and unfilled boxes. PLC is designed as the controller for a constant speed geared DC motor, which rolls the conveyor belt. An interface is designed based on android platform, arduino uno and bluetooth module to control and monitor the packaging process. The novelty of the proposed system is that it has been found suitable and cost-effective for small industries where capital is small. This work can also be implemented for large industries considering the requirements of robustness and application field.

II. Related Works

In recent years, many works on automation and remote monitoring and control of packaging process have been done. Numerous papers are studied and the contribution have been enlisted, which is represented in Table II.

TABLE II
Related Works on Automation of Packaging Process and Remote Control and Monitoring

Author	Contribution
Marie M. Baroro, Joanna & Alipio et al. [6]	Implemented packaging and material handling using PLC and recommended an HMI for controlling purpose.
Abueejela, Yousef & A. Algitta et al. [7]	Implemented automatic packaging machine by controlling the conveyor belt using PLC.
P.Thirumurugan et al. [8]	Implemented metal separating and packaging machine using PLC and a metallic sensor for detecting metals and non-metals.
Gupta, T & Kamboj, S et al. [9]	Proposed a design of packaging process using PLC and SCADA for remote control and monitoring with a software interface to visualize the process. SCADA and PLC inter-connection has been proposed for re-programming the PLC.
Shyr, Wen-Jye, Te-Jen Su and Chia-Ming Lin et al. [10]	Developed and implemented a methodology for remote controlling and monitoring of plant using PLC. A web-server based computer system and Web-Cam are used to remote monitoring and connected computer through the internet can control the plant from the remote place.
J. Garcia et al. [11]	Using SCADA and TCP/IP developed a platform using Motronic software interface for real-time remote monitoring and controlling of the plant and configure new device to add with, re-configure the existed device.
S. Da'na, A. Sagahyroon et al. [12]	Designed and implemented an automated platform through which the PLC-based plant can be remotely monitored and controlled through TCP/IP or by GSM. They also developed a backup control and monitoring system through mobile-based communication by using SMS.
D. S. V. S. Vardhan and Y. S. Narayan et al. [13]	Using feedback sensors for monitoring and PLC for controlling, an automated control and monitor system has been designed.

From Table II, It has been seen that, robust and more costly remote controlling and supervisory systems are used or proposed for industrial applications. However, for small industries or industries where robustness is not required, these above-mentioned proposed designs seems unnecessary. For this reason, the purpose of this paper is to find alternate technology and reduce cost of automation. Moreover, an automation system has been designed which can be easily implemented for low cost industries and for large industries where only monitoring is required rather than robustness.

III. Methodology

The proposed design of automatic packaging machine using PLC LOGO! 230 RCE can be divided into four steps illustrated as following:

A. Start Packaging System

Tapping the "Start" button on android application, a specific data (e.g. "a") sent from mobile. Bluetooth module (HC-05) receives the transmitted data, which are processed by a microcontroller (arduino uno). Arduino uno makes a particular pin as logical "High" for one second, which will control the switching of relay switch (four-array relay switch 5v DC). This will make arduino uno to send a string e.g., "packaging starts" through bluetooth module to the mobile. This relay switch provides logical information to the respective PLC input (e.g. I1). PLC module then operates its logical operations to close or open respective output relay. Electrical dc motor can be used as an output actuator to move box or sample conveyor belt, as shown in Fig. 1.

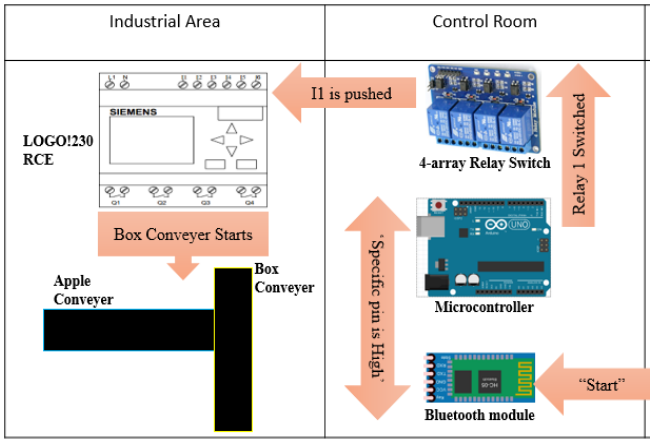


Fig. 1 Proposed procedure of start packaging system.

B. Box Sensing and Counting

Box conveyor run until a box is sensed by the IR sensor. Sensing of a box makes arduino uno to make another pin logical "High" for one second and send another string e.g., "1 box is sensed" to mobile phone including the number of boxes. That pin also controls the switching of a relay switch providing logical information to the respective PLC input (e.g. I4). PLC module stops the box conveyor belt and begins to move sample conveyor belt.

C. Sample Sensing and Counting

Samples (e.g. apple) on the running sample conveyor belt are passed before the IR sensors (mounted on the sample conveyor structure). IR sensors senses and arduino uno process this signal to make a pin logical "High"

for one second. Moreover, arduino send a string to mobile e.g., "1 sample is sensed" and a counter is started to count the sample. The logical "High" pin provides logical information to the respective PLC input (e.g. I3). PLC module stops the sample conveyor belt when three samples have been counted and begin to move box conveyor belt. In such a way, every box is filled with three samples.

D. Stop Packaging System

Box sensing, apple sensing, and counting will be continued until an interruption command is sent from the mobile application. When the "stop" or "exit" button is tapped on the mobile application, the PLC module stops the packaging system and initializes the number of boxes and samples.

IV. Software Implementation

Software implementation includes implementation of PLC ladder diagram, developing an android application and configuring necessary programs (remote control and monitoring system) for arduino.

A. PLC Ladder Diagram

For programming and operation of the presented prototype, a PLC ladder diagram is applied by using LOGO! Soft V8. For a faster, uncomplicated and space-saving solution, LOGO! is a perfect choice. LOGO! Soft Comfort makes possible of easier installation with a minimum wiring effort [14]. This ladder diagram, as shown in Fig. 2 provides two stages of the packaging operation. The input, output, relay switch and counter configuration are shown in Table III.

TABLE III
INPUT, OUTPUT AND RELAY CONFIGURATIONS OF PLC

Input				Output	
I1	I2	I3	I4	Q1	Q2
PLC starts	PLC stops	Apple or sample sensed	Box sensed	Apple or sample conveyor belt	Box conveyor belt
Relays				Counter	
M1		M2		C001	
Relay contacts		Relay contacts		Apple Counter	

1) Carrying Empty Boxes to Desired Location

Flag "make contact" M1 is connected with "make contact" I1 (packaging starts) to provide an OR gate operation. The output of the OR gate and "break contact" I2 (packaging stops) provides an AND logic gate operation to operate flag output coil M1. Flag "make contact" M1 and flag "break contact" M2 provide an AND logic gate operation to operate relay output coil Q2 (box conveyor). Flag "make contact" M2 is connected with "make contact" I4 to provide an OR gate operation. The output of the OR gate and "break contact" C001 provide an AND logic gate operation to operate flag output coil M2. Therefore, once packaging starts (I1 logically high), box conveyor (Q2 relay output coil energies) starts to move unless "packaging stops" (I2 logically high) command is applied or a box is detected (flag M2 is logically high).

2) Packaging Samples into Boxes

Relay output coil "make contact" Q1 is connected with "make contact" I4 and provides an OR gate operation. The output of OR gate, relay output coil "break contact" Q2, the flag "make contact" M1 and sample counter "break contact" C001 provide an AND logic gate operation to operate relay output coil Q1 (apple or sample conveyor). An apple or sample counter C001 is introduced. C001 will be reset if relay output coil Q2 (box conveyor) is logically high. Flag make contact M1, relay output coil make contact Q1 (apple or sample conveyor) and "make contact" I3 provide an AND logic gate operation to count the number of samples.

B. Developing an Android Application

Android-based PLC remote control and monitoring application is developed using Android Studio (emulator) [15]. This application provides a single row display for wireless monitoring and three control buttons, as shown in Fig. 3.

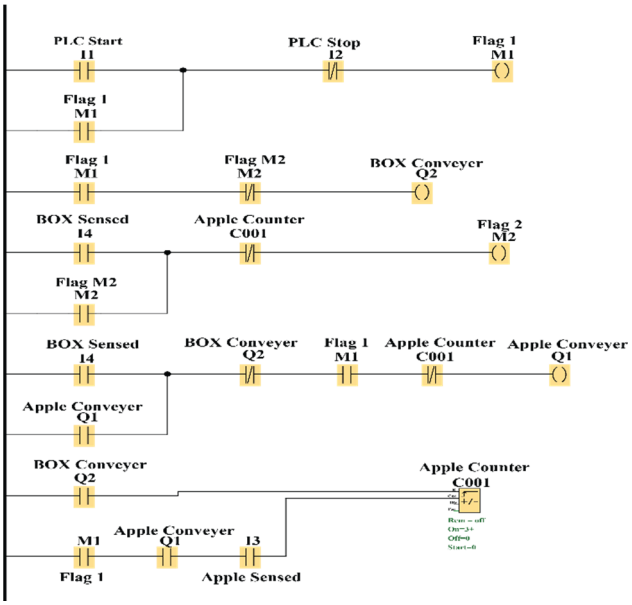


Fig. 2 Ladder diagram of the automatic packaging system.

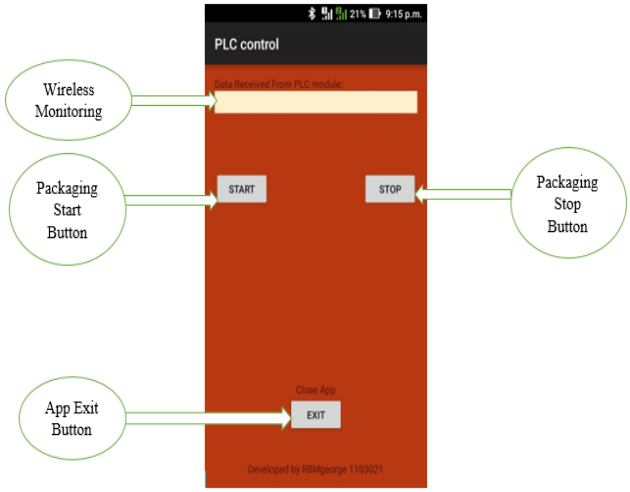


Fig. 3 An android application of automatic packaging & monitoring system.

C. Configuring Remote Control and Monitoring Programs to Arduino

Arduino, an open-source electronics prototyping platform that is based on flexible, easy-to-use hardware and

software. Arduino is a tiny computer system that can be programmed with specific instructions to interact with various forms of input and output and can be reprogrammed as many times depending on the microcontroller used [16]. Arduino uno performs a various function including data exchange between mobile and PLC, counting, calculate and monitor samples and provides necessary data string to the mobile application.

V. Hardware Implementation

The required equipment for complete hardware implementation of automatic packaging and monitoring system is shown in Fig. 4.

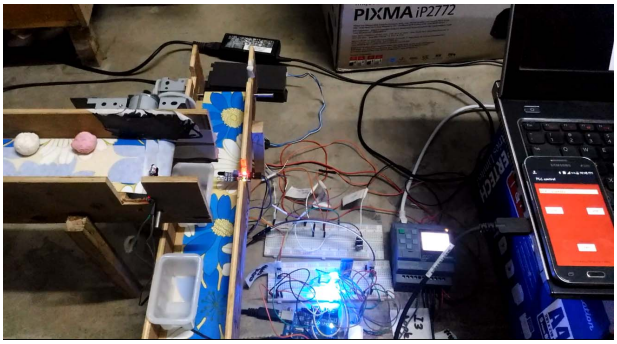


Fig. 4 Overall view of automatic packaging & monitoring system.

The major hardware equipment of the automatic packaging system are the following:

A. LOGO! 230 RCE PLC Module

LOGO! 230 RCE is a small PLC module that is made for home appliances. It provides higher system stability and performance rather than a microprocessor. It comes with easy programming features i.e. ladder diagram that can be programmed using the installation software LOGO! Soft V8. The inputs (220V AC pulse) are provided by a four-array relay switch to the PLC according to the command of arduino uno and the outputs (relay switch) that are connected with the conveyor motor circuits.

B. Arduino, Bluetooth Module and Android phone

Arduino uno is used to communicate between mobile phone and PLC (LOGO! 230 RCE) via bluetooth module. Moreover, IR sensors are directly connected to the arduino uno to generate respective logical input of the PLC module. Data strings provided by arduino uno are shown in Table IV.

TABLE IV
Data Strings Provided By Arduino Uno

Operations in android phone and IR sensors	Data String	
	Data send by arduino	Data receives by arduino
"Start" button	"Packaging Starts"	"a"
"Stop" or "Exit" button	"Packaging Stops"	"b"
Box is sensed	"X box is sensed"	
Sample is sensed	"X sample is sensed"	

C. Conveyor Belts

Two conveyor belts are made from wood, dc motors with proper gearing, belts, and sensors. Belts are used for transporting samples from one location to another, which would be packaged into a specific plastic box later. Through the entire process, dc motors are used as actuators to move the conveyor belts. Proper gearing structure is included for high torque with moderate speed. Two IR (infrared receiver) sensors are used, one for box detection another for sample detection. Detection timing of IR sensors is modified by arduino programming. Even arduino sends necessary data strings to the mobile phone via bluetooth module for every detection of box and sample. The conveyor belt system making procedure is shown in Fig. 5.

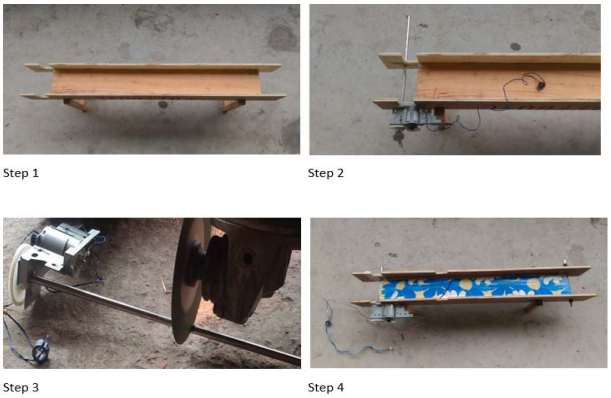


Fig. 5 Making procedure of conveyor belt.

VI. Cost Analysis

An approximate analysis of the cost of equipment used in the proposed design and existed automation is represented in this section. DC motor, conveyor belt are the common equipment in both proposed automation and existed automation. Therefore, their cost is not considered in the analysis. Other equipment that are used in proposed automation and existed automation are compared with respect to their cost and comparison of cost is represented in Table V. It is seen that a reduction of approximately 1700 USD can be possible by using the proposed automation.

TABLE V
Cost Comparison

Proposed Automation		Existed Automation	
Equipment	Cost (USD)	Equipment	Cost (USD)
PLC (LOGO! 230 RCE)	243	PLC (LOGO! 230 RCE)	243
Arduino Uno	24	SIEMENS HMI	700
Bluetooth Module	9	Network Module/ Communication System (TCP/IP, Internet, GSM)	800*
Mobile with Android OS	70	Computer	300
IR sensors	10	CAM/ Sensors	50*
Total	356	Total	2093

*Approximate cost of Network Module and CAM

VII. Conclusion

An automated packaging process with remote control and monitoring capability has been implemented. The main purpose of this design implementation is the reduction in the cost of automation. For lower capital industry or small industries, large automation or high-configured devices are not economical. A reduction of almost 1700 USD become possible using the proposed design. The proposed design provides low-cost HMI to interact with the machine from

remote places using the mobile. A notification or alert system and communication path have been developed that monitoring and control of plant can be done from a distant place from the plant. This ensures the reliability and errorless process of the plant. The proposed design has successfully tested using numerous trial and get a satisfactory result with greater performance of reliability, robustness, and less-error in packaging. This design is feasible for low-capital industries as well as can be implemented for large industries.

REFERENCES

[1] S-L. Jämsä-Jounel, "Future trends in process automation," *Annual Reviews in Control,* vol. 31, no. 2, pp. 211-220, 2007.

[2] N. P. Mahalik, "Advances in packaging methods, processes and systems," *Challenges, MDPI, Open Access Journal,* vol. 5, no. 2, pp. 1-16, Oct. 2014.

[3] T. Samad, P. McLaughlin, and J. Lu, "System architecture for process automation: review and trends," *Journal of Process Control,* vol. 17, no. 3, pp. 191–201, Mar. 2007.

[4] N. P. Mahalik, "Processing and packaging automation systems: a review", *Sensing and Instrumentation for Food Quality and Safety,* vol. 3, no. 1, pp. 12, 2009.

[5] "IEC, International Standard IEC 61131-3: Programmable Controllers Part 3: Programming Languages. International Electro Technical Commission", January 2003, [PDF], Available: http://d1.amobbs.com/bbs_upload782111/files_31/ourdev_569653.pdf. [Accessed: 1 Feb. 2019]

[6] J. M. M. Baroro, M. Alipio, M. L. T. Huang, and T. M. Ricamara, "Automation of packaging and material handling using programmable logic controller," *International Journal of Scientific Engineering and Technology,* vol. 3, pp-767-770, 2014.

[7] Y. M. Abueejela, A. A. Algitta, S. Mustafa, F. Ibrahim, and N. Abdalruof, "Automated packaging machine using PLC," *International Journal of Innovative Science, Engineering & Technology IJISET,* vol. 2, pp. 282-288, 2015.

[8] P. Thirumurugan, R. Aravind, M. A. Kumar, S. D. Manjunath, and R. Kalaiselvan, "Automatic metal separation and packaging using PLC," *IJISET - International Journal of Innovative Science, Engineering & Technology,* vol. 5, no. 2, Feb. 2018.

[9] T. Gupta, and S. Kamboj, "Development of automatic packaging system using PLC and SCADA for industries," *International Journal of Mechanical Engineering and Technology,* vol. 9. pp. 1277-1287, 2018.

[10] W. Shyr, T. Su and C. Lin, "Development of remote monitoring and a control system based on PLC and WebAccess for learning mechatronics", *International Journal of Advanced Robotic Systems*, vol. 10, no. 2, Feb. 2013.

[11] J. Garcia et al., "Reconfigurable distributed network control system for industrial plant automation," *IEEE Transactions on Industrial Electronics,* vol. 51, no. 6, pp. 1168-1180, Dec. 2004

[12] S. Da'na, A. Sagahyroon, A. Elrayes, A.R. Al-Ali, and R. Al-Aydi, "Development of a monitoring and control platform for PLC-based applications," *Computer Standards & Interfaces*, vol. 30, no. 3, pp. 157-166, 2008.

[13] S. V. S. Vardhan and Y. S. Narayan, "Development of an automatic monitoring and control system for the objects on the conveyor belt," *2015 International Conference on Man and Machine Interfacing (MAMI), Bhubaneswar,* pp. 1-6, 2015.

[14] "LOGO!Soft Comfort Online Help- Industry Support Siemens", November 2017, [PDF], Available: https://support.industry.siemens.com/cs/attachments/100782807/Help_en-US_en-US.pdf?download=true. [Accessed: 25 Feb. 2019]

[15] "Android Studio Overview homepage on Android Developers", 2016, [Online], Available: https://developer.android.com/studio/intro/index.html. [Accessed: 20 Feb. 2019]

[16] "Arduino-About Us", 2016, [Online] Available: https://www.arduino.cc/en/Main/AboutUs. [Accessed: 20 Feb. 2019]

Development of an Automated Low-cost Book Scanner and Translator

Sabila Nawshin, Samir Kumar Das, Najia Hossain, Joyita Das Mela, Abu Shafin Mohammad Mahdee Jameel and Salekul Islam
Department of Computer Science and Engineering
United International University
Dhaka, Bangladesh
snawshin161234@bscse.uiu.ac.bd, bipondas516@gmail.com, najiahossain133@gmail.com, joyitamela@gmail.com, mahdee@cse.uiu.ac.bd, salekul@cse.uiu.ac.bd

***Abstract*— Book digitization is the procedure of converting paper books or documents into digital media such as images, text files, electronic books (e-books) and audio books. On this paper we have proposed a low-priced book scanner and translator using Raspberry Pi. We have built a complete prototype that is able to read a book, translate parts of it to English if necessary and scan the whole book in the process to create a digitized soft copy of the book. Thus, a printed book might be converted to an audio book or PDF. For the conversion process, we interface an external webcam with the Raspberry Pi. The pictures taken from a book are processed and converted to text with the Tesseract OCR. Any part, that is not in English, is translated using the Google translator API, then the text is converted to audio file with eSpeak TTS engine. The processed images of the pages are combined and converted into e-book after the whole book is scanned. The entire process was made spontaneous by including automatic page turning mechanism with the hardware implementation.**

Keywords— Image processing, optical character recognition, text-to-speech, book scanner, book translator, book reader

I. INTRODUCTION

On this era of information and technology, digital books are gaining popularity over paper ones. Digital books are easily accessible, more available, easy to store and require little maintenance. Most of the recently published books now provide digital version as well. In a paperless world, almost all documentation is now carried out by storing data in digital format. The problem that arises is with the older books and documentations that only has printed copies. We are faced with the tedious task of converting those numerous books and documents into e-books and audiobooks (where necessary). Major steps of the conversion process include scanning individual pages, processing the scanned images, converting it to text and then to audio (for audio book), combining and converting the processed images of all the pages to a single e-book (e.g., in PDF format) or audio book (e.g., in MPEG format). Various commercial companies have come forward to solve the problem and have built book scanners that are able to automatically convert a given paper book to a digital one. The price of a document scanner (without automatic page-turner) that is presently available in the market is too high starting from 200 to few thousands USD. In this paper, we are presenting a low-cost alternative to the existing book scanners with some additional features.

Optical Character Recognition (OCR) is the process of identifying text characters from printed or scanned text images and converting it to editable text [1]. The Tesseract OCR is an open source optical character recognition engine developed at HP in between 1984 to 1994 [2]. It has since been modified and improved for greater accuracy for the few last decades. It was made open source in the late 2005 and is now maintained by Google. It provides support for various languages such as English, Japanese, Hindi, Spanish, French, German, etc. The Tesseract OCR has been used to convert scanned book pages to editable text files in our project.

The artificial production of human speech is called Speech Synthesis [3]. Text-to-speech (TTS) Engines are a form of speech synthesis which were first developed to aid the visually impaired by offering them spoken words with text. A TTS engine has two parts, front-end and back-end (Synthesizer) [4]. The front-end converts the text to symbolic linguistic representation which is then converted to sound by the back-end [5]. In our project, the text file prepared by the OCR has been converted to audio using the eSpeak TTS engine.

There are several available book scanners in market at the moment, like CZUR ET16 Plus smart book scanner, FUJITSU Image Scanner ScanSnap SV600, Aibecy Adjustable High Speed USB Book Image Document, Aibecy Portable High Speed Document Camera Scanner etc. [6–8]. Most of the book scanners are not available in Bangladesh and the cost of the scanners (including shipping) is higher than what the general people can afford. We are presenting a low-cost alternative for the book scanners that can be built from easily available ingredients. Our scanner has added features like translation and audio book conversion that none of the existing products offer and production cost of the scanner is below 70 USD.

The rest of the paper is organized as follows: section II describes the proposed method, section III describes the software development process, section IV presents the hardware implementation process and section V provides a comparative analysis (between existing products and the proposed one) and discusses the humanitarian impact of this project. Finally, section VI concludes the paper by mentioning the future work we have planned.

II. PROPOSED METHOD

At first, an image is taken using the camera interfaced with a small single-board computer. The image is then processed and supplied into the OCR software to convert it to a text file. The text file is supplied to Google translator API to translate the necessary parts. Then the processed text is read using TTS engine and is recorded for the audio book. The page is then turned using the page turning mechanism implemented on the hardware. The process is repeated until all the pages are done. After that, the images of the pages are converted into a single PDF file. Fig 1 describes the process through a simple flow chart.

For the software implementation, Raspberry Pi 3 (Model B) with Raspbian Operating system has been used as the single-board computer. Raspberry Pi 3 (Model B) has been

chosen for the implementation because of its processing power (1.2 GHz 64-bit quad core processor) and on-board WiFi capabilities [9]. The processing power is needed to use the OCR and TTS engine and the WiFi capabilities has made it possible to use Google translator API, which needs Internet connectivity to operate.

We have used the Python 3 programming language for developing the underlying software part. The OS library has been used to execute the commands that operate the OCR and TTS.

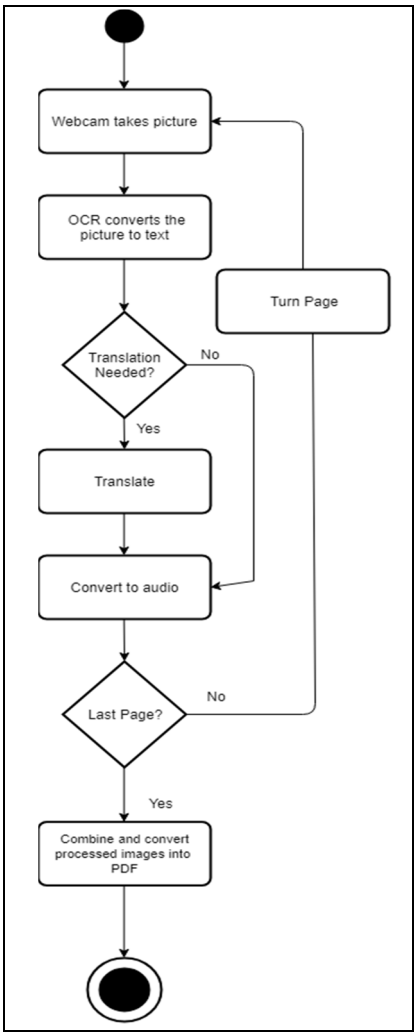


Fig.1 Working procedure

III. DEVELOPING THE SOFTWARE

For the successful operation of the system, we need to develop both software and hardware parts. The hardware or the mechanical part is controlled by the software and thus, the synchronization of their operation is crucial. Some major software used are the Tesseract OCR software to convert the image to editable text for processing, the eSpeak TTS engine to convert the text to audio, and the ImageMagick software to convert the images to e-book.

Tesseract OCR has been used on our project to convert the scanned pages to text for further processing. When an image with text is given to the Tesseract OCR, it is first converted into binary Images from which the character outlines are extracted using connected component analysis [2]. Then, Character outlines are converted into blobs that are organized into text lines [2]. The lines are converted into words and then the recognition of the words start as two-pass process [1]. An endeavor has been made in the first pass to recognize the words from text. Next, the successfully recognized words are passed to an adaptive classifier [2]. Adaptive classifiers are algorithms used to classify unclassified data using an adaptive iterative process that allows to learn from the input data (Called training data). The correctly classified words from the first phrase are used as the training data for the adaptive classifier in the second phrase. The trained classifier is then used to effectively recognize the remaining words from the image [2]. Fig. 2 describes the whole process in a simple flow chart.

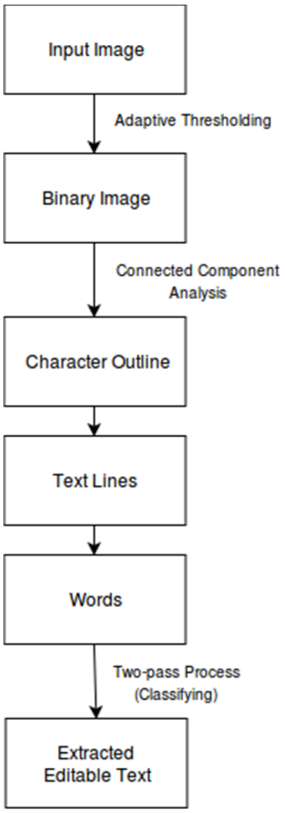


Fig. 2 OCR working procedure flow chart

To convert a paper book to an audio book, text-to-speech functionalities are required in our project. The editable text, prepared by the Tesseract OCR, has been used as input for the eSpeak TTS engine to convert it to speech, which in turns become an audiobook. The eSpeak TTS engine contains two parts: the front-end and the back-end. The front-end converts raw text to equivalent written-out words, then assigns phonetic transcriptions to each word and finally divides and marks the text into prosodic units such as phrases, clauses, and sentences [3]. These are also called symbolic linguistic representation. The back-end (called Synthesizer) then converts these symbolic linguistic representations prepared by the front end to appropriate sound [3]. Fig. 3 describes the process through a block diagram.

For paper book to eBook conversion process, the images of the pages are processed slightly and saved in a common directory. When images of all the pages are taken and processed, they are combined and converted to a single PDF

using ImageMagick software, which contains the functionality needed for the conversion.

The OCR software and the TTS engine are manipulated in our project using the OS library available on Python 3 programming language. It provides needed functionalities to execute commands (similar to the command prompt of an Operating System) provided in the form of string from inside the code. We have used this to make the whole conversion process spontaneous by executing the necessary commands on some simple loops and conditions. The pseudocode of the program is given in Algorithm 1.

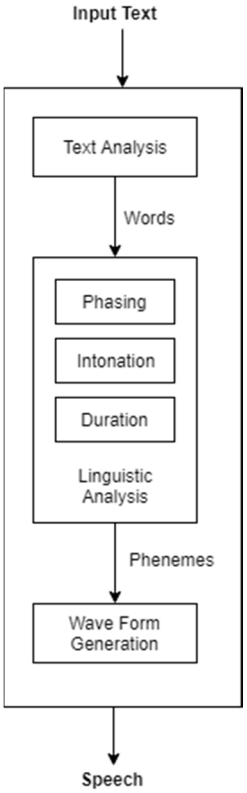


Fig.3 TTS engine working procedure

Algorithm 1: Pseudocode of the developed program

```
1.  Initialize N to total number of pages in the book
2.  for i = 0 to N do
3.      Command OS to take a picture and save it in directory D
4.      Command OS to use OCR to convert the picture to text
5.
6.      Text = read text file into string
7.      new_Text = Text, after being translated by Google
                   Translator API
8.
9.      Audio = Audio returned TTS with new_Text as input
10.     Save Audio on directory D1
11.
12.     Instruct DC and servo motors to move (to turn page)
13.     Instruct DC and servo motors to return to initial position
14.
15. Command OS to use ImageMagick to convert images saved
                   on D to a single PDF
16. Audio_book = Merged Audios from directory D1
17. Save PDF and Audio_book in home directory
```

IV. HARDWARE IMPLEMENTATION

Since the design goal of the project was to develop a low-cost solution, we have carried out the hardware implementation using easily available parts in an affordable price. The steps we have followed are:

1) We have made a box like structure out of cocksheet and fixed a stand on it.
2) The wires and the Raspberry pi has been fixed within the box.
3) An adjustable webcam (1.0 megapixels) has been fixed on the stand to take images of the book.
4) For the automated page turning mechanism, we have used a thin plastic board structure with a small wheel (with a DC motor to operate it), servo motor with an extended head and a motor driver to control the motors from the Raspberry pi.

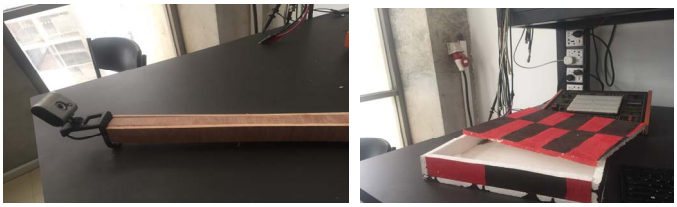


Fig.4 a. (left): The stand with the webcam, b. (right): The box like structure on which the stand was fixed

Fig. 4(a) shows the stand with the adjustable webcam, Fig. 4(b) shows the box like structure, Fig. 5(a) shows the page turner and Fig. 5(b) shows the assembled project.



Fig.5 a. (left): Page turner, b. (right): Assembled Project

The page turner on our book scanner contains two rectangular plastic boards, the first board is fixed on a movable structure and contains a servo motor with an enlarged head. The structure can be moved in accordance to the book's length. The second board is on the top of the first one and can be moved upward or downwards. It can be adjusted according to the book's height. A small wheel is fixed on this part. To turn a page, the page turner first turns the wheel on the direction opposite to the direction at which the page is bound. This causes the page on the front to go up a bit and then the head of the servo pushes the page away from that height, successfully turning it. Fig. 6(a), Fig. 6(b), Fig. 6(c) and Fig. 6(d) demonstrate the process sequentially.

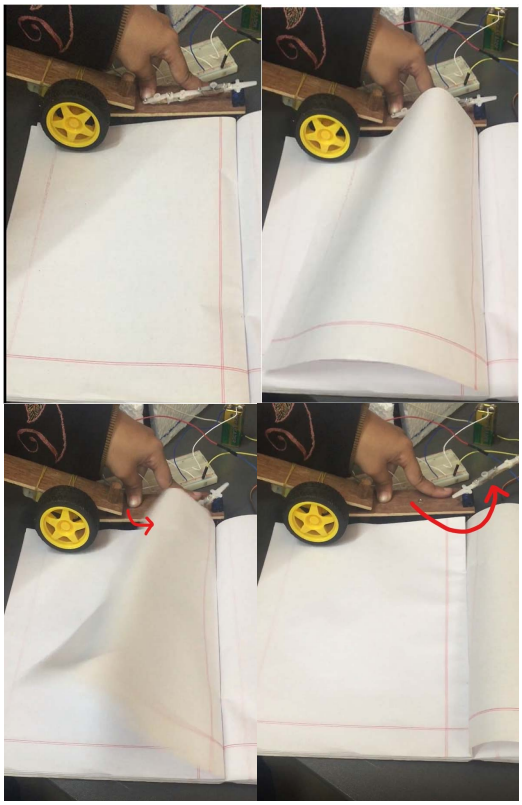


Fig.6 a. (upper-left): Initial state, before turning the page, b. (upper-right): After the wheel was turned, c. (lower-left): After servo starts moving in the direction of the arrow, d. (lower-right): After the servo is done moving and the page is turned

V. Discussion

A. Comparative Analysis

To build our project, the needed elements are Raspberry pi, 8GB micro SD card, an old webcam, a small wheel, a DC motor, a servo motor, 9V batteries, few thin plastic boards and a cocksheet. The Raspberry Pi costs 30 USD, the micro SD card costs 10 USD, the old webcam costs under 5 USD and the rest of the elements sum up to around 15 USD [9]. In total, the production cost for this project is around 60 USD. Most of the book scanners are not available in Bangladesh. Many from the renowned brands do not ship to Bangladesh as of date. Out of the few available ones, the CZUR ET16 Plus CZUR Book Scanner costs 400 USD, the Fujitsu Image Scanner ScanSnap SV600 costs around 500 USD and the Aibecy Adjustable HD High Speed USB Book Image Document Camera Scanner costs 175 USD, which cannot be afforded by smaller organizations and general people in Bangladesh [6–8]. Our implementation of the book scanner can be marketed under 100 US dollars with some fine tuning and further improvements on speed. It should be noted that none of the mentioned book scanners provide automated page turning mechanism and translation facilities.

B. Humanitarian Impact

Reading is a challenging task of the visually impaired people. Our solution provides conversion of paper book to audio book, and thus can also be used as a book reader. The price range of the book scanner is within the affordability of general people. The visually impaired people also can easily afford it. With little help from others, the book scanner can be set up and once the book is given, it will not need further human supervision. Due to its automated features, the book would be read page by page while it is being converted to an audio book. The converted audio book can also be stored and played at later times without the paper book, making it possible to build a collection of books for the visually impaired persons. It can also be used to read to children who has yet to learn reading. The low cost and the affordability of the book scanner and translator enables it to serve blind people from across the world.

VI. Conclusion and Future Work

We have built a low cost Raspberry pi based book scanner and translator that can convert paper books to e-books and audio books. The whole process is automated and does not need human supervision aside from the initial setup. The working procedure of the book scanner is simplistic, and the hardware part can effortlessly be built from easily available elements. Given a paper book, the pages are sequentially scanned (pictures are taken and then are processed) and converted to audio. Future work will include improvements on the scanning speed and quality of the audio and scanned images. The scanned images are combined and converted to PDF using another software that includes many features that we have not explored for this project. The project can be further improved by customizing the open source software to include the PDF conversion function. The image processing for the PDF conversion is done manually now, we plan on using the OpenCV library on python to make it automatic. Though our presented project has many places to improve on, it can serve its purpose as an book scanner in its current state. It is affordable and is specially designed for the use of people in low-income countries such as Bangladesh where book scanning technologies are still out of reach for the general people.

Acknowledgements

The authors would like to thank the Department of Computer Science and Engineering, United International University (UIU) for providing necessary experimental platform for research and analysis for the completion of the work

References

[1] Patel, Chirag, Atul Patel, and Dharmendra Patel. "Optical character recognition by open source OCR tool tesseract: A case study." *International Journal of Computer Applications* 55.10 (2012): 50-56.

[2] R. Smith, "An Overview of the Tesseract OCR Engine," in *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, Parana, 2007, pp. 629-633

[3] Van Santen, Jan P. H., Sproat, Richard W., Olive, Joseph P., Hirschberg, Julia, *Progress in Speech Synthesis* , Springer, 1997

[4] Paul Taylor, *Text To Speech Synthesis*, New York: Cambridge University Press, 2009

[5] John Holms and Wendy Holms, *Speech Synthesis and recognition*, New York: Taylor and Francis, 2001

[6] "*CZUR ET Smart Book Scanner,*" Accessed on: Feb. 04, 2019. [Online]. Available: https://www.czur.com/product/et16/

[7] "*FUJITSU Image Scanner ScanSnap SV600,*" Accessed on: Feb. 04, 2019. [Online]. Available: http://www.fujitsu.com/global/products/computing/peripheral/scanners/scansnap/sv600/

[8] "*Aibecy Adjustable HD High Speed USB Book Image Document Camera Scanner,*" Accessed on: Feb. 04, 2019. [Online]. Available:

https://www.aliexpress.com/item/Adjustable-HD-High-Speed-USB-Book-Image-Document-Camera-Scanner-Dual-Lens-10-Mega-pixel-2/32875202324.html

[9] "*Raspberry Pi, Raspberry Pi Model B,*" Accessed on: Jan. 20, 2019. [Online]. Available: https://www.raspberrypi.org/products/model-b/

A Strategic Routing Analysis for Agro Sensor Communications in Mobile Ad hoc Networks

Sampad Banik[1], Md Munjure Mowla[1], and Iftekhar Ahmad[2]
[1]Department of ETE, Rajshahi University of Engineering & Technology, Bangladesh
[2]School of Engineering, Edith Cowan University, Australia
Email: sampad.banik530@gmail.com[1], rimonece@gmail.com[1], i.ahmad@ecu.edu.au[2]

***Abstract*—Recently, agro sensor communications are getting popularity with the advent of wireless sensor networks (WSNs) and mobile ad hoc networks (MANETs). Infrastructure less MANETs are considered as a potential solution for agricultural domains to provide robust and effective operations in wireless networks by enabling routing characteristics at each mobile node. The efficient use of routing protocol is pivotal for network management in MANETs. In this research, we demonstrate a strategic routing analysis to evaluate the performance of different routing protocols, e.g., proactive and reactive in terms of various network quality of service (QoS) parameters for multiple data rate enabled agro sensor mobile nodes. Two different network scenarios (low-node and high-node) are simulated for various network parameters. Simulation results indicate that OLSR protocol provides higher throughput than others for all nodes. This investigation also shows that GRP has the highest network load capacity than other protocols for agro sensor applications. Another finding from this research is that proactive protocols have the lower average end to end delay and MAC delay than reactive protocols.**

***Index Terms*—Wireless sensor networks (WSN), Mobile ad hoc networks (MANET), Routing protocols, Agro sensor, Traffic Sources.**

I. Introduction

Agriculture is one of basic priorities in any country to ensure food security for peoples' lives. The proper maintenance of the agricultural field is required to increase the productivity. However, in developing or under-developed countries, agriculture mainly relies on some natural and environmental functionalities such as seeding, harvesting, irrigation, flood, drought, etc. Random climate changing, limited cultivation capacity, and other stress factors affects agricultural production heavily. In these respects, flood risk management and prediction, real-time control and treatment of crops provided by technology, i.e., wireless sensor networks (WSNs) can be a solution for better production by using improved agro-climatological estimation and smart sensors monitoring [1]. Although there are a range of different applications of WSNs such as situation realization, sensing intruders, identification of rival's movement on land and marine area, unmanned aerial networks, and surveillance of battlefields; this network is also intended for monitoring an agricultural environment for transmitting and receiving information from a particular domain, process them and throw it to the sink. The collaboration of different nodes is a requirement to ensure the communication between sender and receiver through intermediate nodes. Due to the limitation of installing regular infrastructure-based network models in agriculture fields, mobile ad hoc networks (MANETs) are considered as an appropiate style of WSNs to sense, gather, distribute, synthesize, and monitor environmental risks [2]. The increasing demands of various applications (crop surveillance video, crop production, weed, insect, nutrient management, field guide, pest management, diseases, etc) in agricultural domains motivate us to evaluate the impact of multiple data rates (CBR, VBR, and FTP) enabled nodes on transferring information through a MANET. Interested readers are encouraged to go through [3] for the network analysis of multiple data rate applications in case of wireless mobile communication.

In this research, the performances of different routing protocols (e.g., proactive and reactive) in MANETs are evaluated for multiple data rates enabled sensors to be used for agricultural fields. We consider an agricultural field, where ten and twenty sensor nodes (any quantity) are operating in battery power. The sensor nodes are put in a fixed distant area of an agricultural field for measuring the moisture and PH of land. For coordination of each node in this distributed MANET environment, the routing is the vital concern. This research work mainly focuses on obtaining an optimistic result in perspective of different network parameters (e.g. throughput, end to end delay, network load, MAC delay, re-transmission, jitter, and packet drop) either by using reactive or proactive routing protocols for multiple data rates enabled sensors.

Due to dynamic changes of network topology in MANETs, routing protocols need to adapt quickly. The routing protocols of MANETs are broadly divided into two categories: reactive routing protocols and proactive routing protocols. Reactive protocols does not start the a route discovery process until a destination route is required, for example, Ad hoc On-Demand Distance Vector (AODV) and Dynamic Source Routing (DSR). On the other hand, in proactive protocols, tables are generated to portrait the whole network topology. These tables are revised systematically to maintain an update routing information for all nodes, for example, Optimized Link State Routing Protocol (OLSR) and Geographical Routing Protocol (GRP). However, reactive routing protocols work differently compare to proactive routing protocols. The evidence from this study suggests that AODV protocol results higher throughput for all cases. This investigation also shows that DSR has the highest network load capacity than other protocols. Another

finding from this research is that proactive protocols have the lower average end to end delay and MAC delay than reactive protocols in all cases.

Paper organization is as follows. Section II mentions the related works. Section III illustrates the simulation environment. Simulation results are discussed in Section IV. Section V concludes this paper.

II. Related Works

As a prospective tool towards the development of agriculture, WSNs are simplified by the universal characteristics of function, self-organized small nodes, adaptable and cost-efficient technology.

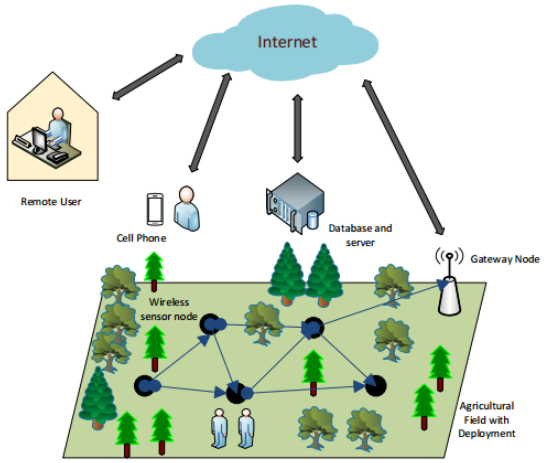


Fig. 1: Agro Sensor Deployment in a typical MANET.

Fig. 1 shows a typical MANET deployed for agricultural applications. Akyildiz highlighted the applications of WSNs to boost conventional farming [4].

Ojha reviewed the dynamic applications of WSNs and associated challenges for WSNs deployment to enhance farming [1]. Some of the popular applications of WSNs in agricultural domains are enlisted as: irrigation arrangement system, husbandry systems governing, disease and pest control, controlled use of fertilizers, cattle movement superintendence, ground water quality managing, greenhouse gases monitoring, remote control and recognition etc. Shiravale proposed a WSN for agricultural which helps the farmers in collecting weather data from meteorological department, planning for the right crop, communicating with agricultural experts [2]. Panchard explored the potentiality of WSNs in the context of agricultural fields for small landowners and farmers while they faced water shortage problems [5]. Author in [6] introduced an IEEE 1451 based framework which works on the acquisition of smart data in precision agriculture. Le demonstrated a WSN for precision agriculture in which the farmers can superintend the agricultural andenvironmental parameters [7]. Lightweight unmanned aerial vehicles (UAVs) were used to capture videos sensing which acts like a real-time precision sprayer [8]. In precision farming and ecological governing, a multi-level sensor network was designed for data acquisition and processing [9]. In a similar research, heterogeneous WSNs are used for a flood prediction scenario to make decision support systems [10].

MANETs are suitable alternative solutions of WSNs where regular infrastructure setup is difficult or expensive. A small agriculture field can be used as a MANET as the large sized network mainly consists of terrestrial WSNs or wireless underground sensor networks [1]. Routes between two communicating nodes may consist of multiple paths through other nodes in the network. The information passed through these mobile nodes causes existing links to overflow, discontinue or establish new routes. Thus, finding and maintaining routes in MANETs are challenging tasks. In this context, performance of routing protocols plays a pivotal role in MANETs. Several research have been carried out to assess the performance of routing protocols in MANETs or WSNs for various applications [11]–[17]. Some of the popular simulation are generally used such as Optimized Network Engineering Tool (OPNET), Network Simulator 2 (NS2), OMNET++, and QualNet.

Different types of routing protocols were compared in MANETs for CBR traffic [11]. Thool investigated the performance of DSR and DSDV routing protocols using different network conditions under NS2 simulation scenario. They analyzed performances with respect to packet delivery, end to end delay, and throughput [12]. Kuppusamy studied the features of adhoc routing protocols TORA, OLSR, and AODV based onthe performance metrics such as PDR, end to end delay, routing overhead by increasing number of nodes in the network [14]. Al-Baseer presented a comparative study for DSR, AODV, OLSR and DSDV routing protocols using OMNET++ simulator [18]. Author in [15] highlighted an overview of OPNET Modeler architecture by implementing a new routing protocol. Hussen analyzed the performance of MANET routing protocols for unmanned aerial vehicle networks in terms of end to end delay, throughput, and network load using OPNET modeler [17].

Although, numerous research have been conducted on different routing protocols in MANETs or WSNs for various application scenarios, few of them only focused on routing protocols on agro sensors or agricultural applications. Farjana *et al.* evaluated the MANET routing protocols in an small sized agro sensors network. They analyzed the performance of routing protocols in respect to some network parameters such as average end to end delay (average), network load, and throughput using OPNET software [19]. In a similar research, Siddiquee investigated the optimal dynamic routing protocol for agro sensor communications for two different sized networks; large (20 nodes) and small (10 nodes) by using both OPNET and NS2 while evaluating the same parameters [20]. In their study, they investigated that AODV routing protocol shows enhanced results in large size network and for a small size network, DSR performs better. Both of these papers [19], [20] consider MANET nodes can deliver only FTP traffic. *However, to the best our knowledge,* no research has been carried out on the impact of multiple traffic (CBR,VBR, and FTP) enabled mobile nodes while considering different

routing protocols of MANETs for agro sensor applications. This paper fills the gap and provide a strategic routing analysis.

III. Simulation Environment

We consider an agricultural field as a MANET, where 10 or 20 nodes are deployed and all are coordinating each other shown in Fig. 2 and Fig. 3. A destination node is used as a gateway to connect with the Internet. Simulation aids in analyzing the performance and behavior of communication networks before implementing it as a real application. We have investigated this network scenario using a most popular networking simulation software, e.g., OPNET.

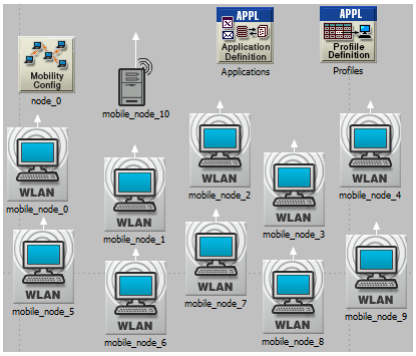


Fig. 2: Simulation setup for 10 nodes.

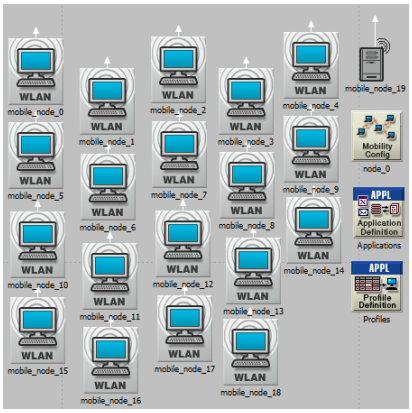


Fig. 3: Simulation setup for 20 nodes.

For this research, we used a popular network simulator, i.e., OPNET Modeler version 17.5 Academic version (Riverbed) for evaluate the network performances. Simulation experiments were performed under three applications (FTP, CBR, and VBR) scenarios that have been created for IPv4. All mobile nodes are enabled with three types traffic. Table I shows the simulation parameters. OPNET Modeler [21] is a adaptable and efficient tool which supports discrete event simulation, statistical analysis, and analytical modeling for the communication networks and distributed systems. Two different network scenarios are considered for this research. We used 10 nodes for the low density network (first scenario) and 20 nodes for the high density network (second scenario). Network Size is fixed on 1000 $\times$ 1000 meters.

TABLE I: Simulation Parameters

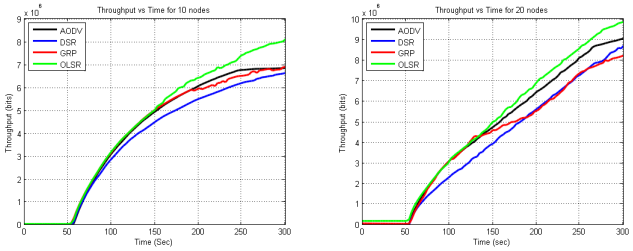
Parameter	**Value**
Simulator	Riverbed OPNET Modeler 17.5
Area	1000 m x 1000 m
Network size	10 nodes, 20 nodes
Protocols	AODV, DSR, OLSR, GRP
Mobility Model	Random Way Point
Speed	5 m/s
Trajectory	vector
Simulation time	300 s
Address mode	IPv4
Traffic Source	File transfer protocol
	Constant bit rate
	Variable bit rate

AODV is a routing protocol that works on on-demand route acquisition algorithm. It conducts two operations, i.e., route discovery and route maintenance. It diminishes required broadcasts number by forming routes in an on-demand scheme. AODV assigns three key route messages: Route Requests(RREQ), Route Errors (RERR) and Route Replies (RREPs) to forward data packets from sender to receiver [22]. DSR is a routing protocol which dynamically discovers a route. In source routing, the header of the packet contains all the information regarding the route as well as the information of nodes from end to end which the packet may pass, this approach guarantee loop free routes for information transfer [23]. OLSR is a link state proactive IP routing protocol optimized to work for MANETs that discovers and distributes link state information of the entire topology of the network to the other nodes in the network. The key feature of OLSR is that it basically uses a Multi-Point Relaying (MPR) to optimize flooding [24]. GRP is proactive routing protocol, follows hop by hop routing type and make periodical frequency updates. In order to keep the information about the neighbor nodes Hello Messages are spread-ed periodically. GRP forwards packets to the destination regarding shortest distance, a node receive a packet send the closest node the packet following the destination route.

IV. Results and Discussion

In this section, various performance parameters of network will be discussed from simulated results.

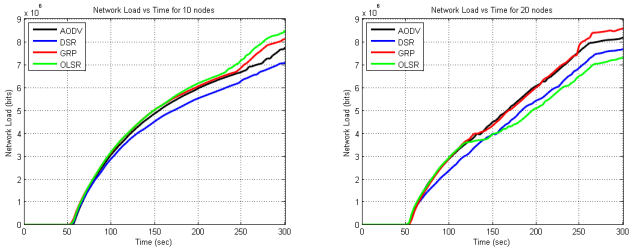
The network *throughput* is defined as the ratio of successfully transmitted information to the time taken for transferring this information. In MANETs, due to mobility of nodes, power constraints, limited bandwidth and traffic load, throughput is immensely affected by the dynamic characteristics of the topology. The discussion of the results begin with the analysis of throughput for four routing protocols (AODV, DSR, GRP, and OLSR) depicted in Fig. 4 in case of low density (10 node) and high density (20 node) networks. OLSR protocol attains higher throughput for both scenarios, whereas DSR protocol has lowest throughput. Simulation was run for 300 s and the average values of throughput were calculated for each case while considering different types of data rate applications.



(a) 10 nodes (b) 20 nodes

Fig. 4: Throughput.

Network load handles the total load sent to WLAN layers. A large amount of incoming traffic load is difficult to maintain and highly affects the delivery of routing packets in MANETs. This also increases the collision of these control packets. In

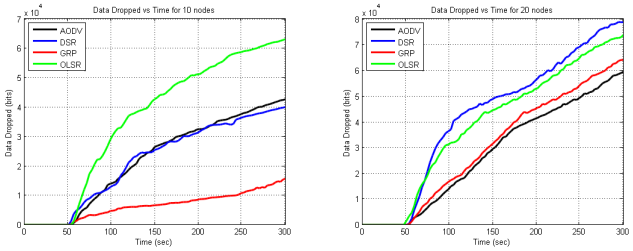


(a) 10 nodes (b) 20 nodes

Fig. 5: Network Load.

case of network load, Fig. 5 exhibits that OLSR has the highest value while AODV and GRP result almost equal (10 nodes). DSR shows the lowest load in this case. With the increment of mobile nodes (20 nodes), GRP provides the highest network load while OLSR results the lowest load. The average network load was calculated from all simulation scenarios.

Data dropped (Packet loss) is the amount of lost packets while transferring data from source to destination in MANETs. This parameter informs the higher layer dropped packets as MAC failed to receive any ACKs for transmitting these packets or segments within a short retry threshold.

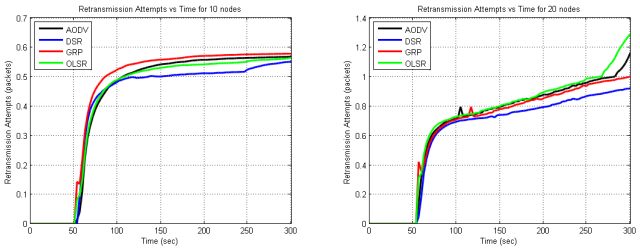


(a) 10 nodes (b) 20 nodes

Fig. 6: Data Dropped.

In the first simulation scenario (10 nodes) portrayed in Fig. 6, it is observed that the highest amount of data dropped is caused by OLSR protocol and less number of packet lost is provided by GRP protocol. For the large network (20 nodes), AODV provides the lower number of packet loss. In case of both scenarios, OLSR performs worst among all protocols.

In a network, the *re-transmission*is required to resend the packets which are already lost or damaged. Network protocols use this basic mechanism for a packet switched communication network to provide reliable communication. Figure 7

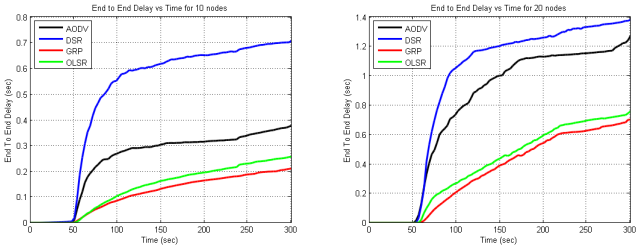


(a) 10 nodes (b) 20 nodes

Fig. 7: Re-transmission.

portrays that DSR has the lowest re-transmission attempts compare to other protocols for both scenarios. The behavior of other protocols fluctuate for high density (20 node) networks.

End to end delay means the transfer time for a packet traveled from the source to target within a network. It considers processing, queuing and propagation time of the link.



(a) 10 nodes (b) 20 nodes

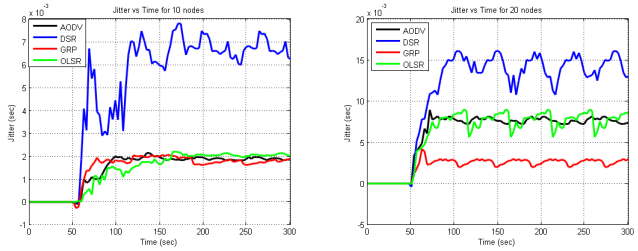
Fig. 8: End to End delay.

The comparative analysis of average end to end delay among four different protocols are illustrated in Fig. 8. GRP protocol shows the lowest end to end delay for small number of nodes (10 nodes) while maximum delay are observed for DSR protocol. For scenario 2 (20 nodes), DSR protocol shows the highest delay while lower delay are observed from both proactive protocols. In case of average end to end delay, DSR performs worst among all protocols for agro sensor communications while considering different type of data rate applications. The proactive protocols shows enhanced results than reactive protocols in case of average end to end delay.

Sometimes, timing differences (*Jitter*) are observed in the network,i.e., signal are deviated at the receiving end compared to sending end packets. This is happened mainly for network congestion and timing drift and it is expected the minimum value.

TABLE II: Average Values.

Protocols	AODV		DSR		OLSR		GRP	
Performance Parameters	**10 Node**	**20 Node**	**10 Node**	**20 Node**	**10 Node**	**20 Node**	**10 Node**	**20 Node**
Throughput (Kbps)	387.52	916.34	262.45	539.98	294.43	990.17	302.30	811.03
Network Load (Kbps)	74.15	110.25	95.26	106.25	84.75	119.67	84.92	110.72
Data Dropped (bps)	14164	3685	14807	7708	24024	8301	6332	5757
Re-transmission (Packets)	0.4169	0.6714	0.3984	0.6102	0.4132	0.6948	0.4381	0.6555
End to End Delay (s)	0.2632	0.5647	0.5554	0.7061	0.0848	0.0591	0.0812	0.0432
Jitter (s)	0.0014	0.0061	0.0049	0.0111	0.0014	0.0061	0.0014	0.0022
MAC Delay (s))	0.2134	0.9162	0.4672	1.1423	0.1286	0.5754	0.1056	0.6701

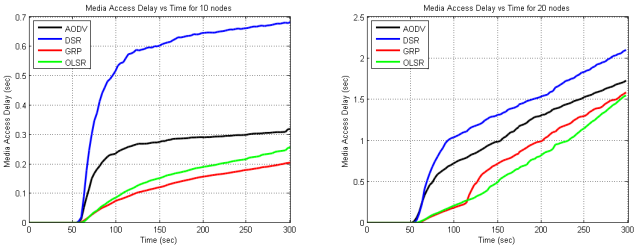


(a) 10 nodes (b) 20 nodes

Fig. 9: Jitter.

Figure 9 illustrates that DSR protocol results highest jitter for both simulation cases. GRP protocol performs lowest jitter in high density (20 nodes) network.

Media access delay is elucidated as the delay of the packets to arrive the MAC layer until it is effectively transmitted out on the wireless means. This parameters is helpful for many real-time applications which have a certain limit of delay. Consequently, it is vital to exhibit low delay for real-time flows.



(a) 10 nodes (b) 20 nodes

Fig. 10: MAC Delay.

From the Fig. 10, it is evident that DSR protocol provides higher MAC delay compare to other protocols in both scenarios. The proactive protocols GRP and OLSR exhibits lower delay than reactive protocols. The average data of four protocols for all above parameters is mentioned in Table II.

V. Conclusion and Future Work

In this research, we evaluated the performance of four different routing protocols in MANET for agro sensor communications. The topological and management issues in agricultural field lead to choose MANET with WSN to assess the performances of protocols. Though several researches have been conducted in comparing the routing protocols, but the application scenario and multi data rate enabled sensor nodes differs our contribution from those. We draw a strong scenario for various number of nodes (10 and 20) with a powerful simulation environments (OPNET) to highlight different performance characteristics such as throughput, end to end delay, network load, jitter, MAC delay, re-transmission attempts, and packet loss. This study has shown that proactive protocols have the lower average end to end delay and MAC delay than reactive protocols for all network size. Another observation is that OLSR protocol provides higher throughput and GRP provides higher network load than other protocols in case of agro sensor communication. Future research will consider the massive density of nodes such as thousand sensors may be mounted into a MANET. In addition, we will focus the cost and energy optimization for routing in these protocols for better QoS applications.

References

[1] T. Ojha, S. Misra, and N. S. Raghuwanshi, "Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges," *Computers and Electronics in Agriculture*, vol. 118, pp. 66 – 84, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168169915002379

[2] S. Shiravale and S. Bhagat, "Wireless sensor networks in agriculture sector-implementation and security measures," *International Journal of Computer Applications*, vol. 92, no. 13, 2014.

[3] M. M. Mowla, I. Ahmad, D. Habibi, and Q. V. Phung, "A green communication model for 5G systems," *IEEE Trans. Green Commun. Netw.*, vol. 1, no. 3, pp. 264–280, Sept. 2017.

[4] I. F. Akyildiz and I. H. Kasimoglu, "Wireless sensor and actor?we refer to entities that can act on the network as actors they are sometimes referred to as actuators in related literature.? networks: research challenges," *Ad Hoc Networks*, vol. 2, no. 4, pp. 351 – 367, 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1570870504000319

[5] J. Panchard, "Wireless sensor networks for marginal farming in india," *Lausanne: Ecole Polytechnique Fedérale de Lausanne, Suisse*, 2008.

[6] M. A. Fernandes, S. G. Matos, E. Peres, C. R. Cunha, J. A. Lpez, P. Ferreira, M. Reis, and R. Morais, "A framework for wireless sensor networks management for precision viticulture and agriculture based on ieee 1451 standard," *Computers and Electronics in Agriculture*, vol. 95, pp. 19 – 30, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168169913000756

[7] T. D. Le and D. H. Tan, "Design and deploy a wireless sensor network for precision agriculture," in *Information and Computer Science (NICS), 2015 2nd National Foundation for Science and Technology Development Conference on*. IEEE, 2015, pp. 294–299.

[8] C. Cambra, J. R. Díaz, and J. Lloret, "Deployment and performance study of an ad hoc network protocol for intelligent video sensing in precision agriculture," in *Ad-hoc Networks and Wireless*, M. Garcia Pineda, J. Lloret, S. Papavassiliou, S. Ruehrup, and C. B. Westphall, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 165–175.

[9] V. Romanov, I. Galelyuka, and Y. Sarakhan, "Wireless sensor networks in agriculture," in *Intelligent Computing and Information Systems (ICICIS), 2015 IEEE Seventh International Conference on*. IEEE, 2015, pp. 77–80.

[10] K. Andersson and M. S. Hossain, "Heterogeneous wireless sensor networks for flood prediction decision support systems," in *Computer Communications Workshops (INFOCOM WKSHPS), 2015 IEEE Conference on*. IEEE, 2015, pp. 133–137.

[11] S. Vanthana and V. S. J. Prakash, "Comparative study of proactive and reactive adhoc routing protocols using ns2," in *2014 World Congress on Computing and Communication Technologies (WCCCT)*. IEEE, 2014, pp. 275–279.

[12] R. Thool *et al.*, "Comparing the performance of proactive and reactive routing protocol in mobile ad-hoc network," in *Industrial Instrumentation and Control (ICIC), 2015 International Conference on*. IEEE, 2015, pp. 394–399.

[13] M. M. Mowla, I. Ahmad, D. Habibi, and Q. V. Phung, "An energy efficient resource management and planning system for 5G networks," in *Proc. IEEE CCNC*, Jan. 2017, pp. 216–224.

[14] P. Kuppusamy, K. Thirunavukkarasu, and B. Kalaavathi, "A study and comparison of olsr, aodv and tora routing protocols in ad hoc networks," in *Electronics Computer Technology (ICECT), 2011 3rd International Conference on*, vol. 5. IEEE, 2011, pp. 143–147.

[15] R. Al-Maharmah, G. Bruck, and P. Jung, "Practical methodology for adding new manet routing protocols to opnet modeler," in *The 5th International Conference on Advances in System Simulation (SIMUL 2013), Vience*, 2013, pp. 73–80.

[16] M. M. Mowla, I. Ahmad, D. Habibi, and V. Phung, "Energy efficient backhauling for 5G small cell networks," *IEEE Transactions on Sustainable Computing*, pp. 1–1, 2018.

[17] H. R. Hussen, S.-C. Choi, J.-H. Park, and J. Kim, "Performance analysis of manet routing protocols for uav communications," in *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE, 2018, pp. 70–72.

[18] A. Al-Baseera, G. Taliba, and A. Bawazira, "Multi-hop wireless network: A comparative study for routing protocols using omnet++ simulator," *Journal of Ubiquitous Systems & Pervasive Networks*, vol. 7, no. 1, pp. 29–34, 2016.

[19] F. F. Khan, T. Samira, A. Jana, and K. N. E. A. Siddiquee, "Performance of agro-sensors: Assessment of optimality in routing protocols of manet in wireless sensor networks," in *Intelligent Control Power and Instrumentation (ICICPI), International Conference on*. IEEE, 2016, pp. 98–102.

[20] K. N. E. A. Siddiquee, F. F. Khan, K. Andersson, and M. S. Hossain, "Optimal dynamic routing protocols for agro-sensor communication in manets," in *Consumer Communications & Networking Conference (CCNC), 2017 14th IEEE Annual*. IEEE, 2017, pp. 1–5.

[21] X. Chang, "Network simulations with opnet," in *Proceedings of the 31st conference on Winter simulation: Simulation—a bridge to the future-Volume 1*. ACM, 1999, pp. 307–314.

[22] H. Bakht *et al.*, "Survey of routing protocols for mobile ad-hoc network," *International Journal of Information and Communication Technology Research*, vol. 1, no. 6, 2011.

[23] M. Rajput, P. Khatri, A. Shastri, and K. Solanki, "Comparison of ad-hoc reactive routing protocols using opnet modeler," in *Computer Information Systems and Industrial Management Applications (CISIM), 2010 International Conference on*. IEEE, 2010, pp. 530–534.

[24] K. Holter, "Comparing aodv and olsr," *folk. uio. no/kenneho/studies/essay/essay. html*, pp. 1–19, 2005.

Return Loss and Gain Improvement for 5G Wireless Communication Based on Single Band Microstrip Square Patch Antenna

Mohammad Faisal[1], Abdul Gafur[2,] Syed Zahidur Rashid[3], Md.OmarFaruk Shawon[4], Kazi Ishtiak Hasan[5] and Md.Bakey Billah[6]
[1, 2, 3, 4, 5,6]International Islamic University Chittagong
Department of Electronics & Telecommunications Engineering
faisal.kml.375@gmail.com[1],agafur_cox@yahoo.com[2],szrashidcce@yahoo.com[3],shawon.omar@gmail.com[4],
kaziishtiaknoman@gmail.com[5],bakeybilla@gmail.com[6]

Abstract— **In this research, a single band micro-strip square patch antenna at 28GHz is proposed. The design of patch antennas are very efficient and widely used in wireless communication due to their lower cost of fabrication, light weight and can operate at microwave frequencies but it offers low efficiency, low gain etc. Future upcoming 5G wireless communication is needed of high gain, good protection from path loss because of their millimetre wavelength of antennas. The main features of this proposed antenna are its very low return loss (S11), high gain, better efficiency and higher bandwidth are of 57 dB, 10.3 dBi , 97% and 4.3GHz respectively which makes it appropriate for 5G wireless communication. The proposed square patch antenna is designed and simulated using CST Microwave studio at operating frequency of 28 GHz. The gain of the antenna is improved by using air substrate. The proposed antenna design gives results of a very low return loss of 57 dB by which the antenna can provide high radiation power in 5G communication.**

Keywords — 5G, mm wave, square patch antenna, 5G antenna, 28GHz, high gain, low S11 parameter

1. INTRODUCTION

With the prosperity of wireless technology, there is an increasing need of microstrip patch antenna with better data transfer speed, efficiency and light weight and lower cost of fabrication [1]. The currently ongoing 4G network technology cannot meet demand of modern wireless technology such as dynamic information access, HD streaming or reducing energy wastage and reducing latency. 5G technology promises mobile data that will be as much as 1 thousand times faster than 4G [2]. Hence 5G network system will provide high data rate, large bandwidth and high capacity. [3]There are several bands aimed for 5G one of them is 28 GHz which is declared by Federal Communication Commission (FCC).Proceeding to these millimetre waves band it will create modern challenges in the design of antennas for the smart phone devices and base stations towers. There are numerous contraventions using millimetre wave frequencies including propagation loss, path loss, signal attenuation and atmospheric distractions. To exceed these challenges, high gain and high directional antennas are needed. The high gain antennas can be applied to overcome the challenges existing at millimetre wave frequencies such as path loss and propagation loss [4]. The problem with using these high frequency bands or millimetre wave bands is that the path loss or free space loss is intense in these bands, which results in the deterioration of signal to interference plus noise ratio (SINR) [5].

The patch antenna with low dielectric constant material is one of the methods to increase the gain [6]. In this design, air is used as the substrate whose dielectric constrains is 1.
In reference paper [1], at operating frequency 60 GHz, they managed to obtain gain 5.48 dB and return loss -41 dB by using H-slot and E-slot. In reference paper [8], work has been done by using air substrate at operating frequency 28GHz and managed to achieve -40dB return loss, 9.8dB gain and 96% efficiency.

In this paper, the proposed patch antenna introduces high realized gain of 10.3dBi and very low reduction loss of 57dB which are the main requirements to overcome the challenges of future 5G wireless Communication. The introduced antenna provides higher realized gain and very low return loss than other referenced works which means our proposed design radiates more efficiently.

2. Design Specifications

The following are the equations needed to calculate width and length of the patch to design the proposed antenna.

$$W = \frac{c}{2f_r}\sqrt{\frac{2}{\epsilon_r+1}} \qquad (1)$$

Where, c= velocity of light ($3 \times 10^8 ms^{-1}$)
fr= resonant frequency
ϵr = dielectric constant (ϵr =1)
The length of the patch can be calculated by using the following formula as shown in equation (2)

$$L = L_{eff} - \Delta L \qquad (2)$$

Where we need to calculate the effective length of the patch to get length of the patch as follows

$$L_{eff} = \frac{c}{2f_r\sqrt{\epsilon_{eff}}} \qquad (3)$$

Where ϵ_{eff} the effective dielectric constants as are stated below

$$\epsilon_{eff} = \frac{\epsilon_r+1}{2} + \frac{\epsilon_r-1}{2}\left(1 + \frac{10h}{w}\right) \qquad (4)$$

Now extension length ΔL as follows:

$$\Delta L = 0.412h\frac{(\epsilon_r+0.3)\left(\frac{w}{h}+0.264\right)}{(\epsilon_r-0.258)\left(\frac{w}{h}+0.8\right)} \qquad (5)$$

The above five equations are used to design the proposed patch antenna and simulated the antenna by using CST Microwave Studio Simulation Software.

Fig 01 illustrates the geometric view of the proposed square patch antenna. As we have seen the substrate is designed by using dimensions of Ls × Wswith air substrate which has dielectric constant and height are 1 and 0.5 respectively.The dimensions used to design the proposed square patch antenna are shown in TABLE 1 and TABLE 2.

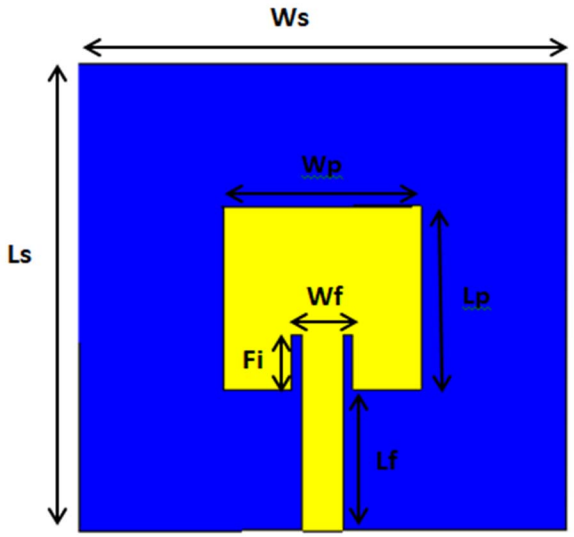
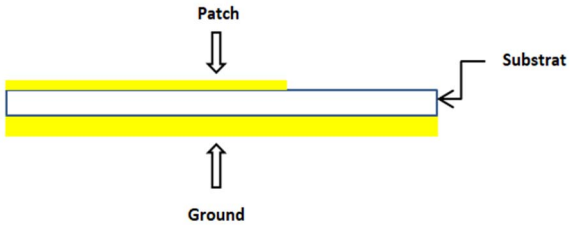
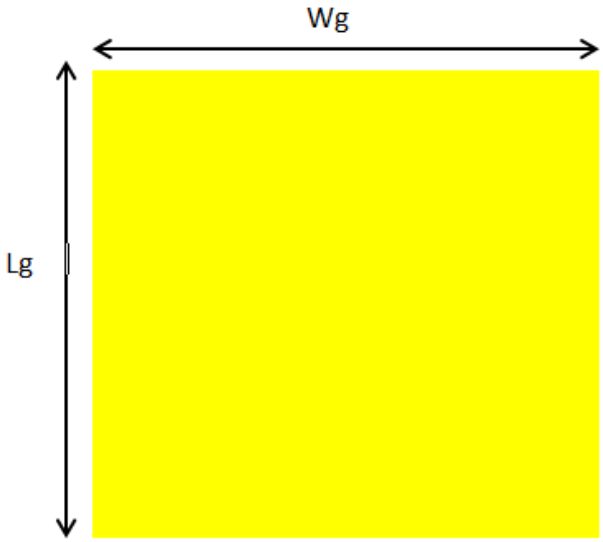


Fig 1: Proposed square patch antenna Geometry.



(a) Side View



(b) Ground View

Fig 02: (a) Side view (b) Ground view

TABLE1

The designed 5G antenna parameters

Symbols of Parameters	Parameters Description	Values in (mm)
L	Patch length	4.8
W	Patch width	4.8
Lg	Length of Ground plane	12
Wg	Width of ground plane	12
Wf	Width of feedline	0.5
Fi	Insertion Feed	1.4

The substrate dimensions used to design the proposed antenna dimension as follows in TABLE 2.

TABLE 2

Parameters Symbols	Description of Parameters	Value (mm)
Ls	Substrate length	6
Ws	Substrate Width	6
hs	Substrate Height	0.5
ϵ_r	Relative permittivity	1

3. SIMULATION RESULTS AND DISCUSSIONS

(A) Return-loss plot:

Return loss generally refers to the reflection coefficient which is denoted by (S11) describes the input output relationship between ports and represents how much power is reflected from the antenna. The return loss plot gives the information of how the feedline is matched to the antenna. If (S11) is 0 dB then it shows that all the power is reflected from the antenna, nothing is radiated. (S11) value should be at least -10 dB to give efficient performance by the antenna. The less the (S11) value the better performance of the antenna. Hence our proposed antenna has (S11) value of -55 dB at 28 GHZ frequency which is very less value and very less loss of power and maximum power is radiated to antenna. Fig 03 shows the plot for Frequency Vs S parameters in (dB).

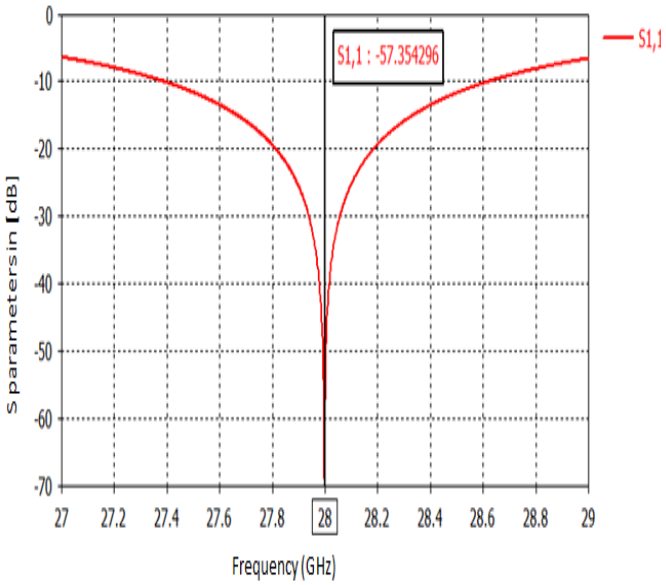


Fig 3: Frequency Vs S parameters in (dB)

(B) Voltage Standing Wave Ration (VSWR) Plot:

The plot of Frequency Vs VSWR is shown in Fig 04 which shows the information that how much antenna is matched to the transmission line and how much power is reflected from the antenna. As shown the proposed antenna has a VSWR value of 1.00 at 28 GHZ which satisfies the requirements range 1≤VSWR≤2 that is very efficient and all power is radiated and no power is reflected.

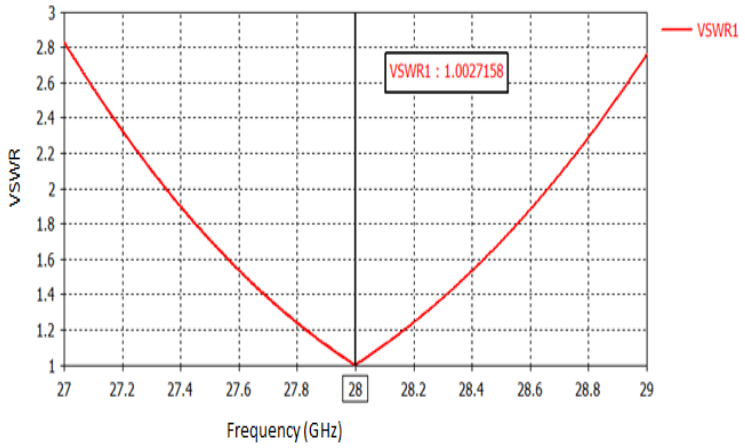


Fig 04: Frequency Vs VSWR

(C) Radiation Pattern:

The antenna radiation pattern represents how much energy is radiated by the antenna. The respective radiation pattern can be seen in 3D and 2D as shown in Fig 05 and Fig 06.

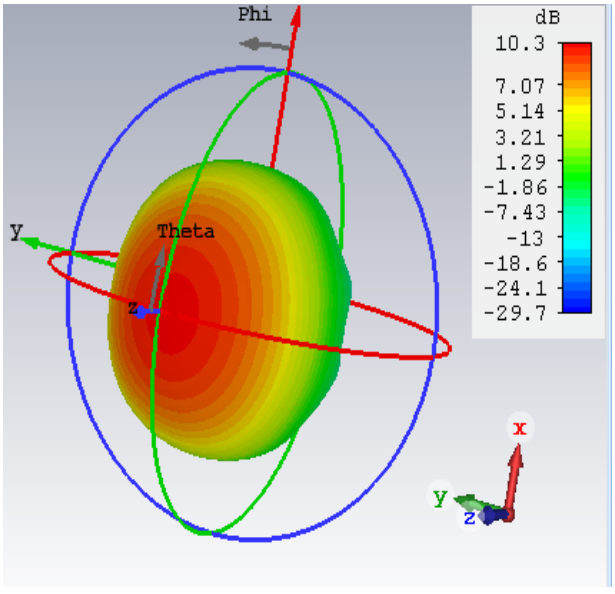


Fig 05(a): 3D radiation pattern

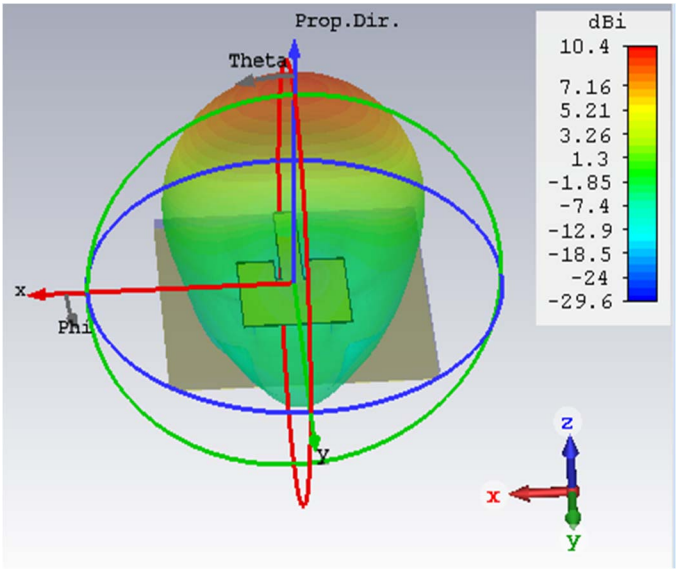


Fig 05(b): Fairfield Transparent 3D radiation pattern

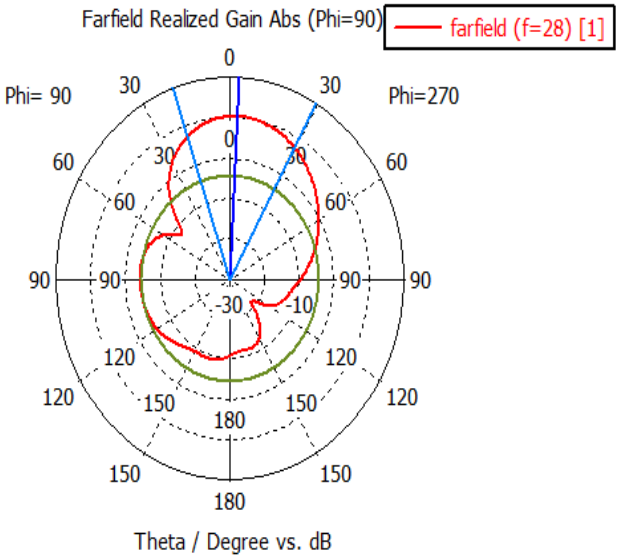


Fig 06: (b) 2D at 28 GHZ when $phi = 90°$

The 2 dimension (2D) radiation plot of the proposed patch antenna can be seen as shown in Fig 6(a) and (b) for both phi=0 degree and phi= 90 degree respectively.

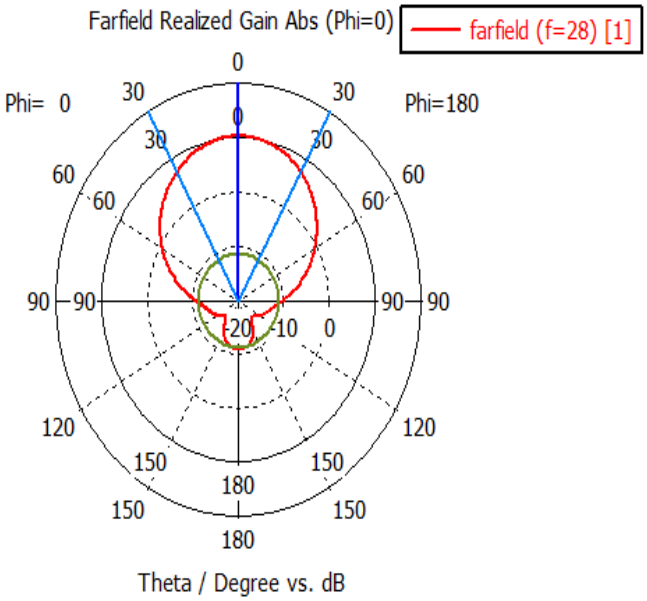


Fig 06: (a) 2Dat 28GHZ when $phi = 0°$

(D) Efficiency plot:

The following figure illustrates the plot of both radiation efficiency and total efficiency Vs Frequency which shows that 97% radiation efficiencies of power for both that exhibit the antenna radiated maximum power and very few signal or power loss at peak 28GHz.

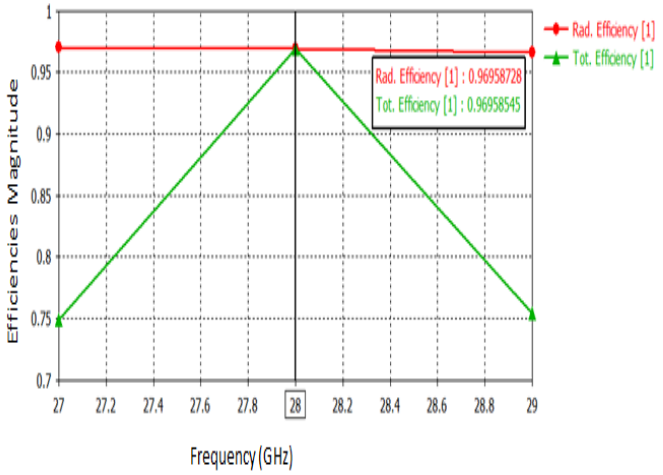


Fig 07: Radiation Efficiency

TABLE 3

Final Characteristics of the proposed antenna

Proposed 5G Antenna	Resonant Frequency	Return Loss	Gain	Bandwidth	Efficiency
	28	-57dB	10.3 dBi	4.3 GHz	97 %

TABLE 4

Comparison of our proposed antenna with existing 5G antennas

Reference Papers	Resonance Frequency (GHz)	Return Loss (dB)	Gain (dB)	Efficiency
[7]	28	-38.1	6.16	91%
[1]	60	-41	5.48	-
[4]	28	-23	7.8	-
[8]	28	-40	9.8	96%
Proposed	28	-57	10.3	97%

4. Conclusion

Our proposed antenna was designed and simulated using CST Microwave Studio. From our simulation result it is found that the antenna has a gain of 10.3 dBi which is higher than the most antenna designed as shown in the comparison TABLE 4. The proposed antenna bears a return loss of -57 dB which means that a very low amount of power is reflected from the antenna and most of the power is radiated to the antenna. Our introduced antenna has a bandwidth of 4.3 GHz and the efficiency is 97%. Hence our simulation results exhibit that it is on per standard of future 5G wireless communication.

REFERENCES

[1] J. Saini and S. K. Agarwal, "Design a single band microstrip patch antenna at 60 GHz millimetre wave for 5G application," *2017 International Conference on Computer, Communications and Electronics (Comptelix)*, Jaipur, 2017, pp. 227-230.

[2] I. Gharbi, R. Barrak, M. Menif and H. Ragad, "Design of patch array antennas for future 5G applications," *2017 18th International Conference on Sciences and*

[3] *Techniques of Automatic Control and Computer Engineering (STA)*, Monastir, 2017, pp. 674-678.

[4] D. Imran *et al*., "Millimetre wave microstrip patch antenna for 5G mobile communication," *2018 International Conference on Engineering and Emerging Technologies (ICEET)*, Lahore, 2018, pp. 1-6.

[5] S. U. Rahman, Q. Cao, I. Hussain, H. Khalil, M. Zeeshan and W. Nazar, "Design of rectangular patch antenna array for 5G wireless communication," *2017 Progress In Electromagnetics Research Symposium - Spring (PIERS)*, St. Petersburg, 2017, pp. 1558-1562.

[6] D. M. Pozar, Microwave Engineering, 3rd ed. Hoboken, NJ, USA: Wiley,2005.

[7] H. Jaafar, M. T. Ali, S. Subahri, A. L. Yusof and M. K. M. Salleh, "Improving gain performance by using air substrate at 5.8GHz," *2012 International Conference on Computer and CommunicationEngineering (ICCCE)*, Kuala Lumpur, 2012, pp. 95-98.

[8] Muhammad Mostafa Amir Faisal, Mohammad Nabil, Md.Kamruzzaman"Design and Simulation of a Single Element High Gain Microstrip Patch Antenna for 5G Wireless Communication" 2018 2nd Int. Conf. on Innovations in Science, Engineering and Technology (ICISET) 27-28 October 2018, Chittagong, Bangladesh.

Performance Evaluation of SDN Controllers in Wireless Network

1st Saiful Islam
Dept. of ECE
HSTU
Dinajpur, Bangladesh
saiful.hstu.sii@gmail.com

2nd Md. Ariful Islam Khan
Dept. of ECE
HSTU
Dinajpur, Bangladesh
arifulk.hstu@gmail.com

3rd Sanjida Tasnim Shorno
Dept. of ECE
HSTU
Dinajpur, Bangladesh
shornohstu29@gmail.com

4th Sumonto Sarker
Dept. of ECE
HSTU
Dinajpur, Bangladesh
sumonto@hstu.ac.bd

5th Md. Abubakar Siddik
Dept. of ECE
HSTU
Dinajpur, Bangladesh
siddik.ece@hstu.ac.bd

***Abstract*—The ongoing relentless innovation of technologies like Big data, cloud computing, virtual reality is creating the need for more secure, further reliable, more scalable ultra fast networks. For meeting these objectives Software Defined Networking (SDN) arises. Distinguished control plane and data plane have made SDN more dependable, fast, secure and increased its widespread acceptance. The crucial part of such a network is the controller which handles the flow control and makes the network more effective. So, for ensuring better performance of the network, selection of the right controller according to the need is imperative. Thus performance evaluation of available SDN controllers is a high-priority. Though there are many researches pursued on the performance analysis of SDN controllers in terms of wired network, there is a lack of studies involving controller efficiency in wireless network. In this paper we have examined the performance of some popular SDN controllers like Ryu, POX, ONOS and Floodlight in an emulated wireless network. Analysing the test results we have asserted our observation regarding the controller performance which can be of considerable importance while determining the right SDN controller for wireless network.**

***Index Terms*—SDN, Mininet-WiFi, controller performance, OpenFlow, wireless, SDWN.**

I. Introduction

Software Defined Network (SDN) is considered the future of networking as it is shaping the era of wired and wireless communication into a new image. The principle factor that differentiates SDN from traditional network is that the separation of control and data plane. Whereas, in traditional networks decision making and data forwarding elements are embedded together, in SDNs they are distinguished between control and data plane. The emergence of centralized controlling capability has introduced controllers which play as a vital part of the network. It can be compared with the brain of human body as it makes the decisions and keeps the network up and running [1]. The better performance, increased security, reliability, flexibility and scalability benefits have made Software Defined Networks more efficient than conventional networks and increased its acceptance to be the standard for internet of the future and Next Generation Networks (NGN) [2]. The future of communication networks mostly involve wireless data transmission. High speed wireless network is enabling technologies like IoT, cloud computing, artificial intelligence, wearable technologies etc to become more effective and efficient. Uses of mobile devices such as smartphones, tablets, laptops rely much on wireless networks. So it is imperative to develop techniques to improve the performance and security of such networks. In this case, SDN is contributing to make wireless communication more versatile in an enormous scale. Thus the concept of Software Defined Wireless Network (SDWN) is introduced [3]. Software Defined Wireless Network paves the way for implementing many future technologies like D2D (Device to Device) communication, NFV (Network Function Virtualization), Edge computing and so on. But performance of a SDN vastly depends on selection of the controller. As operating systems denote performance of computers, controllers play the same role for Software Defined Networks [4]. So performance analysis and comparison of the SDN controllers is a must for ensuring the right choice. Although there have been many researches conducted upon the comparative analysis of SDN controllers, there is hardly any proper paper on the performance analysis of Software Defined Network controllers from the perspective of wireless networks. In this paper, we have selected some of the popular SDN controllers such as- Ryu, POX, ONOS and Floodlight to measure their performance in an emulated SDWN. We have taken Mininet-WiFi as a tool for emulating a wireless network and tested the controllers and considered several parameters those could help us to derive some conclusions. These are: Delay Jitter Throughput.

The rest of the paper includes: section II-Background, which describes previous works related to the topic. This section also overviews important features of the chosen SDN controllers. Section III-Experimental setup, which discusses the tools

used to set up a wireless environment and to measure the performance of the controllers. Section IV-Discussion, which includes the results of the experiment conducted and important implications from the result analysis. Section V-Conclusion and Future work, which concludes the study and provides directions for further research.

II. Background

A. Related Work

As selection of the right controller bears huge importance for ensuring consistent performance of the network, there have been a good number of researches on the performance evaluation of SDN controllers. Bholebawa et al performed comparison in [1] between two leading controllers- Floodlight which is Java-based and POX which is based on Python. The authors have designed several topologies in Mininet such as single, linear, tree and custom. In these topologies RTT and throughput are measured for the two controllers. They concluded that Floodlight shows better performance than POX regarding both of these criteria. Although Floodlight is suggested for achieving efficient network performance, POX is also recommended for its ease of use.

In paper [2], Shamim et al selected three SDN controllers Ryu, POX and Pyretic for performance evaluation. Authors used Mininet as a tool for simulating a wired SDN. The parameters for measuring the performance was RTT or Round Trip Time. It designates the length of time taken to send a signal plus the length of time required for an acknowledgement of that signal to be received. The paper indicates among the three controllers, Pyretic performs better than both Ryu and POX.

In [4], Fancy et al. conducted analysis of performances of two popular SDN controllers- Floodlight and POX. They have considered some basic parameters like delay and throughput. The controllers were tested in different topologies. After observing the results, it is concluded that Floodlight shows far more better performance than POX. However, Floodlight uses mentionable amount of memory space for execution because of its inbuilt java files. POX is better in this case as it is based on python.

Jany et al. [5] evaluated the performance of OpenFlow based wired and wireless network. Authors used the tool Mininet-WiFi to emulate a wireless network. They used Ubuntu Xterm to acquire client-server interfaces and also used iperf test to measure basic parameters like delay, bandwidth, jitter etc. In the Software Defined Wired and Wireless Network (SDWWN), a single controller administered all the switches and access points together. Analyzing the results, authors concluded that the network shows promising results and creates further opportunities to conduct research upon SDWWN using Mininet-WiFi.

Shalimov et al. performed the comparison of SDN controllers- Beacon, MuL, Mestro, POX, NOX, Floodlight and Ryu based on certain criteria like throughput, latency, reliability, security etc. in paper [6]. In this study it appears that Beacon, Floodlight, NOX, POX and Ryu are capable of lasting long-term test when the traffic load is average. MuL and Maestro do not perform the same way. They also performed security analysis by sending malformed packets which can terminate TCP sessions and eventually causing a failure of a network portion or even the entire network.

In [7], comparative performance analysis of some SDN controllers is carried out by Tootoochian et al. NOX-MT, Beacon and Maestro are chosen to conduct the research. NOX-MT appeared to be ahead of others according to the analysis. Although these controllers are no longer popular.

Khattak et al. evaluated performance of two SDN controllers OpenDayLight and Floodlight in [8]. They used Cbench which is a handy tool for benchmarking SDN controllers. In that study they concluded ODL includes memory leakage problem. On the other hand, Floodlight has average higher response time than ODL.

In paper [9], comparison of five popular controllers based on various features is provided by Khondoke et al. Authors have considered several criteria for evaluation such as- GUI, programming language, operating system, virtualization, southbound interfaces, REST API support, OpenFlow version support etc. Although the results vary when the priority of criteria is changed, RYU was voted as the best amongst other controllers by the authors.

In [10], Rastogi et al. carried out the comparative performance analysis of two popular python based SDN controllers- POX and RYU. The authors conducted the study putting traffic handling capabilities into perspective. Mininet was used for emulation and D-ITG for traffic generation. The research concludes that POX provides better results in case of layer 1 switching. Whereas, in case of layer 2 switching RYU appears to be better than POX.

In [11], Beacon, Maestro, IRIS, OpenDaylight, Floodlight, Ryu, POX, NOX, ONOS, MUL, Libfluid-based SDN controllers are assessed by Salman et al. Authors concluded that amongst these open-source controllers Maestro is the topmost considering the latency performance. Mul and Libfluid-based were ahead of others when it comes to throughput.

In [12], both feature oriented comparison and performance analysis of some very popular SDN controllers are done by mamushiane et al. The controllers include OpenDayLight, Ryu, ONOS and Floodlight. After analyzing the results it is stated that Ryu exhibited better performance than others in terms of latency. On the other hand, ONOS proves itself the best when throughput performances are considered. Authors also recommend using OpenDayLight for its enriched features.

Among all these studies it appears that only the last paper [5] includes performance evaluation of a Software Defined Wireless Network. Although in this paper, authors evaluated the performance of the network rather than of any controllers. Thus it appears to us that a research based on the performance evaluation of certain SDN controllers in terms of wireless network comprises considerable significance.

B. SDN Controllers

A Software Defined Network consists of some radical components such as south-bound interfaces, north-bound interfaces, controllers, network applications etc. Among these, south-bound interface works as an interface between controller and infrastructure. It accumulates the status of the network, updates the rules regarding packet forwarding to the switching devices at the infrastructure layer. On the other hand, the north-bound interface functions as an interface between controller and application. It receives policies from SDN applications and communicates them to the controller. These policies are given in high level programming languages [13]. Network applications decides the path for a packet and commands the controller to install necessary forwarding rules in the switching devices. Therefore, SDN controller is the key for implementing the control logic which is gained from the applications and defined by the operator of the network [14]. Controllers can be centralized or decentralized in terms of architecture. They can be built based on several high level programming languages like java, python, C++ etc. Along with delivering the forwarding devices with forwarding rules the controller also provides other services like management of topology, host tracking, security services etc. for the network [15]. There are many SDN controllers developed to this day for various purposes using several languages. We have taken some popular controllers in consideration. They are brought up in Table I with their common features.

TABLE I
COMMON FEATURES OF SDN CONTROLLERS

Name	Ryu	POX	Floodlight	ONOS
Architecture	Centralized multi-threaded	Centralized	Centralized	Distributed
GUI	Yes	Yes	Web/Java-based	Web-based
Programming Language	Python	Python	Java	Java
Platform support	Linux	Linux, Mac, Windows	Linux, Mac, Windows	Linux, Mac, Windows
Network	Campus/ Research	Campus/ Research	Campus	Telecom/Data center
REST API	Yes	No	Yes	Yes

***The features are obtained from [16], [17], [18], [19].**

III. EXPERIMENTAL SETUP

A SDWN separates the control plane from the data plane of the network. It occurs outside the Access Points or APs. There are stations which are managed by Access Points (APs) and the station perform connection with Access Points (APs) through both authentication and association. APs are in charge of transmission and reception of traffic over the wireless network and they receive related instructions from the stations [5].

Mininet-WiFi: It is a tool for emulating Software Defined Wireless Network. It is derived from the Mininet SDN emulator [20] through adding wireless functionality using virtualized WiFi stations and access points. It has several limitations like limited (100) WiFi nodes, not supporting WDS etc. Despite these limitations it is useful for emulating SDWN as it has capabilities for performing all normal SDN emulations like Mininet network emulator [21].

OpenFlow: Mininet-WiFi supports OpenFlow (OF) which is one of the pioneer SDN standards. It explicates the communication protocol which empowers the controller to interact directly with the forwarding plane of the networking devices. It was first introduced by Stanford University in 2008 and version 1.0 was released in 2009. There are several versions released since then. OF is managed by Open Networking Foundation (ONF) which is an organization led by the users for maintaining open standards and SDN adoption [22].

Iperf: We have used iperf for generating traffic which is a tool intended for performance measurement and network troubleshooting. Using iperf, TCP and UDP bandwidth performance can be measured. It produces the traffic on either TCP/IP or UDP and provides measurement according to bandwidth, jitter, delay etc. [23].

In our machine, we have Ubuntu 18.04 LTS installed. We also have used Xterm [24] which is a terminal emulator for Ubuntu and other UNIX-like operating systems. It provides interfaces for both client and server.

The topology we have used is illustrated in figure 1 (Fig. 1). The topology consists of three stations, three access points and a SDN controller which supports OpenFlow. We can acquire this topology by default in Mininet-WiFi using this command-

sudo mn –wifi controller=remote,ip=127.0.0.1 –topo=linear,3,protocols=OpenFlow1.3

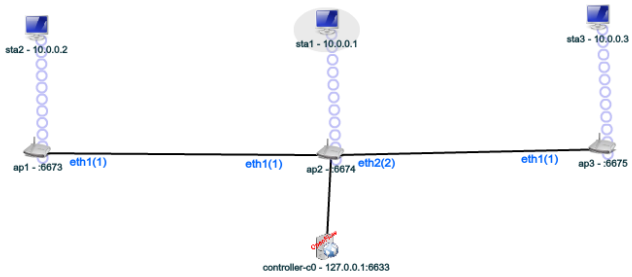


Fig. 1. Samlple topology

Delay: We have examined the packet delay between server and client. Server-client is declared using iperf command. Packets are transmitted and received between the two ends. tcpdump command is used to analyze and save the packet information. The difference between transmission-time and reception-time indicates the delay.

Jitter: The variation in the delay regarding packet flow between server and client is indicated by jitter. For

measuring the jitter we have used iperf command. Essential commands include:

iperf -s -u -p 5566 -i 1

iperf -c 10.0.0.1 -u -b 10M -t 15 -p 5566

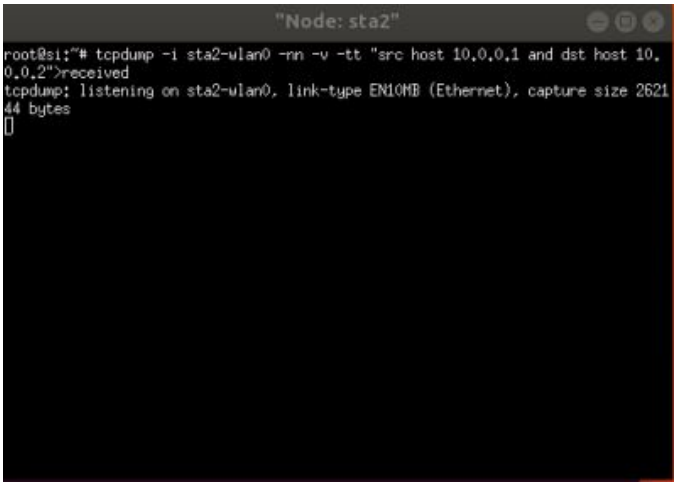


Fig. 2. Receiving terminal for tcpdump

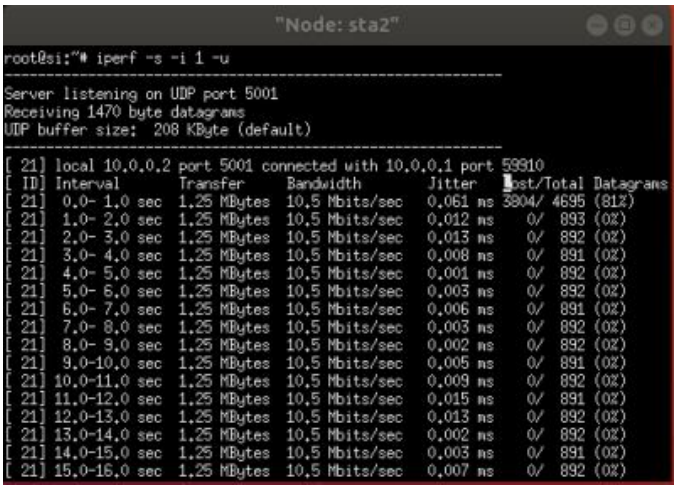


Fig. 3. Server-side terminal for iperf

Throughput: The amount of successful data transferred between two end systems in a given time period is called throughput. We have measured the TCP throughput in Mbits/s. Iperf command is further used here for generating traffic.

IV. Discussion

We have established connection between the controller and Mininet-WiFi. After the establishment, we have waited for a certain duration (20-30 seconds) for the connection to become stable. Then we collected the data associated with delay, jitter and throughput. After collecting the data we have plotted them in several graphs.

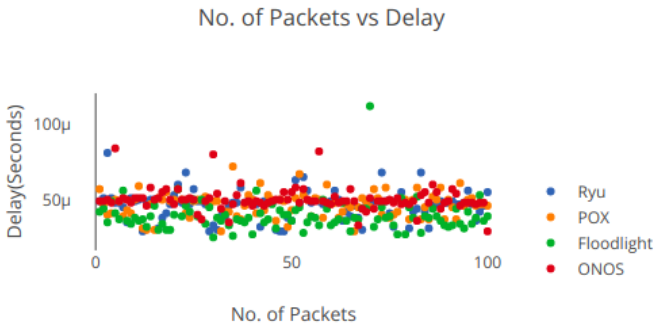


Fig. 4. Delay Comparison

A. Delay Comparison

In the above figure (Fig. 4), the X-axis represents the number of packets and the Y-axis represents delay in seconds. It can be also seen that we have taken packet numbers up to 100 and delay up to 100 micro seconds. We have considered End to End transmission delay in our experiment. It can be observed that among all the SDN controllers-

- Floodlight has the lowest delay. Thus it has outperformed the others in terms of delay.
- On other hand, ONOS has the worst delay performance.
- The rest of the controllers have average performance.

B. Jitter Comparison

The X-axis in the figure below (Fig. 5) represents time in seconds and the Y-axis represents jitter in milliseconds. We have measured jitter for 30 seconds time. From the definition given earlier we know that jitter is the inequality in time delay while transmitting packets of data in a network. There is a certain level of variation (below 30 ms according to Cisco Systems, Inc.) in irregular packet delay that is regarded as acceptable jitter. From the figure we can clearly state that all the controllers show acceptable jitter performance. It appears that-

- Floodlight has the minimum jitter. So it exhibits the best jitter performance.
- On the contrary, Ryu displays the worst jitter performance.
- Among the others, the jitter caused in POX is less than ONOS.

C. Throughput Comparison

The above figure shows the performance of the SDN controllers in terms of throughput. The X-axis entitles time in seconds and the Y-axis entitles throughput in Mbits/s. We have allocated 10 Mbits/s throughput for 30 seconds time. As the throughputs of the SDN controllers do not differ much from one another, it can be stated that they all show the same throughput performance.

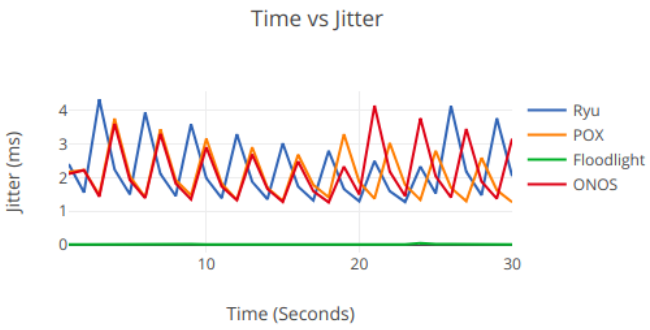


Fig. 5. Jitter Comparison

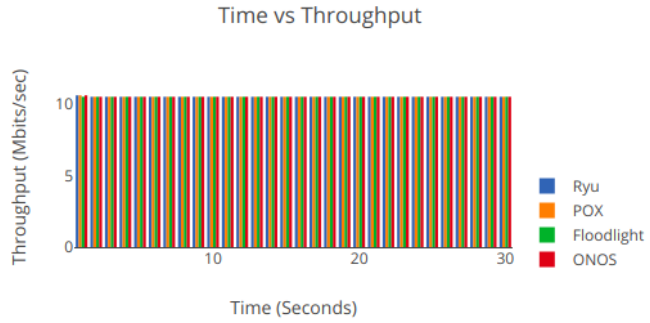


Fig. 6. Throughput Comparison

V. Conclusion and Future Work

From the experimental evaluation we can cease to a decision that Floodlight performs the best among the four SDN controllers in a SDWN in terms of delay and jitter. We can also conclude that in terms of delay, ONOS shows the worst performance. On the contrary, Ryu displays the worst performance when we consider jitter. However, POX exhibits an average performance in all terms. Although we have taken these three criteria in consideration, there are various other properties based on which controller-performance can be evaluated. Such as workload, datagram loss, security etc. We have excluded these parameters. We would like to measure controller performance in terms of these criteria in future. Moreover, we intend to deploy a SDWN in our campus using Floodlight.

References

[1] Bholebawa, I. Z., Dalal, U. D. (2018). Performance Analysis of SDN/OpenFlow Controllers: POX Versus Floodlight. Wireless Personal Communications, 98(2), 1679-1699.

[2] Shamim, S. M., Shisir, S., Hasan, A., Hasan, M., Hossain, A. (2018). Performance Analysis of Different Open Flow Based Controller Over Software Defined Networking. Global Journal of Computer Science and Technology.

[3] Costanzo, S., Galluccio, L., Morabito, G., Palazzo, S. (2012, October). Software defined wireless networks: Unbridling SDNs. In Software Defined Networking (EWSDN), 2012 European Workshop on (pp. 1-6). IEEE.

[4] Fancy, C., Pushpalatha, M. (2017, December). Performance evaluation of SDN controllers POX and floodlight in mininet emulation environment. In 2017 International Conference on Intelligent Sustainable Systems (ICISS) (pp. 695-699). IEEE.

[5] Jany, M. H. R., Islam, N., Khondoker, R., Habib, M. A. (2017, December). Performance analysis of OpenFlow based software defined wired and wireless network. In Computer and Information Technology (ICCIT), 2017 20th International Conference of (pp. 1-6). IEEE.

[6] Shalimov, A., Zuikov, D., Zimarina, D., Pashkov, V., Smeliansky, R. (2013, October). Advanced study of SDN/OpenFlow controllers. In Proceedings of the 9th central eastern european software engineering conference in russia(p. 1). ACM.

[7] Tootoonchian, A., Gorbunov, S., Ganjali, Y., Casado, M., Sherwood, R. (2012). On Controller Performance in Software-Defined Networks. Hot-ICE, 12, 1-6.

[8] Khattak, Z. K., Awais, M., Iqbal, A. (2014, December). Performance evaluation of OpenDaylight SDN controller. In Parallel and Distributed Systems (ICPADS), 2014 20th IEEE International Conference on (pp. 671-676). IEEE.

[9] Khondoker, R., Zaalouk, A., Marx, R., Bayarou, K. (2014, January).Feature-based comparison and selection of Software Defined Networking (SDN) controllers. In Computer Applications and Information Systems (WCCAIS), 2014 World Congress on (pp. 1-7). IEEE.

[10] Rastogi, A., Bais, A. (2016, December). Comparative analysis of software defined networking (SDN) controllersIn terms of traffic handling capabilities. In Multi-Topic Conference (INMIC), 2016 19th International (pp. 1-6). IEEE.

[11] Salman, O., Elhajj, I. H., Kayssi, A., Chehab, A. (2016, April). SDN controllers: a comparative study. In Electrotechnical Conference (MELECON), 2016 18th Mediterranean (pp. 1-6). IEEE.

[12] Mamushiane, L., Lysko, A., Dlamini, S. (2018, April). A comparative evaluation of the performance of popular SDN controllers. In Wireless Days (WD), 2018 (pp. 54-59). IEEE.

[13] Xia, W., Wen, Y., Foh, C. H., Niyato, D., Xie, H. (2015). A survey on software-defined networking. IEEE Communications Surveys Tutorials, 17(1), 27-51.

[14] Kreutz, D., Ramos, F. M., Verissimo, P. E., Rothenberg, C. E., Azodolmolky, S., Uhlig, S. (2015). Software-defined networking: A comprehensive survey. Proceedings of the IEEE, 103(1), 14-76.

[15] Raj, P. (2018, February). Topology-based analysis of performance evaluation of centralized vs. distributed SDN controller. In Current Trends in Advanced Computing (ICCTAC), 2018 IEEE International Conference on (pp. 1-8). IEEE.

[16] RYU SDN Framework Ryubook 1.0 documentation.Retrieved from https://osrg.github.io/ryu-book/en/html.

[17] Installing POX POX Manual Current documentation. Retrieved from https://noxrepo.github.io/pox-doc/html.

[18] Project Floodlight.. Retrieved from https://floodlight.atlassian.net/wiki/spaces/ floodlightcontroller/pages/ 1343542/Getting+Started.

[19] Guides - ONOS - Wiki. Retrieved from https://wiki.onosproject.org/display/ONOS/Guides.

[20] Mininet: An Instant Virtual Network on your Laptop (or other PC).Retrieved from http://mininet.org/.

[21] Fontes, R. R., Afzal, S., Brito, S. H., Santos, M. A., Rothenberg, C. E. (2015, November). Mininet-WiFi: Emulating software-defined wireless networks. In Network and Service Management (CNSM), 2015 11th International Conference on(pp. 384-389). IEEE.

[22] What is OpenFlow? Definition and How it Relates to SDN. Retrieved from https://www. sdxcentral.com/ sdn/definitions/what-is-openflow/.

[23] GUEANT, V. iPerf - The TCP, UDP and SCTP network bandwidth measurement tool. Retrieved from https://iperf.fr/.

[24] XTERM Terminal emulator for the X Window System. Retrieved from https://invisible-island.net/xterm/.

Low-Cost Underwater Wireless Communication System Development Using Piezo-Ceramic Transducer

Md. Ragib Shaharear,[1] H.A.Shakhawat Hossen Prayash,[2] Md. Farhanul Islam,[3] and Tuhin Dey[4]
[1,2,3,4]Department of Electrical & Electronic Engineering,
[1,2,3,4]Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh.
[1]shaharearshuvo@gmail.com, [2]habib.prayash@gmail.com, [3]md.farhanulislam@gmail.com [4]deytuhin-eee@sust.edu

***Abstract*—In recent years, the need for reliable, short-ranged underwater wireless communication has been proliferated because of the surge in underwater exploration. Underwater Wireless Sensor Networks (UWSN) can be used in vehicles and sensors to perform collaborative monitoring and data collection task. Due to the high density and contamination, limited bandwidth, high propagation delay, media access control, and power constraint underwater communication is a very challenging task. Researchers developed different research methodologies to overcome this problem. In this paper, an acoustic communication-based system is proposed by using Piezo-Ceramic transducers and practically developed. The acoustic wave can travel easily in water without much attenuation. Piezo-Ceramic transducer creates vibration according to the electrical pulses and does not need much modification of real data signal. This paper describes a wireless low-cost acoustic system for underwater communication. The system is built with some filters and amplifiers and power booster circuits. Finally, a cost effective underwater communication system is proposed that has been successfully implemented on a submarine bot.**

***Keywords*—Acoustic system, Underwater communication, Power amplifier, Signal Processing, Receiver, Transmitter, Ultrasonic transducer.**

I. INTRODUCTION

A commonly used electromagnetic signal, radio signal's electromagnetic wave cannot pass through the water smoothly due to high attenuation and it's strength decrease and which is noticeable in the water. Optical signal attenuate for reflection, dispersion, etc. But sound wave can propagate through the water very well, so an acoustic system can be used to overcome this problem. The underwater communication system can be used to communicate with a submarine, in marine research, oceanography etc[1]. In recent years, rapid progress has been made in clean deep water communication and shallow water communication but it is considered very challenging. High-speed communication in the underwater acoustic channel is very difficult because of some factors such as multi-path propagation, small available bandwidth and high signal attenuation co-factor, refractive properties of the medium, severe fading, large Doppler shifts [2]. The available bandwidth in an underwater acoustic communication channel is severely limited for transmission loss which increases with both frequency and range[3]. An ultrasonic transducer is used here to send and receive data for reducing attenuation co-factor. In recent years many research works are done for underwater communication. For example, the Persistent Littoral Undersea Surveillance Network (PLUSNet) describes multi-sensor and multi-vehicle antisubmarine warfare (ASW) technique for underwater communications[4]. A short-range shallow water network is proposed for observing pollution in Newport Bay, CA[5]. A network of acoustic modems is proposed for low range, low power acoustic communications for seismic monitoring system development[6]. In this paper, in order to make underwater communication smooth a low power consumption digital acoustic system and cost effective underwater communication system is proposed and developed. Our proposed system is a new acoustic technology system in the world of underwater communication which is different from others and helpful for smooth underwater communication. Our underwater system consists of a transmitter, receiver, voltage booster, power amplifier and signal processing unit. In section II total system overview is discussed, Section III transmitter circuit full system is discussed, Section IV receiver circuit is discussed, Section V data analysis, difficulties and its solution and result analysis are discussed, Section VI conclusion section, here our future work and overall review of paper are discussed. Our main focus is to make this system cost effective, according to data Linkquest, Benthos, and DSPComm underwater communication system cost is more than $8000 [7]. But we make some change in our ultrasonic transducer to use it properly and also design our transmitter, receiver, voltage booster and power amplifier circuit so that cost will be minimized and the total cost to make our underwater wireless communication system is $625. Some problems are found when we started working with the ultrasonic transducer. Authors will describe the problems and their solutions in different sections.

II. SYSTEM OVERVIEW

The present work demonstrates a stable low-cost underwater wireless communication system. It has been successfully tested on eight feet long water pool. The communication system is divided into two parts, namely- transmitter and receiver. The transmitter consists of three stage power amplifier. On the other hand, the receiver consists of three stage noise filter and three stages signal amplifier. A custom made transducer is used in both receiver and transmitter of the system. Piezo Ceramic Transducer named SMC20D17H5 is used which costs about $35.



Fig. 1: Piezo-ceramic Transducer without coating (left) and with coating (right)

This ultrasonic transducer [8] is used for transmitting and receiving data. If an electrical (voltage) signal is applied to it, it generates radial mode vibration and if it senses vibrations, it generates an analog electrical (Voltage) signal. It has very low internal damping. When an electrical pulse is applied, it takes 4-5 cycles to propagate the wave. Its typical output impedance is 50 Ohms and the center frequency is around 47KHz. It can create radial mode vibrations means sound waves. But the materials used in the transducer (silver electrode) is very sensitive to the water at the presence of voltage. When 40v DC voltage is applied to the silver electrode of the ultrasonic transducer, the electrode starts to react with the water and within a few minutes, it has been damaged completely. To solve this chemical reaction problem, a resin coating is applied to encapsulate the silver electrode from the water and the problem is solved.

It is tested on a 100uV input signal & 15V (p-p) output signal is obtained.

III. Transmitter

In this section, transmitter converts electrical signals into ultrasounds and this ultrasonic sounds can propagate through the water easily. Transmitter unit can be divided into four stages:

1) Signal processing
2) Voltage booster
3) Power amplifier
4) Transducer

A. Signal processing

First of all the data is converted into a character string and send to the main processing unit of the transmitter, which is actually a micro-controller. The character string is converted into an 8-bit binary number. The 8-bit binary is divided into two 4-bit binary segments.
A start bit is added to those 4-bin binary segment separately to make the starting clear to identify and error detection.

B. Voltage Booster

For long range transmission, ultrasonic transducer requires higher voltage. As the primary power is supplied from a 12v battery, the voltage has been boosted to 40v to increase the range. The is LM2577 is used for the boosting purpose [9].

C. Power Amplifier

For power amplification, a boosted voltage from the boost converter is used. As a transmitter transducer requires enough power to deliver the signal through the water so Darlington pair circuit is used to increase the current. [10].

IV. Receiver

The receiver converts the ultrasonic sound signals into electrical signals. The receiver is a very high gain multistage signal amplifier. The receiver unit can be divided into some stages :

1) Transducer
2) Amplifier
3) Filter
4) Clipper
5) Microcontroller

It is a three stage amplifier. It's Gain is more than 100000. It has been tested on 100uV to 15 Volts. It has some high pass and low pass filters.

A. Amplifier

Three stage high gain amplifiers are used. In the first stage, the signal is amplified by a transistor PN2222. To determine the gain following equations are used[11].

$$Gain = \frac{V_{out}}{V_{in}} = \frac{\triangle V_E}{\triangle V_B} = -\frac{R_E}{R_C} \tag{1}$$

Then the signal is passed through a bandpass filter. In the second stagethe pre-amplifier's signal is amplified again by an op-amp LM318. Finally, this signal reached to the third stage through another bandpass filter and amplified to an output of peak to peak 12-15 volt.

V. Data Transfer Process

Frequency Shift Keying (FSK) modulation is used for the underwater communication system.

- 4 bit communication system is used (can be extended to 8 bit).
- Start bit: 47kHz frequency, 3 clock cycles time.
- Binary 0: 47kHz frequency, 1 clock cycles time.
- Binary 1: 47kHz frequency, 2 clock cycles time.

In the developed system every clock cycle uses 40 kHz frequency for 50 ms. Here 50 ms is the optimal time for that carrier frequency as lowering time will cause more distortion. Start-bit contains 3 clock cycle, Binary 0 and 1 take 1 and 2 clock cycles and there is a gap one cycle between every bit transmission. So the total time is,

*Maximum time = 50*start bit + 50*4*binary 1 + 50*4*Gap*
*or, Maximum time = (50*3 + 50*4*2 + 50*4*1) ms*
or, Maximum time = 750 ms
*So, time for 1-byte (8-bit) = 750*2 ms = 1500 ms.*
So, The minimum data rate = 1/1.5 byte/second = 0.67 byte/second

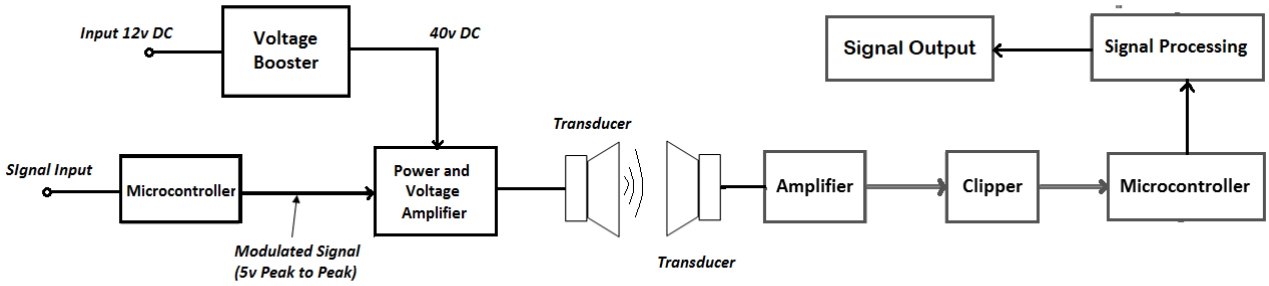


Fig. 2: Block diagram of one way signal transmission & receiving system.

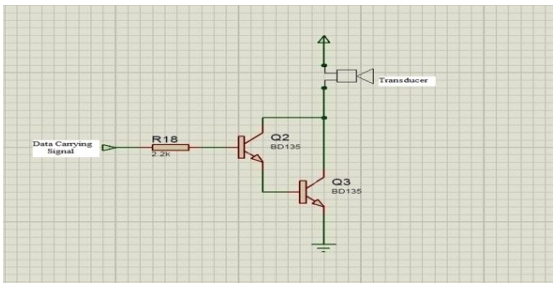


Fig. 3: Power Amplifier Circuit Diagram

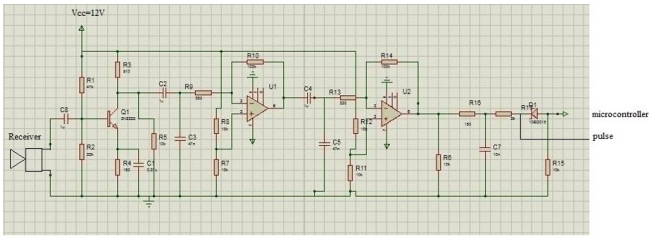


Fig. 4: Signal Receiver Circuit Diagram

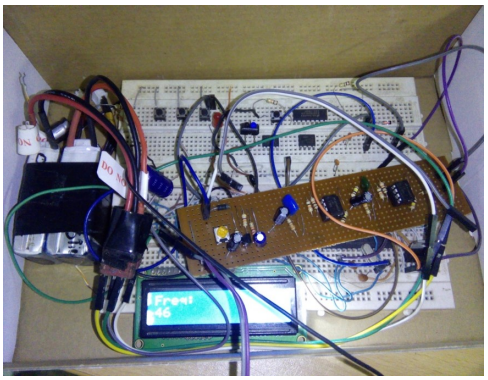


Fig. 5: Receiver Circuit Setup

A. *Command flow diagram of data transmitter*

In figure command flow diagram of data transmitter is shown.

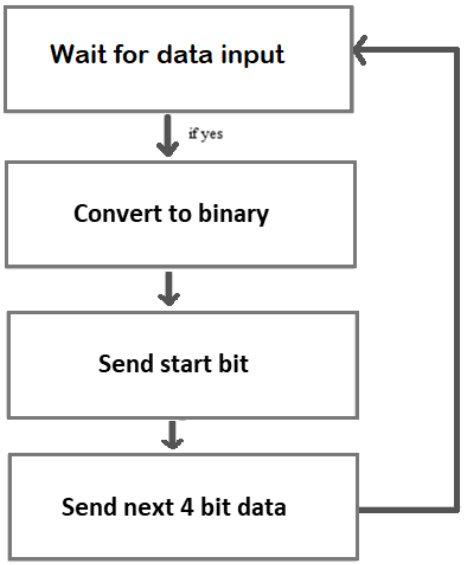


Fig. 6: Flow diagram of data transmitter (used in submarine bot)

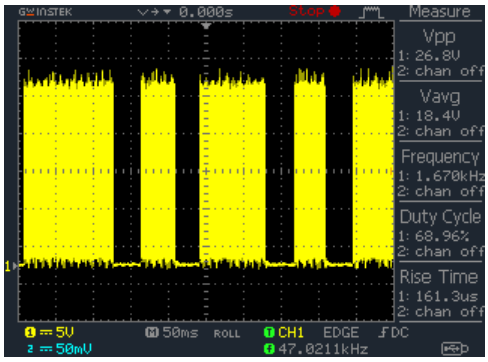


Fig. 7: Transmitted Signal

B. Command flow diagram of data receiver

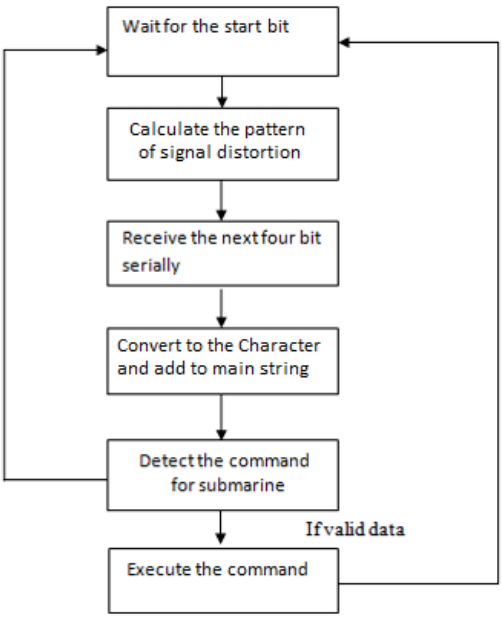


Fig. 8: Flow diagram of data receiver (used in submarine bot)

In figure command flow diagram of data receiver is shown.

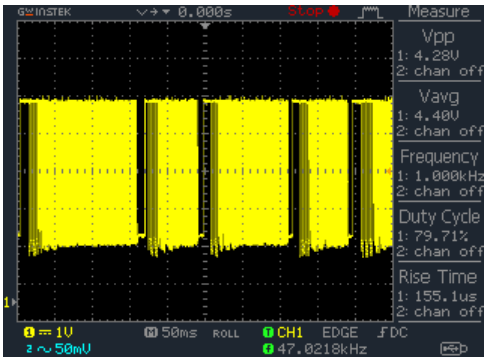


Fig. 9: Received Signal

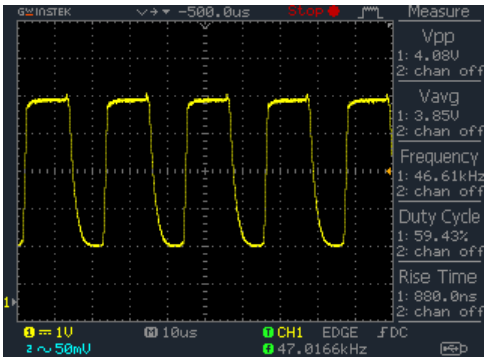


Fig. 10: Received Signal

C. Error Correction

The start bit is predefined and is known to the transmitter and receiver unit. If any distortion occurs, it affects both the start-bit and data bit in the same way. Calculating the distortion in the start-bit, an algorithm is developed to attempt a recovery in all individual bits. However, the recovery can work on only for a certain amount of distortion.

VI. Conclusion

In this work, main focus was to establish a smooth communication for the underwater acoustic channel. Frequency Shift Keying (FSK) modulation technique is used here for data transmission. Four bit communication system is developed and studied here. The data transfer result is very smooth which has been seen shown in the oscilloscope graph. The future plan of this project is to increase the range of communication distance by implementing powerful voltage booster and making the error correction algorithm more efficient and also working with higher frequency to get increased data transfer rate.

Acknowledgement

This work is funded by University Research Centre and supported by the Department of Electrical and Electronic Engineering of Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh. Authors are grateful to the University Research Centre and Department of Electrical & Electronic Engineering of Shahjalal University of Science and Technology for funding and lab supports.

References

[1] Chitre, M., Shahabudeen, S. and Stojanovic, M., 2008. Underwater acoustic communications and networking: Recent advances and future challenges. Marine technology society journal, 42(1), pp.103-116.

[2] Ling, J., Yardibi, T., Su, X., He, H. and Li, J., 2009. Enhanced channel estimation and symbol detection for high speed multi-input multi-output underwater acoustic communications. The Journal of the Acoustical Society of America, 125(5), pp.3067-3078.

[3] Stojanovic, M., 1996. Recent advances in high-speed underwater acoustic communications. IEEE Journal of Oceanic engineering, 21(2), pp.125-136.

[4] Grund, M., Freitag, L., Preisig, J. and Ball, K., 2006, September. The PLUSNet underwater communications system: acoustic telemetry for undersea surveillance. In OCEANS 2006 (pp. 1-5). IEEE.

[5] Jurdak, R., Lopes, C.V. and Baldi, P., 2004. Battery lifetime estimation and optimization for underwater sensor networks. IEEE Sensor Network Operations, 2006, pp.397-420.

[6] Wills, J., Ye, W. and Heidemann, J., 2006, September. Low-power acoustic modem for dense underwater sensor networks. In Proceedings of the 1st ACM international workshop on Underwater networks (pp. 79-85). ACM.

[7] LinkQuest, Inc. Underwater Acoustic Modems[Online].

[8] Lee, S-H., and T. J. Royston. "Modeling piezoceramic transducer hysteresis in the structural vibration control problem." The Journal of the Acoustical Society of America 108.6 (2000): 2843-2855.

[9] Cao, Xinping, et al. "Electromagnetic energy harvesting circuit with feedforward and feedback DCDC PWM boost converter for vibration power generator system." IEEE Transactions on Power Electronics 22.2 (2007): 679-685.

[10] Sidney, Darlington. "Semiconductor signal translating device." U.S. Patent No. 2,663,806. 22 Dec. 1953.

[11] Jeppson, K.O., 1994. Modeling the influence of the transistor gain ratio and the input-to-output coupling capacitance on the CMOS inverter delay. IEEE Journal of Solid-State Circuits, 29(6), pp.646-654.

Mobility Aware Cache Management in 5G Future Generation Wireless Communication System

Mayeesha Fairooz Ahmad
Electrical and Electronic Engineering
East Delta University
Chittagong, Bangladesh
143001010@eastdelta.edu.bd

Mohammad Arif Hossain
Computer Engineering
University of Central Florida
Orlando, FL, USA
arif.hossain@knights.ucf.edu

***Abstract*— Cache management has become one of the key issues in future generation 5G wireless communication network. The feature of high-speed mobility along with vast users of the 5G network increases complexity in cache management. Depending on the cache location and user's mobility, the handoff has become even more prevalent in the content-centric 5G network architecture. As a result, an increased number of dropping and blocking might occur hindering the desired quality of service (QoS). Keeping that in mind, in this paper we have proposed a mobility-aware handoff priority-based request admission control (RAC) policy for managing cached content requests by utilizing limited bandwidth. By reducing the dropping probability and ensuring lesser force termination of already admitted requests; our system provides a better QoS. This paper investigates the scope of the betterment of QoS by reducing the dropping probability in a mobility-aware 5G heterogeneous network (HetNet). The analytical results show better performances than the general scheme in handling the effects of user mobility.**

Keywords—5G, cache, mobility, handoff, dropping probability, blocking probability, bandwidth utilization, request admission control.

I. INTRODUCTION

In the near future, user demands such as- increased capacity, improved data rate, decreased latency, and better quality of service are needed to be addressed with future generation communication technologies. The 5G future generation systems are expected to provide enhanced device- and network level capabilities. It is anticipated that the upcoming 5G network should fulfil requirements such as- 1000 times greater mobile data traffic volume compared to current traffic, 100 times more connections, 10Gbps peak data rate, above 10Mbps affordable per-user data rate anywhere and anytime, and 100Mbps for some particular user, lower latency, delay in milliseconds (ms) for the end-to-end user plane, and within 10ms for the control plane, higher spectrum efficiency, 3-5 times higher than that today for large-area coverage and even higher for special cases, lower cost, nearly 1/1000 of cost per bit than today, higher energy efficiency, almost 1/1000 of energy consumption per bit than today [3]-[5], [7].

While 100 times more user connectivity is an expectation from 5G technologies, this network shall also incorporate high mobility requirements as an integral part. For instance, the high-speed deployment which focuses on continuous coverage along the track in high-speed trains is reported in [6], and the higher user mobility as one of the use case families is proposed by NGMN (Next Generation Mobile Network) [12]. For 5G, the targeted maximum user speed at which a defined QoS can be achieved is 500km/h [6]. Whereas a lot of things can be developed in the field of 5G technology, in this paper the focus has been centered on significantly reducing the dropping probability of handoff requests due to high speed user mobility and improving bandwidth utilization.

The sections of this paper are organized as follows-

Chapter II discusses the previous works related to this one. Chapter III has information about the methodology that is developed for this work while chapter IV illustrates the experimental evaluation of the stated work before rounding it off. We conclude the paper, indicating the scopes of future works, in chapter V.

II. RELATED WORKS

Over the years, a lot of research works has been done to improve the quality of service (QoS) in wireless cellular communication technology. In reference [8], the developed CAC scheme enables to minimize the call dropping probability, as well as it improves bandwidth utilization. Although the proposed scheme blocks more new calls, it reduces the number of handoffs and average call duration and forced call termination probability as compared to the non-prioritized bandwidth-adaptive scheme. Just like in the voice-centric (aka call-centric) systems, Admission Control scheme can be equally significant in the emerging content-centric heterogeneous 5G wireless communication system where the provisioning of quality of service (QoS) is a major concern. In recent years, with the development and increase of mobile devices, we see an exceptional global growth of mobile data traffic (reached 7.2 exabytes per month in 2016) [9], [10]. Due to enormous users and their numerous demands, it has become a mandate to make the best use of the limited resources that assure the highest standard of Quality of Service (QoS). To meet the guaranteed level of QoS for the maximum number of users, maximum utilization of bandwidth and some other factors like- call dropping probability, call blocking probability etc. are needed to keep under consideration. The lower performances of these factors may retrograde the overall cellular networks performances [11]. In such cases, releasing some bandwidth from the already admitted non-real-time traffic calls and the released bandwidth is used to handle handoff calls, and then the resulting probability of dropping a handoff call is lesser than the probability of blocking a new call [8]. Reference [16] also suggests a channel borrowing scheme for efficient resource management. Studies have shown that an effective admission control scheme can evidently reduce blocking or dropping probability along with higher bandwidth utilization by ensuring the best use of limited resources [15], [8]. As the user mobility is rapidly increasing and predicted to reach about 500kmph in 5G [1], [2], [6]; managing handoff requests is a challenge in the content-centric 5G network architecture. Under high mobility scenarios, there shall be quite frequent handoffs; new algorithms are needed to reduce the number of handoffs, the handoff latency, and the handoff failure probability [7]. While the fifth generation (5G) mobile networks, in fact, can enhance mobile network

performances, yet reducing the latency of network services remain a major challenge. Because of the escalating requirements of the mobile applications and services on prompt delivery of contents or real-time user interaction delays occur [13]. To overcome latency issues, different caching techniques have been proposed and those have proven caching as to be an effective way for improving users' quality of experience (QoE) [14]. Hence, the efficient management of contents cached in a 5G network is worth the efforts being put in research in this regard.

In our work, to meet the guaranteed QoS in the content-centric next-generation 5G wireless communication technology, we focus on effective cache management in high-mobility scenarios and utilizing the limited resources of bandwidth with request admission control.

III. METHODOLOGY

A. Network Architecture

The design of the proposed scheme is a two-tier heterogeneous network (HetNet) consisting of macrocell base stations (MBSs) and ultra-densely deployed several small cell base stations (SBSs) under each MBS. A model of the proposed network architecture has been displayed in Figure 1. In 5G HetNets, MBSs are responsible for network coverage with control signaling, whereas SBSs are expected to be densely deployed to boost network capacity in a "plug-and-play" manner. The topologies of both the MBSs and SBSs are modeled as regular hexagonal cells. MBSs and SBSs use orthogonal spectrum bands to avoid inter-tier interference [17]. Both MBSs and SBSs are connected with the core network through wired backhauls. Each user is connected to either an SBS or an MBS depending on the location of the cached content for which the user has requested.

B. Request Admission Control (RAC) Scheme

The flow diagram in Figure 2 represents the proposed cache management based on a handoff prioritized Request

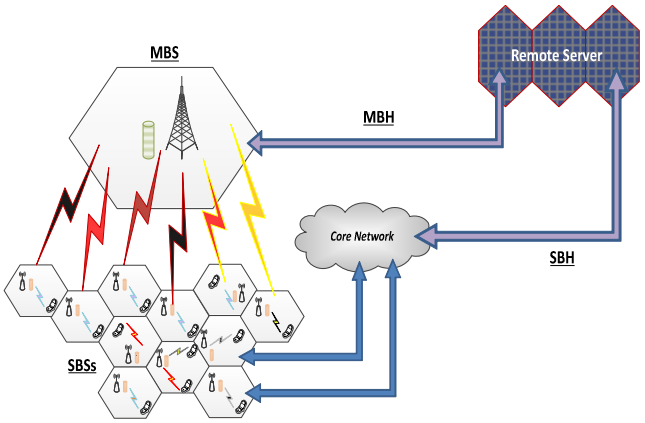


Figure 1: System architecture for the proposed scheme

Admission Control (RAC) policy. Here, as we can see when a request (content request) arrives at the system node, if the content is already cached in the system then depending on the cache location the scheme is followed.

Generally, a channel receives two types of requests; new request and request being handed over from another channel. In our study, we consider handoffs which are initiated by crossing a cell boundary due to the user's mobility. So, for a handoff request, the client is served with priority. Through the available Radio Access Network (RAN) bandwidth of the respective SBS or MBS reserved for handoff requests serve this type of clients coming into the system owing to mobility issues. In case of unavailability of the reserved bandwidth (BW), handoff requests are served by the remaining non-reserved BW of the requested BS. When even that is not available, then the request is to be served either by the non-reserved RAN BW of respective MBS of the requested SBS. In that way, some unused bandwidth is borrowed from the MBS in order to serve the frequent handoff users of popular contents at SBSs. When handoff request enters a channel at BS where no remaining RAN BW is available, then backhaul (BH) connection is established to accept that request.

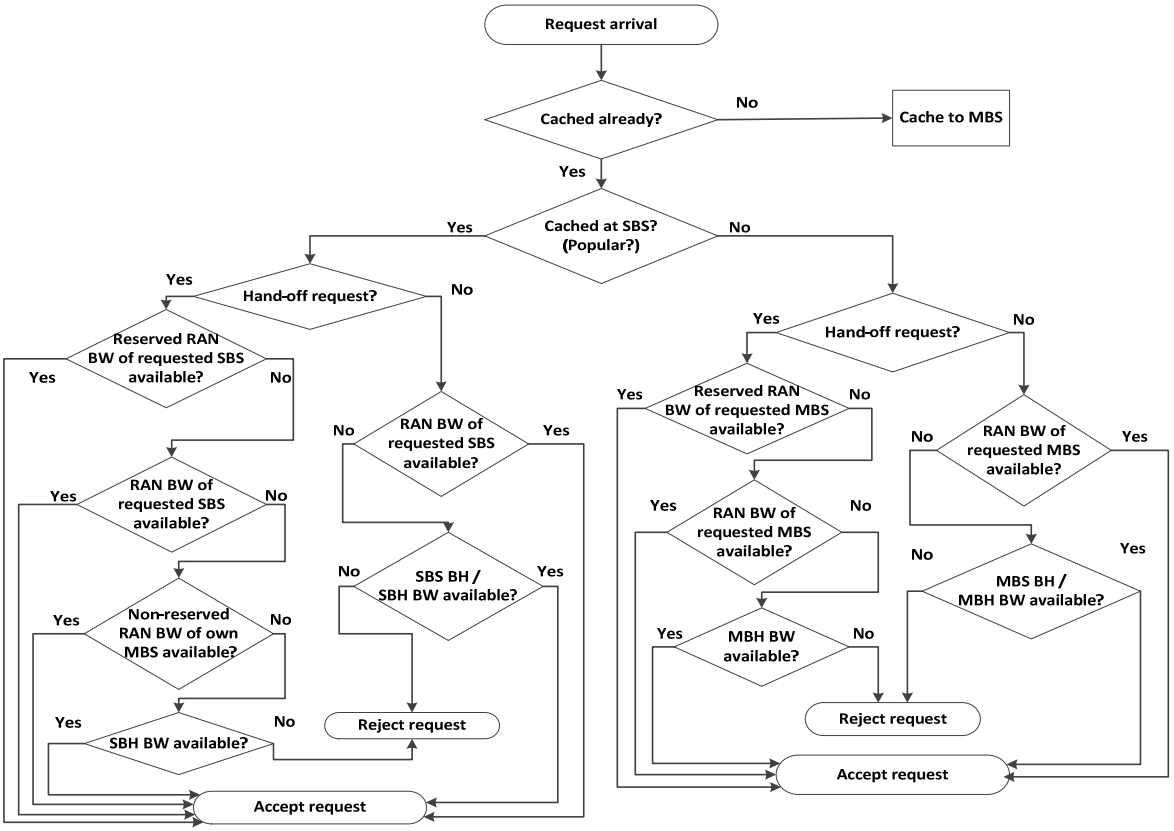


Figure 2: Flow Diagram of the proposed Request Admission Control (RAC) scheme

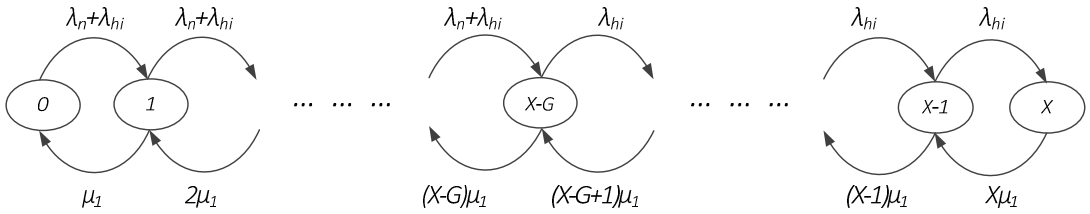
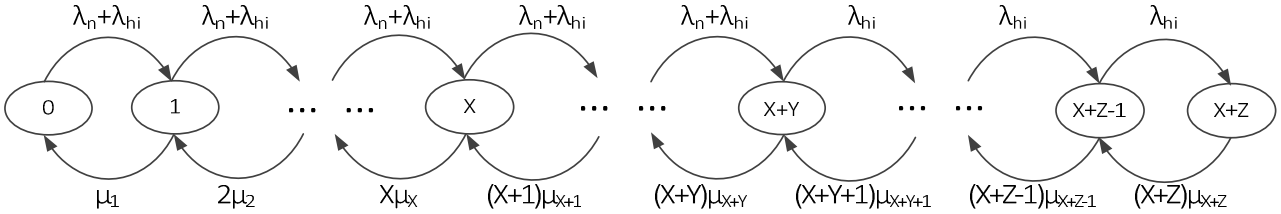


Figure 3: The Markov chain for the general scheme



$0 \leq Y < Z$

Figure 4: The Markov chain for the proposed mobility aware HO-prioritized RAC scheme

For new requests (not handoff), if the requested content is cached then it is accepted if 1) the non-reserved RAN BW of the requested SBS/MBS is available or 2) through BH network to the core network. Otherwise, the requests are rejected.

The RAC illustrated in Figure 2 explains how a handoff (HO) request is given priority over a new request with the best utilization of the limited bandwidth resources. As a result, our proposed approach can achieve a significant reduction in dropping probability which is a critical factor in maintaining QoS. Also, sharing of non-reserved bandwidth of a parent MBS with its own SBSs can provide with improved BW utilization to the system. Nevertheless, as a consequent of prioritizing HO requests, more new requests are to be rejected in this method. Rejection of some new requests will cause a minor increase in the probability of blocking. But, as our approach intends to majorly address user mobility, rejecting some new requests will not be a severe issue in the growing 5G future generation network. Because declining some new requests for accommodating HO ones, said otherwise, is guaranteeing less force termination of user requests in the system. That again contributes to the betterment of QoS.

C. *Queueing Analysis*

The proposed policy is a queuing system that is analyzed mathematically by using the birth-death process of Markov chain queueing analysis. Figure 3 shows the Markov chain of the general scheme. General scheme is defined as a system where no request is being prioritized over another and no bandwidth is being shared. Therefore, bandwidth is not being utilized to the fullest. On the other hand, bandwidth sharing takes place in our proposed scheme as depicted with the Markov chain in Figure 4. As a result, the number of states is increased so that more users can be served without being terminated at handoff situations, most likely occurring because of their high-speed mobility. This adaptive nature of the proposed approach allows for better resource utilization.

In the proposed mobility-aware handoff prioritized RAC scheme, the mathematical equations are used with the help of equations from related works [8] and [15]. The terminology of the queuing analysis is introduced in Table I.

For the system states $0 \leqslant k \leqslant X$, when there is enough bandwidth in the system, all the requests are allocated the requested bandwidth according to the content cache location. For system states $X+Y+1 \leqslant k \leqslant X+Z$, users with handoff requests are served with priority. This is the reserve BW in the proposed plan to treat HOs from user mobility. And from states $X+1 \leqslant k \leqslant X+Y$, the shared bandwidth is available which is allocated to the HO requests once being shared.

For queuing analysis, the probability that the system is in state k, is given by equation (1). In the proposed scheme, a new request is blocked if the system is in the state (X+Y) or larger. However, a handover request is dropped only if the system is in the state (X+Z). Thus, from equations (1) − (7) are used in computing the proposed policy.

TABLE I. NOMENCLATURE

Symbol	Definition
$P(k)$	Queueing probability at k^{th} state
$P(0)$	Queueing probability at 0th state
λ_n	Arrival rate of new request
λ_{hi}	Arrival rate of handoff request
λ_t	Total arrival rate of a state
μ_s	Departure rate of each state from 0 to X
μ_r	Departure rate of each state from (X+1) up to (X+Z)
$1/\mu$	Average call duration
$1/\eta$	Average cell dwell time
$P(H)$	Request probability
$P(B)$	New request blocking probability
$P(D)$	Handoff request dropping probability
X	Number of non-reserved states (bandwidth) for accommodating both new and handoff requests
Z	Number of reserved states (bandwidth) for accommodating only handoff requests
BW	Bandwidth utilization factor

$$P(k) = \begin{cases} P(0)\dfrac{(\lambda_n + \lambda_{h_i})^k}{k!\mu_s{}^k}; & 0 \le k \le X \\ P(0)\dfrac{(\lambda_n + \lambda_{h_i})^k}{k!\mu_s^k \prod_{r=X+1}^{k} \mu_r}; & X \le k \le X+Y \\ P(0)\dfrac{(\lambda_n + \lambda_{h_i})^{X+Y}(\lambda_{h_i})^{\{k-(X+Y)\}}}{k!\mu_s^X \prod_{r=X+1}^{k} \mu_r}; & X+Y \le k \le X+Z \end{cases} \quad (1)$$

where, $$P(0) = \left[\sum_{k=0}^{X} \frac{(\lambda_n + \lambda_{h_i})^k}{k!\mu_s^k} + \sum_{k=X+1}^{X+Y} \frac{(\lambda_n + \lambda_{h_i})^k}{k!\mu_s^X \prod_{r=X+1}^{k} \mu_r} + \sum_{k=X+Y+1}^{X+Z} \frac{(\lambda_n + \lambda_{h_i})^{X+Y}(\lambda_{h_i})^{\{k-(X+Y)\}}}{k!\mu_s^X \prod_{r=X+1}^{k} \mu_r} \right]^{-1}$$

$$P(\mathrm{H}) = \frac{\eta}{(\mu + \eta)} \quad (2)$$

$$\lambda_t = \lambda_n + \lambda_{h_i} \quad (3)$$

$$\lambda_{h_i} = \frac{\mathrm{P(H)(1-P(B))}}{(1-(\mathrm{P(H)}(1-P(\mathrm{D}))} \lambda_n \quad (4)$$

$$\mathrm{P(B)} = \sum_{k=X+Y}^{X+Z} P(\mathrm{k}) \quad (5)$$

$$P(\mathrm{D}) = \mathrm{P(X+Z)} \quad (6)$$

$$BW = \frac{(\lambda_n(1-\mathrm{P(B)})) + ((1-\mathrm{P(D)})\lambda_{h_i})}{(\mu+\eta)(\mathrm{X+Z})} \quad (7)$$

In the non-prioritized scheme, where there is no priority of handover calls over new calls, the total number of states remain X. But in the proposed RAC HO priority scheme the total number of states increases with the bandwidth sharing and reservation combination.

IV. Experimental Evaluation

The performance analysis has been done in MATLAB for the proposed scheme in comparison with the conventional scheme that does not include the idea of priority for mobility. Since the mobility of the user is prioritized in here, our primary focus remains on reducing the probability of dropping of handoff requests, notably. Along with the curves of dropping probability, we observed the patterns of the curves of blocking probability and bandwidth utilization to show the benefits of the proposed approach as well as its minor drawbacks.

At first, we compare the of probability of dropping, *P(D)*, probability of blocking, *P(B)*, and bandwidth utilization, *BW* of the proposed policy with the general non-prioritized approach with respect to the increasing total request arrival rate, λ_t (requests/sec).

For Probability of dropping, we analyzed a comparative study as illustrated in Figure 5. Here, it is shown that the dropping probability of handoff requests due to user mobility is significantly reduced in the proposed mobility aware RAC scheme. It has been made possible with the prioritized scheme. In the non-prioritized scheme, with the increasing number of requests coming into the system, the probability that a HO request will be dropped goes higher easily. But as per the proposed scheme, the reserved bandwidth allocated for HO requests and also the shared bandwidth works in favor to reduce the dropping probability. Hence, it can be an effective policy to employ in the massive high-speed user scenarios of 5G.

For Probability of blocking we get results as shown in Figure 6. It is shown that the blocking probability of new requests has slightly increased in the proposed mobility aware RAC scheme, but only by a maximum of 14%. This increase is a tradeoff that is to be done to prioritize the high-speed mobility in 5G scenario. But, as it is occurring due to accommodating more HO requests in the system, the policy ensures that less number of requests are forcefully terminated which were already admitted into the system and thus confirms the better quality of experience to the connected users.

For Bandwidth utilization the results are represented in Figure 7. The figure shows that the bandwidth utilization of the proposed mobility aware RAC scheme is slightly lower than the general scheme, although the growth pattern is more realistic. But since the values are very close as the arrival rate goes higher, it indicates that the proposed scheme could be a preferred option in highly dense areas.

Subsequently, we observe the before mentioned factors with respect to the increasing value of new request arrival rate, λ_n (requests/sec). This observation is meaningful since the effect of mobility i.e., hand-off probability, *P(H)* is directly related to the new request arrival rate, λ_n.

With the increasing new request arrival rate, the dropping probability of handoff requests of the two schemes is illustrated in Figure 8. Here, it is graphically showed that the dropping probability of handoff requests due to user mobility is significantly reduced (by almost 35%-40%) in the proposed mobility aware handoff priority based RAC scheme. Such minimization of dropping probability has occurred since our policy gives priority to HO requests which is not the case in the non-prioritized general scheme. In a general scheme, when higher number of requests comes into the system, the handoff probability increases as well. But in our case the reserved bandwidth allocated for handoffs serve the purpose of minimizing dropping probability. Moreover, the bandwidth being shared between an MBS and an SBS works in favor to reduce the effect of user mobility upon system's quality of service. Therefore, the proposed policy is an efficient one in the 5G high-speed user scenario.

In Figure 9, it is shown that the blocking probability of new requests has slightly increased in the proposed mobility aware RAC scheme, that also by a maximum of 21%. This increase happened due to acknowledging handoffs with higher priority over new requests. But, as our approach intends to serve more users with higher mobility, this will not be a serious problem in the emergent 5G future generation network. With significantly better dropping probability, one of our major concerns is being treated successfully. That is, requests are less likely to be terminated while being handed over from one base station (BS) to another. Thus, lesser termination is another way of improving QoS through efficient management of cache requests.

Figure 10 depicts that the Bandwidth utilization in the proposed RAC scheme gradually grows high with the increasing arrival rate. The growth pattern is also more realistic than the general scheme. The somewhat lower value in our scheme in the low arrival rate region is not an issue. The reason is that in 5G the arrival rate will be higher than the usual. So, Figure 10 connotes that the proposed RAC scheme is a preferred option in the 5G high-speed and high-arrival rate scenario.

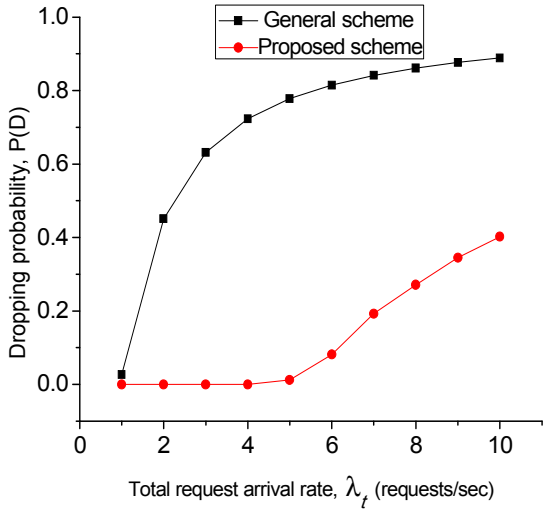


Figure 5: Comparison of handover request dropping probability

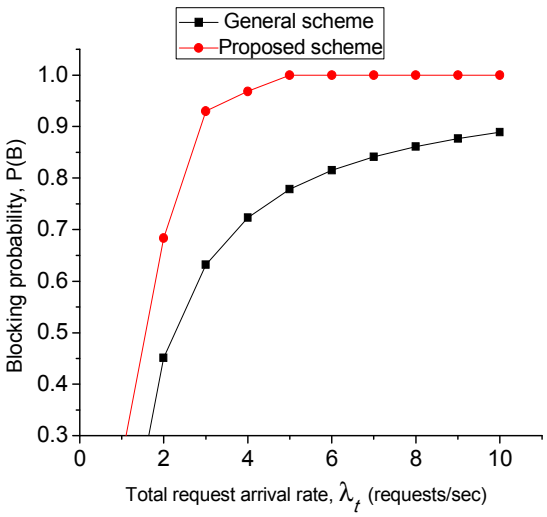


Figure 6: Comparison of new request blocking probability

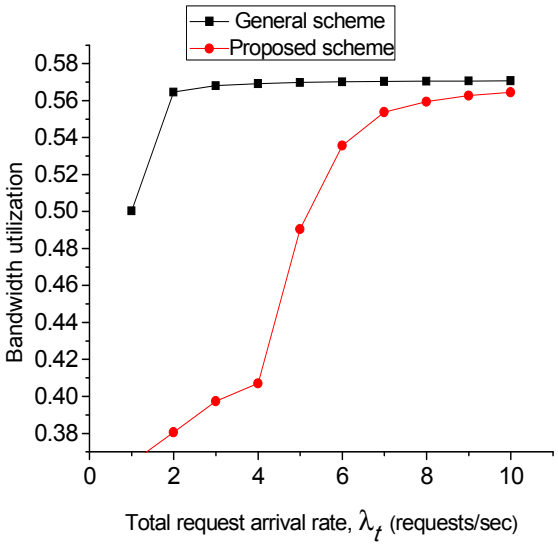


Figure 7: Comparison of bandwidth utilization

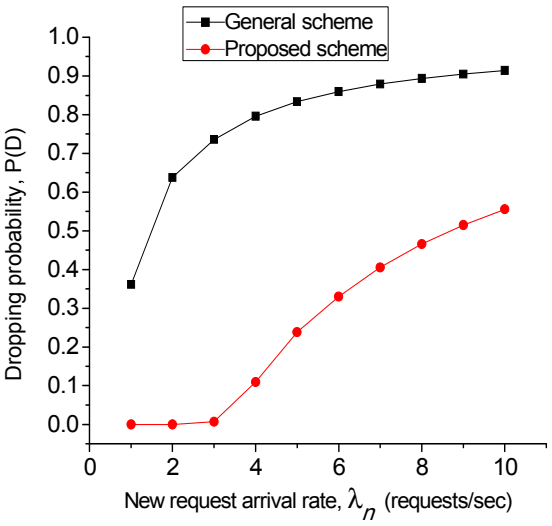


Figure 8: Comparison of handover request dropping probability

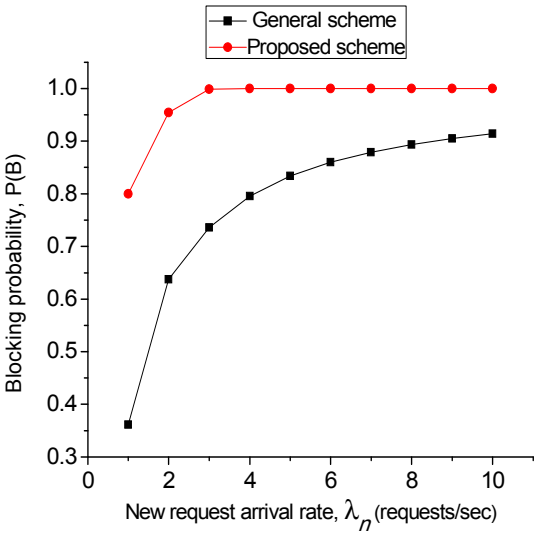


Figure 9: Comparison of new request blocking probability

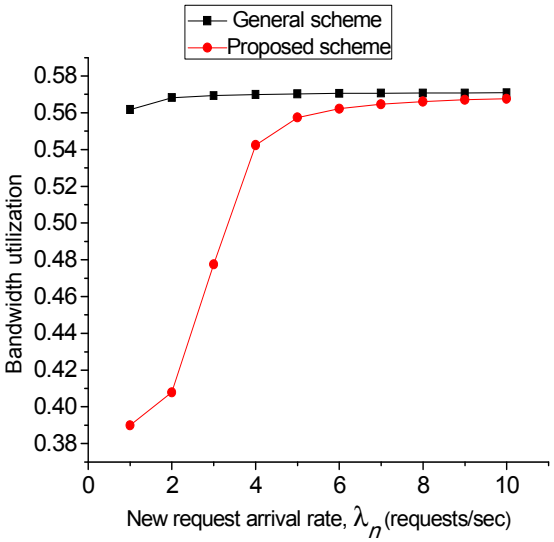


Figure 10: Comparison of bandwidth utilization

Now, from the discussed numerical analysis and graphical performance evaluation, we can conclude that from the user mobility aspect of 5G communication the proposed RAC HO prioritized scheme successfully reduces the dropping probability occurring due to handoff requests from high-speed mobile users on a significant scale. Lower dropping probability ensures a lesser number of force terminations of the already admitted requests. And the slight increase in blocking probability is not an alarming concern since the already admitted requests are facing less forced termination because of the new requests in the considered 5G scenario. Consequently, the mobility of the user is dealt with priority and requested cache is being efficiently allocated in the proposed scheme.

V. CONCLUSION

This paper has presented a handoff priority-based Request Admission Control (RAC) scheme. The proposed scheme allows for efficient cache management for content requests in a user mobility-aware heterogeneous fifth generation (5G) system by utilizing the existing bandwidth resources. The proposed scheme majorly addresses the area of prime effect of high-speed user mobility, which is the probability of dropping of hand-off requests. As the experimented results show, a minor increase in blocking probability of new requests was accepted while reducing the probability of dropping of hand-off requests significantly. Moreover, in this handoff prioritized request admission control (RAC) based cache management indicates further improvements in bandwidth utilization.

We will focus on enhancing the accuracy and adaptability of the proposed scheme in our future works. Besides, we will concentrate on simultaneously minimizing two performance metrics- the probability of dropping a handoff request and the probability of blocking a new request in order to develop a scheme for class-based cache request distribution and higher bandwidth utilization.

REFERENCES

[1] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective." *IEEE Wireless Communications* 21.3 (2014): 118-127.

[2] I. Chih-Lin, S. Han, Z. Xu, Q. Sun, and Z. Pan, "5G: rethink mobile communications for 2020+." *Phil. Trans. R. Soc. A* 374.2062 (2016): 20140432.

[3] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies." *IEEE access* 3 (2015): 1206-1232.

[4] M. Fallgren and B. Timus, "Scenarios, requirements and KPIs for 5G mobile and wireless system." *METIS deliverable D* 1 (2013): 1.

[5] Industry Proposal for a Public Private Partnership (PPP) in Horizon 2020 (Draft Version 2.1), Horizon 2020 Advanced 5G Infrastructure for the Future Internet PPP. [Online]. Available: http://www.networks-etp-eu/fileadmin/user_upload/Home/draft-PPP-proposal.pdf

[6] 3GPP TR 38.913: "5G; Study on Scenarios and Requirements for Next Generation Access Technologies", June. 2016. [Online]. Available: https://www.etsi.org/deliver/etsi_tr/138900_138999/138913/14.02.00_60/tr_138913v140200p.pdf

[7] P. Fan, J. Zhao, and I. Chih-Lin, "5G high mobility wireless communications: Challenges and solutions." *China Communications* 13.2 (2016): 1-13.

[8] M. Z. Chowdhury, Y. M. Jang, and Z. J. Haas, "Call admission control based on adaptive bandwidth allocation for wireless networks." *Journal of Communications and Networks* 15.1 (2013): 15-24.

[9] X. Liu, J. Zhang, X. Zhang, and W.Wang, "Mobility-aware coded probabilistic caching scheme for MEC-enabled small cell networks." *IEEE Access* 5 (2017): 17824-17833.

[10] Cisco, Visual Networking Index. "Global Mobile Data Traffic Forecast Update, 2015–2020 White Paper." *Document ID*958959758 (2016).

[11] M. Z. Chowdhury, M. A. Hossain, S. Ahmed, and Y. M. Jang, "Radio resource management based on reused frequency allocation for dynamic channel borrowing scheme in wireless networks." *Wireless Networks* 21.8 (2015): 2593-2607.

[12] N. G. M. N. Alliance, "5G white paper." *Next generation mobile networks, white paper* (2015): 1-125.

[13] J. G. Andrews, et al. "What will 5G be?." *IEEE Journal on selected areas in communications* 32.6 (2014): 1065-1082.

[14] D. Liu, B. Chen, C. Yang, and A. F. Molisch, "Caching at the wireless edge: design aspects, challenges, and future directions." *IEEE Communications Magazine* 54.9 (2016): 22-28.

[15] M. A. Rahman, M. A. Hossain, S. Ahmed, and M. Z. Chowdhury, "A new guard-band call admission control policy based on acceptance factor for wireless cellular networks." In 2014 *International Conference on Informatics, Electronics & Vision (ICIEV)*, pp. 1-5. IEEE, 2014.

[16] M. A. Hossain, S. Ahmed, and M. Z. Chowdhury, "Radio Resource Management for dynamic channel borrowing scheme in wireless networks." In 2014 *International Conference on Informatics, Electronics & Vision (ICIEV)*, pp. 1-5. IEEE, 2014.

[17] S. Zhang, N. Zhang, P. Yang, X. Shen, "Cost-effective cache deployment in mobile heterogeneous networks." *IEEE Transactions on Vehicular Technology* 66.12 (2017): 11264-11276.

An Intelligent Assistive Tool for Alzheimer's Patient

Kazi Shahrukh Omar[1], Afia Anjum[2], Tahrima Oannahary, Md. Rezaul Karim Rizvi, Diana Shahrin, Tasmiah Tamzid Anannya, Sanjida Nasreen Tumpa, Md Mahboob Karim, Muhammad Nazrul Islam, Md Fazle Rabbi
Department of Computer Science and Engineering, Military Institute of Science and Technology (MIST), Dhaka-1216, Bangladesh
Email: {[1]qshahrukh41, [2]afia.anjum70}@gmail.com

***Abstract*—Alzheimer's Disease (AD) is a chronic neurodegenerative disease that causes to develop dementia. Alzheimer's patients find it hard to remember recent events, reason and even to recognize people they know. As the disease advances, symptoms can include difficulty with language, disorientation including getting lost, mood swings, loss of motivation, lack of self-awareness and overall behavior. Though a limited number of IT based solutions exist to provide support for Alzheimer's patients, but most of these provide very isolated services either for the patients or for the caregivers. The objective of this research is to propose an assistive tool for Alzheimer's patients and their caregivers to provide support like health monitoring, assist to find lost items, provide reminder to take medicine and assist to monitor patient's location. A light-weighted evaluation study was carried out with 15 participants. The evaluation study showed that the proposed system was effective and usable for the patients and their caregivers.**

***Keywords*—Alzheimer, mobile application, heart rate, assistive tool, intelligent system, sensors.**

I. Introduction

Alzheimer's disease is a progressive disease in which the destruction of memory cells causes slow depreciation in memory, thinking and reasoning skills. A great deal of difficulties are faced by an Alzheimer's patient including short term memory loss, aggression, loss of communication skill, repetitive behaviour, anxiety, agitation etc [1].

According to [2], 24.3 million people have Alzheimer's disease today, with 4.6 million new cases every year (one new case every 7 seconds). The number of people affected will get doubled every 20 years to 81.1 million by 2040. The increasing rate of Alzheimer's patient is not uniform for all over the world. The experts have forecast the increasing rate by 100% between 2001 and 2040 for the developing countries whereas the rate is more than 300% for India, China, and their south Asian and western Pacific neighbors. The number of patients facing dementia caused by Alzheimer's Disease is increasing in the entire world including Asian subcontinent [3].

With the advancement of technology, it is seen that, technology plays a vital role as caregivers for those people who cannot take care of themselves [4], [5]. Therefore, the objective of this research is to develop an intelligent assistive tool for Alzheimer's patients to ease their daily life by measuring their heart rate, locate their personal belongings, monitor medication taking status and find their current location, while these necessary states of patients will be delivered to their family members and caretakers through a mobile application.

The rest of the paper is organized as follows. Section II discusses the related researches previously done in this area. Section III presents the conceptual framework of the system. The detail implementation of the proposed system is discussed in Section IV followed by the evaluation of the system in section V. Finally, section VI concludes the article.

II. Literature Review

A significant number of research has been carried on the design, development and evaluation of the assistive tool for the special need people, blind persons, aging population, and for the Alzheimer's patients [6]–[9]. This section will briefly introduce some work focusing to the development of assistive tool for Alzheimer's patients.

A limited number of papers were conducted focusing to mobile applications and software systems. In [10], Yamagat et al. showed that tablets and iPads could be utilized with Alzheimer's patients portraying favorite music and family photographs, as it helped them remember. In another study, Armstrong et al. [11] highlighted the potential explicit and implicit interaction techniques for Alzheimer's patients to receive care using smart phone buttons, touch screens, motion sensing and voice recognition. Similarly, in [9], Donnelly et al. developed a mobile application to provide a virtual caregiver which would give frequent memory cues through video reminders for mild AD patients. A medicine dispensing system is proposed in [12] that reminded the user to take medicine at the predetermined time interval and informed caregiver if they had taken the medicine or not. Some other studies [13]–[15] discussed the smart medication systems that would help patients who require medications several times a day. These devices would deliver medication in the proper chronological order and at correct time intervals.

Again some studies were conducted focusing on hardware systems. For example, an IOT based wearable ECG wireless sensor system was proposed in [16] which monitors whereabouts of Alzheimer's patients and triggers alarms to caretakers in case of emergencies. The use of pervasive computing technology to provide better live-at home facility for the elderly citizens are discussed in [17] and [18]. However, their system neglected the possibility of the outdoor experience and specific cases of visiting outdoor by the old people. In

TABLE I
SUMMARY OF PREVIOUS WORKS

References	*Technology*	*Objective*
[9]	Internet based technology	To give video reminder about daily activities to the patient
[12]	Real time clock, timer, sensor to detect presence of bottle, audible alert	To work as a medication reminder and overdose safeguard
[13]	Computer controlled medication dispensing system	To ensure the proper medication and timing
[15]	Mechanical technology	To indicate the number of pills that has been taken
[16]	IOT based wearable ECG wireless sensor	To monitor whereabouts of Alzheimer's patients and trigger alarm in case of emergencies
[17]	Voice interface, smart phone, ultrasonic location interface	To make the patient do their daily works and taking care of patient's medication
[19]	Multi-agent systems together with the use of RFID technology	To guarantee that patient is getting the right care from the assigned nurse and doctor
[20]	Electronic pedometer	To monitor a person's movement and show heart rate
[21]	Sensors, wireless communication	To record the patient's movement and daily behavior
[22]	GPS and GSM technology	To track the location of patient
[23]	Microchip Wireless Technology	Tracks patient in a range

[19], researchers presented an autonomous intelligent agent for monitoring Alzheimer's patient's health properly. In [20], a fitness monitoring device was proposed that included a wireless heart rate monitor which was coupled to the electronic pedometer without any wiring. The wireless heart rate monitor provided a heartbeat signal indicative of an exertion level of the user, while a visual display was provided to display the pedometer functions and the heart rate signal. In [21], a device was described, which was attached to the patient's body, monitoring the movement and behaviour of the patient and informing the caregivers whenever unexpected or alarming incidents occurred. According to [22], real-time GPS-GSM tracking could be helpful for dementia patients to ensure their safety when they go outside of their house. In [23], authors emphasized on memory loss of an Alzheimer's patient who tends to forget things and came up with the solution of Wifi and wearable technologies. Google map-based tracking system and notification to caretaker was an efficient way to help these people.

A summary of the previous works is shown in Table I. In sum, the previous studies suggested separate modules to assist the patients from different angles. It is certain that no collaborative system was developed to help the patients integrated with physical health monitoring, smart medication system, smart tracker for lost items and locating patient both inside and outside home. In addition, inclusion of caregiver and smart phones were also missing parts of the puzzle. Thus, this research aimed to develop an autonomous system integrating all necessary features (constant monitoring, reminder for taking medicine, monitoring heart condition and finding lost items) including relatives and caregiver to make Alzheimer's patient's life easier.

III. CONCEPTUAL FRAMEWORK

A conceptual framework was proposed to develop an intelligent assistive tool for Alzheimer's patients as shown in Fig. 1. The framework consists of four modules: heart rate monitoring, lost item finder, smart medicine box and lost patient locator. The functionalities of these modules can be accessed/used from a single mobile application that can be used by the patient's caregiver or nurse. Each of these modules will perform independently to help Alzheimer's patients.

A. Heart Rate Monitoring

The heart monitoring device allows to monitor the heart rate of the patient and notify the concerned person in case of irregular heart rate. If the heart rate of the patient rises above a threshold value, the nurse/concerned person will be informed about the patient's condition. Based on the situation the nurse will take first aid measures and call on doctor in case of emergency (see Fig. 2).

B. Lost Items Finder

It is hard for Alzheimer's patients to keep track of their important things like spectacles, keys, wallet. So, a tracker is attached to their important accessories so that whenever they lose something, they can find it using a mobile application which will buzz an alert to that accessory (see Fig. 3).

C. Smart Medicine Box

Alzheimer's patients might forget to take their medicines on time. Thus, continuous monitoring of patient's medication is another vital thing. Whenever the patient will take medication, smart medicine box will activate a WiFi module and send data of medicine intake to server. If the patient does not take medication on time, then a text alert will be sent to the concerned person/nurse (see Fig. 4).

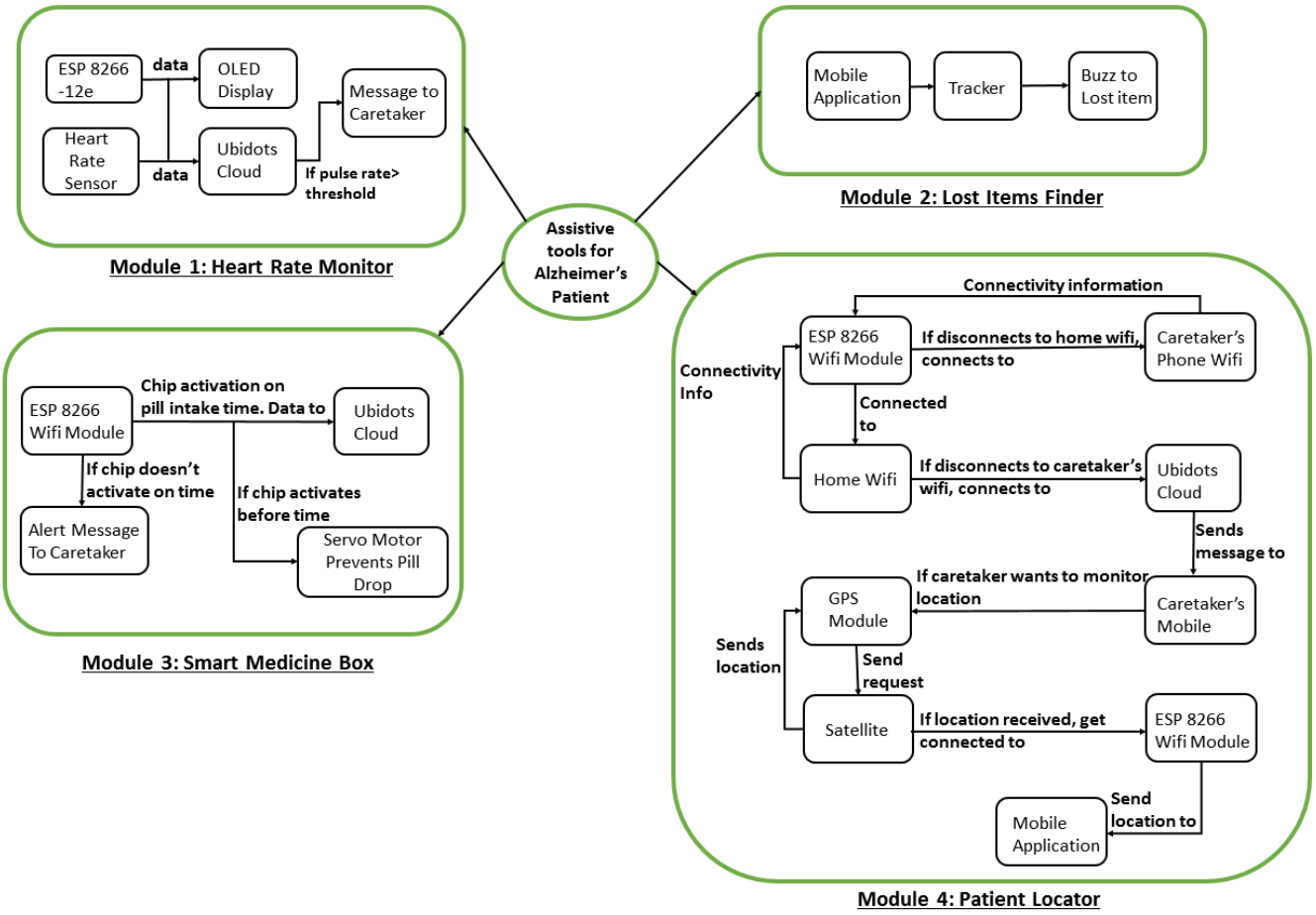


Fig. 1. Conceptual framework of the proposed assistive tool.

D. Lost Patient Locator

If Alzheimer's patient goes out of home unattended, he/she might get lost. The GPS-WiFi tracking system will always keep track of the patient's whereabouts. If the patient goes out of boundary line of the house or is far away from their caretaker, the caretaker will be notified by a text alert and will be able to continuously monitor patient's location through a mobile application (see Fig. 5).

The four individual modules creates such a system environment for the Alzheimer's patients so that their medication and health is on check, they can easily find necessary items if lost and their whereabouts and location can be tracked. It also helps the caretaker of the patient as the system makes sure the patient is always within reach and does not get lost which is a common scenario in case of Alzheimer's patient.

IV. Developing The Assistive Tool

The assistive tool was developed in two parts: hardware (see Fig. 6) and a mobile application. The required hardware components are shown in Table II. The Ubidots cloud platform was used to develop the mobile application. The mobile application holds the monitoring and assistive control of four features that include patient locator, smart medicine box, heart rate monitor and lost items finder. Fig. 7(a) portrays the home screen of Alzheimer's assistant application.

(a) Patient locator - The GPS tracking system was used to keep track of Alzheimer's patient's movements. The tracking system was divided into phases. In the first phase, if the patient goes outside of home with his/her caretaker and gets lost, the caretaker can know the patient's location through a mobile application (see Fig. 7(b)). The GPS module continuously fetches the location of the patient with latitude and longitude and sends the data to caretaker's mobile application via WiFi module.

TABLE II
Required Components of the System

Module	*Components*
Heart Rate Monitor	Esp8266 12e,Pulse sensor, OLED display,Battery, Wires
Smart Pill Box	Esp8266 12e,Servo Motor,Batteries and other power sources,OLED display,Wires
Lost Items Finder	Piezo buzzer, Hc-05 Bluetooth module,Arduino Uno,Battery
Lost Patient Locator	Esp8266 12e,Neo 6m GPS tracker,Battery

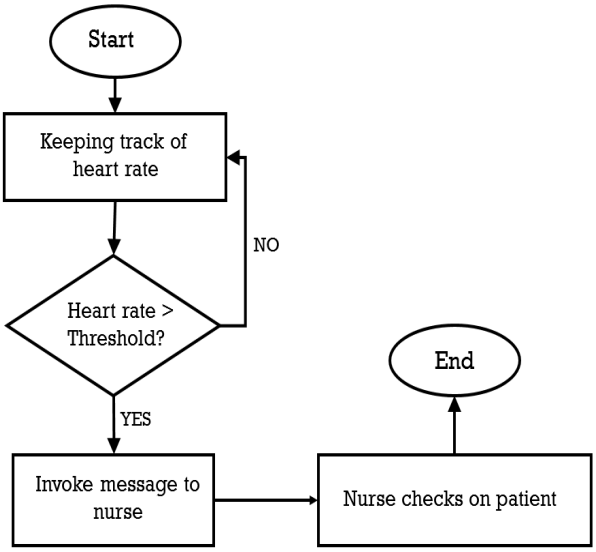


Fig. 2. Heart rate monitor

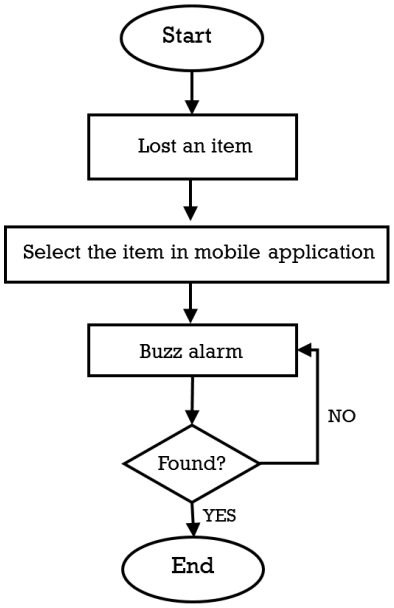


Fig. 3. Lost item finder

The second phase is that, if the patient goes outside of their home alone, the caretaker will be informed through a text alert and can see the patient's location through the application. This is done by an ESP-12E Wifi module which is also connected to the home Wifi. Whenever the patient steps outside of their home, the Wifi module will try to connect with caretaker's mobile network. If it fails to connect, it assumes that the patient has gone outside alone and so it sends a message to the caretaker's mobile to notify him/her.

(b) Smart medicine box - A smart medicine box was developed to keep track of the pill intakes by the patient. Whenever it is time for the patient to take medicine, the

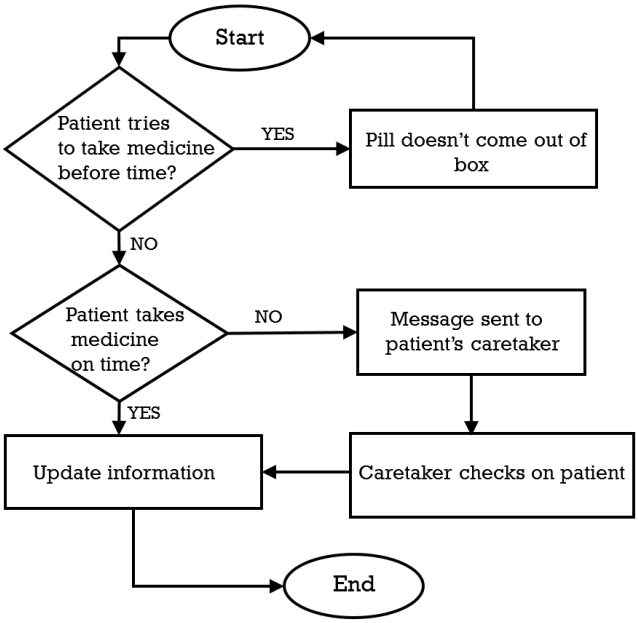


Fig. 4. Smart medicine box

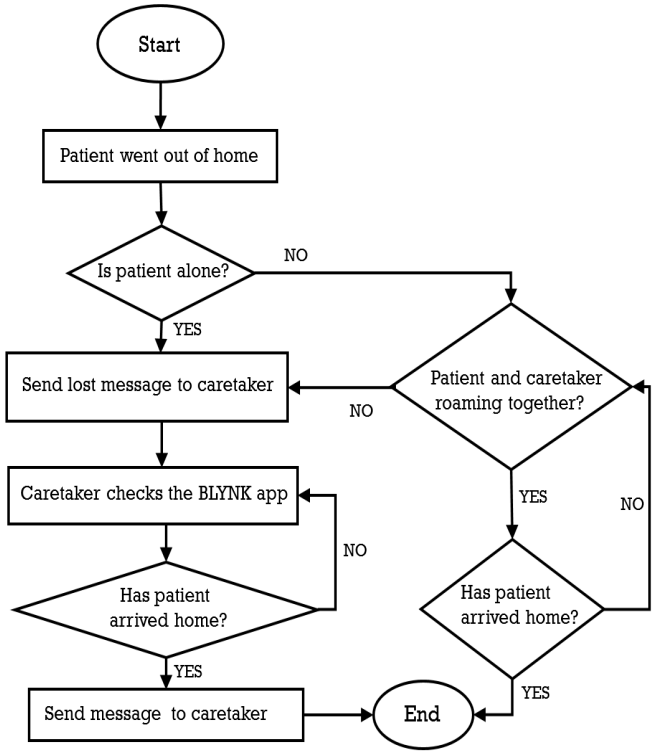


Fig. 5. Lost patient locator

medicine box will automatically open. Also, a buzzer will make sound to remind the patient that it is pill intake time. If the patient still does not take medicine, then an alert message will be sent to patient's caregiver. After the patient takes medicine, the box will be closed again. It will again open

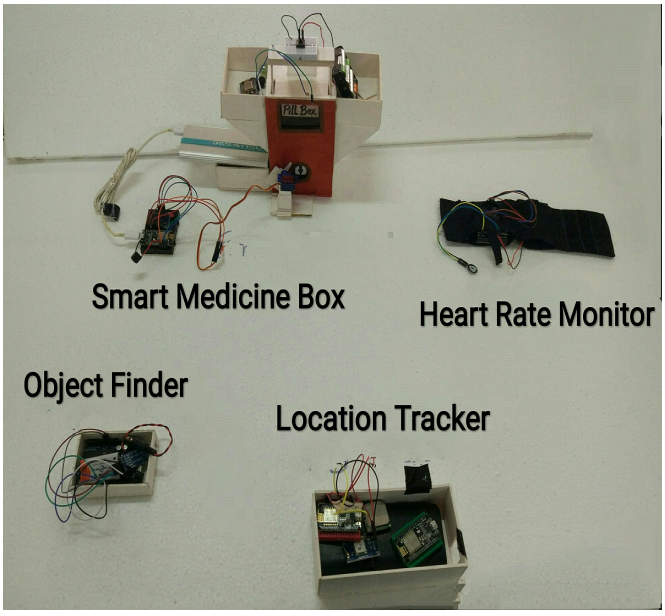


Fig. 6. Hardware parts of the proposed system

when the next pill intake time will come. An OLED display is also attached with the medicine box that displays the time remaining for next pill intake.

(c) Heart rate monitor - The heart monitoring system takes the pulse rate of the patient through pulse sensor attached to patient's fingertip. If the Heart rate exceeds a threshold value (e.g. 150 BPM) immediately a message will be sent to the caretaker.

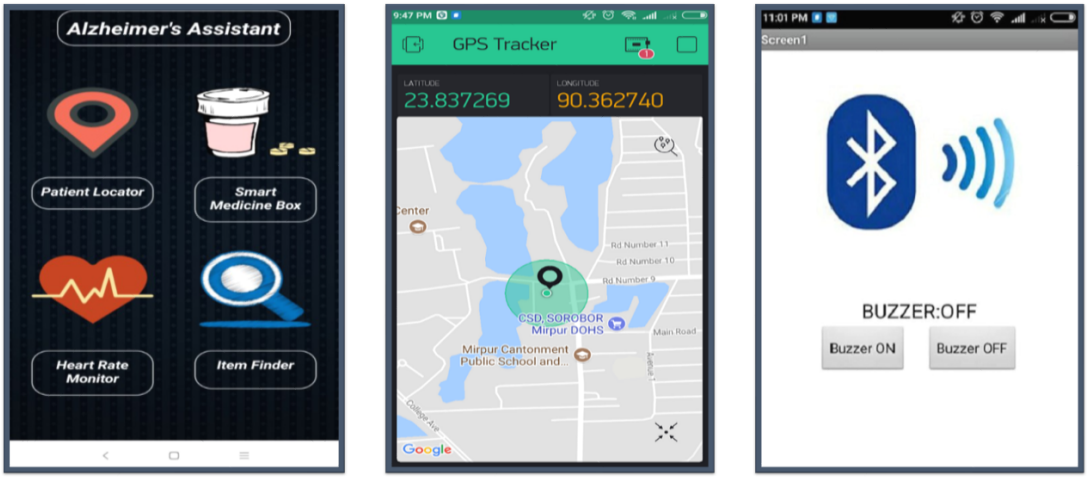
(d) Lost items finder - Alzheimer's patient loses their belongings very easily. The object finder helps the patient to find the lost items. If the lost items are at Bluetooth range of their mobile, the application (see Fig. 7(c)) will ring a buzzer. With the help of the application and by following the sound the lost items can be found.

V. Demonstration and Evaluation

A Focus Group Discussion (FGD) [24] was carried out to evaluate the performance and usability of the proposed assistive tool. The FGD session was replicated with randomly selected 10 final year under-graduate students and 5 faculty members. During the FGD: firstly, the proposed tool was demonstrated; after that we gave them opportunity to explore and use the tool for 20-30 minutes. At the end, participants were asked to provide their opinion about the system's functionality, usability and usefulness. 13 out of 15 (87%) participants found that the system performs accurately each function, while all of them (100%) were opined that the system was easy-to use. Some of them suggested to include more features related to patient's health, while few other participants suggested to improve the safety of the device by putting the circuits inside a box.

VI. Conclusion

In this paper, an assistive tool was designed and developed for Alzheimer's patients. In respect to the earlier works, this research contributed in two folds: firstly, several features are integrated in a single system, and secondly, a mobile application was developed to utilize those features effectively from a single platform. Thus the proposed system provides an integrated round the clock service to the patients both in home and outdoor. Alzheimer's patients and the caregivers of patients will be greatly benefited from the system, since the system will provide support for constant monitoring, provide assistance for taking medicine, look after the heart condition of the patient, and help to find patient's belongings whenever needed.



(a) Home screen of Alzheimer's assistant app

(b) Patient locator

(c) Item Finder

Fig. 7. Alzheimer's assistant app

However, there are still certain limitations of this proposed system. The system was built as a prototype. It was developed in an academic environment and thus no evaluation study with real-users were obtained against it. Also, the smart medicine box module does not work without a Wifi connection. These issues will be addressed in our future work. Apart from these, the system modules can be developed as wearable. Furthermore, few more features can be added to provide more medical/health details of the patient.

ACKNOWLEDGEMENT

This research was supported wholeheartedly by faculties of Department of Computer Science and Engineering, MIST. We would like to show our gratitude for their great support. The first two authors of this paper have contributed equally in terms of implementation of the proposed system, while all authors have contributed to literature review, evaluation study and writing of the article.

REFERENCES

[1] J. Paton, K. Johnston, C. Katona, and G. Livingston, "What causes problems in alzheimer's disease: attributions by caregivers. a qualitative study," *International journal of geriatric psychiatry*, vol. 19, no. 6, pp. 527–532, 2004.

[2] C. P. Ferri, M. Prince, C. Brayne, H. Brodaty, L. Fratiglioni, M. Ganguli, K. Hall, K. Hasegawa, H. Hendrie, Y. Huang *et al.*, "Global prevalence of dementia: a delphi consensus study," *The lancet*, vol. 366, no. 9503, pp. 2112–2117, 2005.

[3] R. G. Logsdon, L. E. Gibbons, S. M. McCurry, L. Teri *et al.*, "Quality of life in alzheimer's disease: patient and caregiver reports," *Journal of Mental health and Aging*, vol. 5, pp. 21–32, 1999.

[4] A. Kintsch and R. DePaula, "A framework for the adoption of assistive technology," *SWAAAC 2002: Supporting learning through assistive technology*, pp. 1–10, 2002.

[5] K. Hawkey, K. M. Inkpen, K. Rockwood, M. McAllister, and J. Slonim, "Requirements gathering with alzheimer's patients and caregivers," in *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2005, pp. 142–149.

[6] T. Zaki, M. N. Islam, M. S. Uddin, S. N. Tumpa, M. J. Hossain, M. R. Anti, and M. M. Hasan, "Towards developing a learning tool for children with autism," in *2017 6th International Conference on Informatics, Electronics and Vision 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT)*, 2017, pp. 1–6.

[7] Y. Lee, C. Chiu, L. Jhang, and C. Santiago, "A self-reliance assistive tool for disable people," in *3rd International Conference on Control and Robotics Engineering 2018 (ICCRE)*, 2018, pp. 26–30.

[8] N. S. Khan, S. Kundu, S. Al Ahsan, M. Sarker, and M. N. Islam, "An assistive system of walking for visually impaired," in *International Conference on Computer, Communication, Chemical, Material and Electronic Engineering 2018 (IC4ME2)*, 2018, pp. 1–4.

[9] M. Donnelly, C. Nugent, S. McClean, B. Scotney, S. Mason, P. Passmore, and D. Craig, "A mobile multimedia technology to aid those with alzheimer's disease," *IEEE multimedia*, vol. 17, no. 2, pp. 42–51, 2010.

[10] C. Yamagata, J. F. Coppola, M. Kowtko, and S. Joyce, "Mobile app development and usability research to help dementia and alzheimer patients," in *Systems, Applications and Technology Conference (LISAT), 2013 IEEE Long Island*. IEEE, 2013, pp. 1–6.

[11] N. Armstrong, C. Nugent, G. Moore, and D. Finlay, "Using smartphones to address the needs of persons with alzheimer's disease," *annals of telecommunications-annales des télécommunications*, vol. 65, no. 9-10, pp. 485–495, 2010.

[12] R. B. Sagar, "Bottle-cap medication reminder and overdose safeguard," Aug. 12 2003, uS Patent 6,604,650.

[13] A. P. Passamante, "Smart automatic medication dispenser," May 19 1998, uS Patent 5,752,621.

[14] A. Naditz, "Medication compliance—helping patients through technology: modern "smart" pillboxes keep memory-short patients on their medical regimen," *Telemedicine and e-Health*, vol. 14, no. 9, pp. 875–880, 2008.

[15] E. H. Aiken, "Indicator for pill bottle," Sep. 14 2004, uS Patent 6,789,497.

[16] M. W. Raad, T. Sheltami, and E. Shakshuki, "Ubiquitous tele-health system for elderly patients with alzheimer's," *Procedia Computer Science*, vol. 52, pp. 685–689, 2015.

[17] S. Helal, C. Giraldo, Y. Kaddoura, C. Lee, H. El Zabadani, and W. Mann, "Smart phone based cognitive assistant," in *UbiHealth 2003: The 2nd International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications*, 2003.

[18] C. Giraldo, S. Helal, and W. Mann, "mpca–a mobile patient care-giving assistant for alzheimer patients," in *First International Workshop on Ubiquitous Computing for Cognitive Aids (UbiCog'02)*, 2002.

[19] J. M. Corchado, J. Bajo, Y. De Paz, and D. I. Tapia, "Intelligent environment for monitoring alzheimer patients, agent technology for health care," *Decision Support Systems*, vol. 44, no. 2, pp. 382–396, 2008.

[20] K. Y. Sham and P. L.-K. Wong, "Fitness monitoring device having an electronic pedometer and a wireless heart rate monitor," Apr. 6 1999, uS Patent 5,891,042.

[21] M. Student, N. Karasikov, D. Tamir, and R. N. Bar-On, "Remote monitoring system for alzheimer patients," Jun. 12 2008, uS Patent App. 11/913,439.

[22] M. H. Kasliwal and H. Y. Patil, "Smart location tracking system for dementia patients," in *Advances in Computing, Communication and Control (ICAC3), 2017 International Conference on*. IEEE, 2017, pp. 1–6.

[23] S. Saranya and P. JesuJayarin, "An efficient tracking device for alzheimer patient using miwi," *International Research Journal of Engineering and Technology*, vol. 4, no. 4, pp. 3365–3371, 2017.

[24] R. A. Krueger and M. A. Casey, "Designing and conducting focus group interviews," [Online; accessed 02-Feb-2018]. [Online]. Available: http://web.worldbank.org/archive/website01028/WEB/IMAGES/SDP_36.PDF#page=10

A Prototype of Automated Vault Locker Solution for Industrial Application

Shafayet Hossain
Dept.of Electrical and Electronic Engineering
Atish Dipankar University of Science and technology
Dhaka, Bangladesh
shafayethossain774@gmail.com

Mazid Ishtique Ahmed
Dept.of Electrical and Electronic Engineering
Atish Dipankar University of Science and technology
Dhaka, Bangladesh
mazid.i.ahmed@ieee.org

Md. Niaz Mostakim
Dept.of Electrical and Electronic Engineering
Atish Dipankar University of Science and technology
Dhaka, Bangladesh
niazmostakim@ieee.org

Abstract— In this paper, an enhanced security solution is proposed for industry grade safety vault system. Electronic security and safety have been offering robust mechanism for implementation in ensuring proper shelter of key documents and goods in offices, banks, institutions, libraries, laboratories etc. Development of various sensors have enabled systems to have preventive and corrective measures in this regard significantly. In order to deliver a concrete security solution for critically important and confidential documents and goods, we proposed an Automated Safety Vault with Double Layered Defense Mechanism. The solution comprised of an Electronic Lock driven by password verification and a Biometric authentication for users using a Fingerprint scanning and sensing tool. Both of these two layers ensured the authenticity of the user by preventing any unauthorized access to the Vault. Moreover, actions of PIR and LDR sensors would also safeguard the vault from any breakout access other than the double layered defenses by being positioned within the periphery of the vault. The system was then implemented in a prototype scope for testing and validation of the proposals. The implemented system and testing data showed that the Automated Safety Vault with all its security features had successful operation. The specification of the whole system as well as the results are documented in this paper.

Keywords—vault, locker, arduino, pir, ldr, safety, fingerprint, biometric, numeric

I. INTRODUCTION

Security is of primary concern and in this busy, competitive world, human being cannot find ways to provide security to his confidential belongings manually. Instead, he finds an alternative which can provide a full-fledged security as well as atomized. In the ubiquitous network society, where individuals can easily access their information anytime and anywhere, people are also faced with the risk that others can easily access the same information anytime and anywhere. Because of this risk, personal identification technology, which can distinguish between registered legitimate users and imposters, is now generating interest. Various attempts are made for providing security for all domiciles. Up to date, complete security is not discovered but continuous enforcement of modern day technologies can always contribute to add the next level.

In automatic security systems generally passwords, identification cards and PIN verification techniques are being used but the disadvantage is that the passwords could be hacked and a card may be stolen or lost. The most secured system is fingerprint recognition because a fingerprint of one person never matches the other. Biometrics studies commonly include fingerprint, face, iris, voice, signature, and hand geometry recognition and verification. Many other modalities are in various stages of development and assessment. Among these available biometric traits fingerprint proves to be one of the best traits providing good mismatch ratio, high accurate in terms of security and also reliable. That's why we are motivated to do this project.

Many researchers have implemented security system and continuously they are working to develop them. Crystalynne D. Cortez et.al designed a microcontroller based biometric locker system with short message service. This proposed system has implemented on finger print sensor and GSM module. Finger print sensor is used to scan finger after matching it with the saved pattern and unlock the locker. GSM module sends a text message consist of an information when an unrecognized fingerprint was encountered [1]. Another finger print-based security system has implemented which is finger print-based door locking system. If the fingerprint matches, the door will be opened automatically otherwise the buzzer connected to an audio amplifier will be activated so that the people near the surroundings will get an alert. This system has proposed by A. Aditya Shankar et.al [2]. Sheikh Izzal Azid et al. have proposed a low cost home security system [3]. In this system, different sensors have used to secure home. They have used motion sensor, smoke sensor, humidity and light sensors. Home security based on ZigBee technology and GSM/GPRS network has implemented by J. Tan, Q. Wang et.al. They implemented this system form real time surveillance and remote monitoring of the home [4]. Another wireless module based security system has been implemented by Mohd Adib B. Sarijari et al. In this system ZigBee IEEE, 802.15.4 standard has used to develop smart home environment [5]. M.K.Hossain et al have implemented a three-layer security [6]. In this security system, they have used NFC tag; fixed password and motion .To open the door these three requirements have to provide. Arduino and GSM based advanced security system has been implemented by C. Malarvizhi et al. In his system chloroform sprayer is used to unconscious the thief inside the home [7]. Another security system based on Arduino had been implemented by Saurabh Kul et al [8]. This system has implemented by using Keypad, fingerprint, motion and PIR sensor and they implemented their system in five separate parts.

Therefore, the study of all these approaches intent to solve the dire problem of security to critical systems of an institution with proper authorization. Such systems are only accessible by the designated users and not by the mass. Moreover, the solution must ensure the obstruction of all possible ways of violation of security within the periphery of the secured area. It is also expected to regulate the access to certain users divided in different using capacity groups.

II. The Proposed Solution of Double Layered Defense Mechanism

The key objective of the proposed solution is to obtain the result of collective decision making based on the data provided by multiple sensors. To attain this performance, a controller is employed to analyze the sensor data and to generate desired controlled variables in order to ensure the security of the vault. Arduino UNO platform, is a modern day phenomenon for achieving prototypes of such controlling tool with various peripherals.

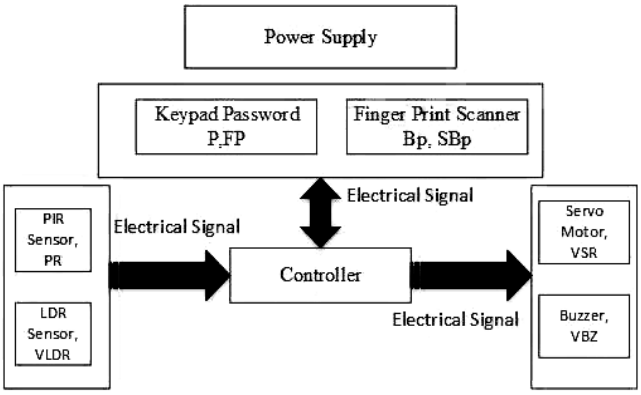


Fig. 1. Block diagram of proposed system.

As shown in fig 1, the controllers played the key role of decision making populated with the biometric and the physical motion detecting sensors and generated the controlled response or regulated to drive the servo motor of the vault door or buzzing the alarm due to unauthorized entry. However, the proposed solution does not only ensure vault security, but also provides certain degree of assistance in designing other similar prototypes involving biometric as well as physical sensing.

In addition to these conditions, the objectives were to develop a prototype featuring:

- Automated vault locking mechanism driven by actuated servo motor controlled and directed by ensuring double layers of safety.
- Allowance of access by verifying securely stored pre-registered password by the user which could confirm the primary layer of safety to the vault in terms of unauthorized entry.
- Pattern recognition of biometric data of user's fingerprint which would be pre-enrolled by scanning the fingerprints of the thumb and index fingers. The patterns were stored in the system database. It ensures the secondary layer safety to the secured vault after successful adherence to the primary layer of password protection.
- Motion detection in case of extreme violation of both layers of protective measures using PIR motion sensor and LDR light dependent sensor to assist physical intervention of security in charge from the law-enforcement agencies.

III. Components of the Prototype Implementation

The implementation of the prototype was based on the combination of the individial blocks of fig. 1. A brief detail of the components which were used to build each block to enable the system to behave to obtain the expected results are as following –

A. Power Supply: 5V DC power supply was prepared using simple bridge rectifier with ripple filters.

B. Keypad: A numeric keypad of 4x4 structures was used as the input method of password entry. Only the numeric spaces of the keypad were used for password generation and verification of the password at the time of operation.

C. Fingerprint Scanner: A biometric fingerprint scanner manufactured by Shenzen CAMA, China was used.

D. PIR Sensor: Used for motion detection within the vault premise.

E. LDR: Used for detection of any unnecessary lights within the vault.

F. Servo-motor: A 10 gram Servo motor with No load speed indicated as 0.12 seconds/60 at a rotary angle of 180 degrees. The operating voltage is with 4.8 – 6 Volts.

G. Buzzer: A locally sourced buzzer used outside the vault premise to alert the security enforcement personnel about any unauthorized entry into the vault.

H. LEDs: Used as Status Indicator LEDs (SIL), placed above the vault door for indication of modes of operations as mentioned in table III.

I. Controller: Two Arduino Uno hardware solution packs were used as the controlling unit of the proposed solution which initiated all the processes indicated in fig. 2, 3 and 4.

IV. Working Principle of the Proposed Solution

As already mentioned in section II, the controller served as the brain of the proposed system which being specified with the sensors' data, set the system trajectory to follow security compliances. Moreover, the components as mentioned in section III, were arranged in such a manner that the controller was driven to follow three separate yet inter-connected algorithms each gathering the sensed variables which are listed in table I. All the sensors delivering the deciding factors to operate and ensure security of the vault are segregated in three layers as well.

TABLE I. SENSOR DETAILS

Sensors	Sensed Variables	
	Variable	*Reference Values/status & Output type*
PIR	Motion_detect	1
LDR	Light_detect	Logic, LOW
Keypad (initial configuration)	Prestored_password	CHAR, ARRAY
Keypad (user operation)	User_password	CHAR, ARRAY
Biometric (initial scan of fingerprint)	Prestrored_finger_print	IMAGE, BIN
Biometric (user operation)	User_finger_print	IMAGE, BIN

TABLE II. LAYER DETAILS

Layers	Layer Sensors	Leading to -
Violation Status	PIR, LDR	Layer 1 or Vault Status
Layer 1	Keypad	Layer 2
Layer 2	Biometric Fingerprint Scanner	Vault Operation or Vault Status

The vault was considered to be a square shaped structure and naturally, in safe and normal operation there would be no presence of any light. Electric lighting system would still be disbaled when the vault was in the locked state and the controller would only lit up the space of the vault when it was securely accessed through Layer 1 and Layer 2 approach.

The first layer was responsible for primary authentication of a user. The verification in this cased was based on the comparison of a prestored password and a password input from the same user through a numeric keypad. The operation of Layer 1 was followed as per the flow chart shown in fig. 2. The authentication followed a procedure of password verification with a trial limit of three input from the user. Failure to provide the correct password within that limit of attempting to access the system would result in system lock down and buzzing the alarm. After successful initalizing of Layer 1, the next layer of authentication using fingerprint scanner would be presented to the user for scanning his\her fingerprint. Upon the complete match with the prestored fingerprint would activate the servo motor attached to lock of the door of the secured vault to be opened as depicted in fig. 3. However, the vault again would move to locked state ending with the triggering off the alarm. Thus the Double Layer security would be ensured for the safe and legitimate use of the vault and safeguard the content or elements to be kept protected.

Apart from the mentioned causes of safety violation, a rather emergency case of infringement could still occur due to forceful break in. If a breach in to the system occurred in such a manner bypassing the layers of password and fingerprint affirmation, a motion detector sensor (PIR sensor) was placed within the perimeter of the vault which is only active while the procedure of the Layer 1 and 2 would be missing from the controller. Similarly, a Light detection Sensor (LDR sensor) was also stationed along side the PIR sensor to assist the controller to decide any unauthorized access existance of a person/object within the secured peripherry of the vault. While any of these sensors were delivering any sensed data to the controller, the locking procedure would behave as it used to the situation of safety enforcement of Layer 1 and Layer 2. This layer of guranting the vault safety was termed as the Violation Status layer and its working principle is shown in fig. 4.

Along with the conditions to switch on or off the buzzer to monitor the security of the vault, a LCD display was also attached adjacent to the door of the vault, in the right side and the numeric keypad as well as the fingerprint scanner was situated in the opposite side at the same elevation. On the other hand, three LEDs were placed at the top of the locking door which indicated the operational modes of the vault and are termed as Status Indicator LEDs (SIL) in this report which are detailed in table III.

TABLE III. MODES OF OPERATION

Modes	SIL	Operational Status
V-111	GREEN	Vault “SECURED”
V-A	BLUE	Vault “UNDER USE”
V-0	RED	Vault “SECURITY BREACH”

V-111 indicated that all the layers and violation status procedures are active and the vault was secured. In contray of that, V-0 indicated that some kind of violation of any of the security feature of the vault had occurred and would update the ooperational status. However, V-A would mean that the vault was being used by a user following both layers of access and it would remain so till the user had completed the requirements of accessing. The switch from V-A to V-111 occurred after the user came out of the vault by mechanically opening the door from the inside and securing it again till the SIL was updated as per the status of V-111. In any case of delayed update between V-111 and V-A after mechanical lock was initiated, the system would again move to V-0 and follow the operation of breach as per vault violation rules shown in fig. 4.

V. OBSERVATIONS FROM THE IMPLEMENTATION

The proposed model was implemented as a prototype for validation of the working criteria to ensure the safety of a vault using double layers of security. The model was built as per the block diagram outlined in fig. 1. The controller was programmed according to the algorithms presented in fig. 2, 3 and 4. The door of the vault operated as per the modes of operation already discussed in earlier sections, following table III. The vault security was tested by instigating forced breaking in as well as inputting false biometric and numeric authentication parameters. In all those scenarios, the buzzer of the system was set off and operational status got reconditioned through the SIL. For the proper and natural operation of the vault security, user verification tools of password protection and biometric sensing were investigated which assured that the prototype was functioning with all its attributes and protocols of double layer security being commissioned.

In fig. 5(a), the implemented prototype can be seen with all the peripheral components and the vault door being energized by the servo motor. The SIL are indicating the states of operation as per the categories mentioned in table III. Rest of the components, including the control unit as well as the motion and light detectors for the detection of security breach were placed within the premise of the vault of this prototype.

All the states of operation indicated by the SIL as well as the texts of LCD display were shown in fig. 5(b), fig. 5(c) and fig. 5(d) respectively.

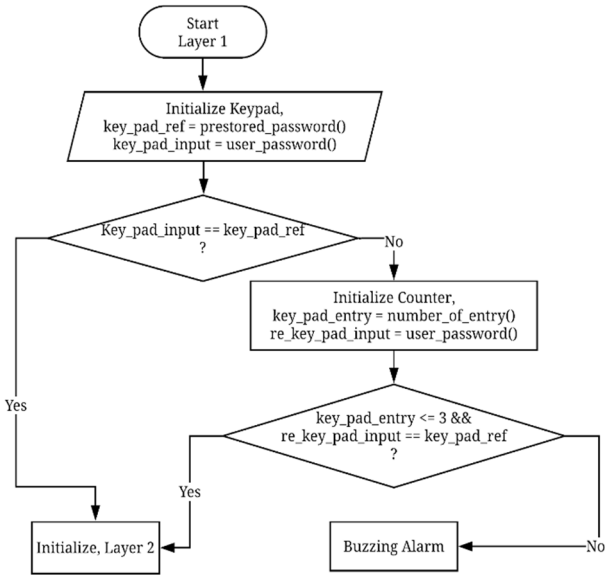


Fig. 2. Flowchart of primary layer (Layer 1) of vault security.

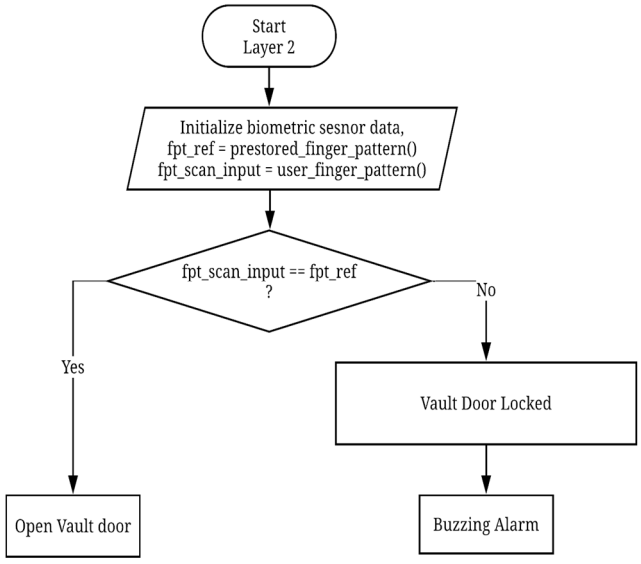


Fig. 3. Flowchart of secondary layer of vault security.

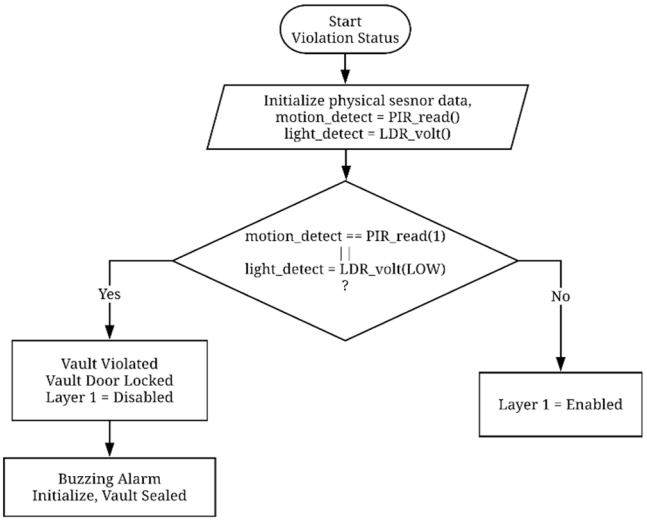


Fig. 4. Flowchart of continuous violation status verification.

The vault door can be noted to be in operation with SIL indicating the BLUE light being after successful entry of numeric password and verification of biometric data through fingerprint scanner as the system prepared itself to be in V-A mode of operation. A forced break-in was performed to test the V-0 state which was also properly working due to the adequate detection of motion and illumination by the PIR and LDR modules respectively. During the user's access of the vault, the door was kept open at an angle of 90^0. After the secured access, the user is required to lock the vault door using the keypad by pressing the "*" key. To initiate the access, the system needed to be enabled by pressing "#" key of the keypad.

In fig. 6, the control unit components can be found installed within the vault premise of the prototype including the PIR sensor and LDR module.

All the components were used and configured according to manufacturer's specification. It can also be added that the prototype was implemented in a very reasonable budget. The cost of implementation can be found in table IV which can be scaled for commercial manufacturing.

TABLE IV. COST OF IMPLEMENTATION

Components	Manufacturer	Cost (in BDT)	Cost (in USD)
Fingerprint Scanner	Shenzen CAMA	2100	25.07
LCD Display 16*2	Shenzen Shanhai	150	1.8
PIR	Ningbo Best Group	90	1.07
LDR Module	Fut-electronic Tech. Co.	120	1.43
Arduino Uno (2 units)	Sparkfun Industries & Adafruit	800	9.55
Keypad, Cables and accesories	Locally sourced	350	4.18
Total:		3610	43.10

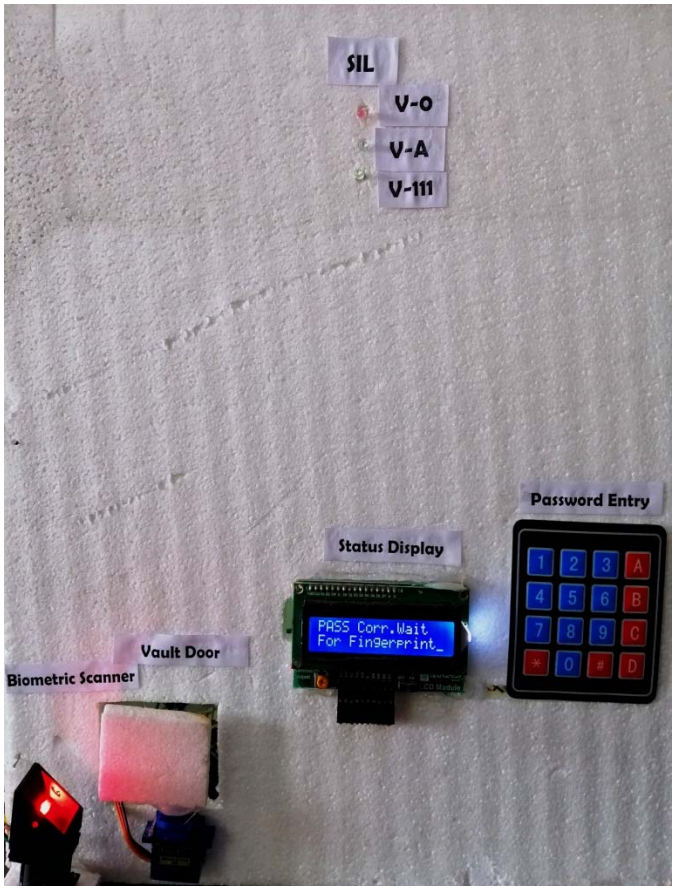


Fig. 5(a). Implementation of the prototype.

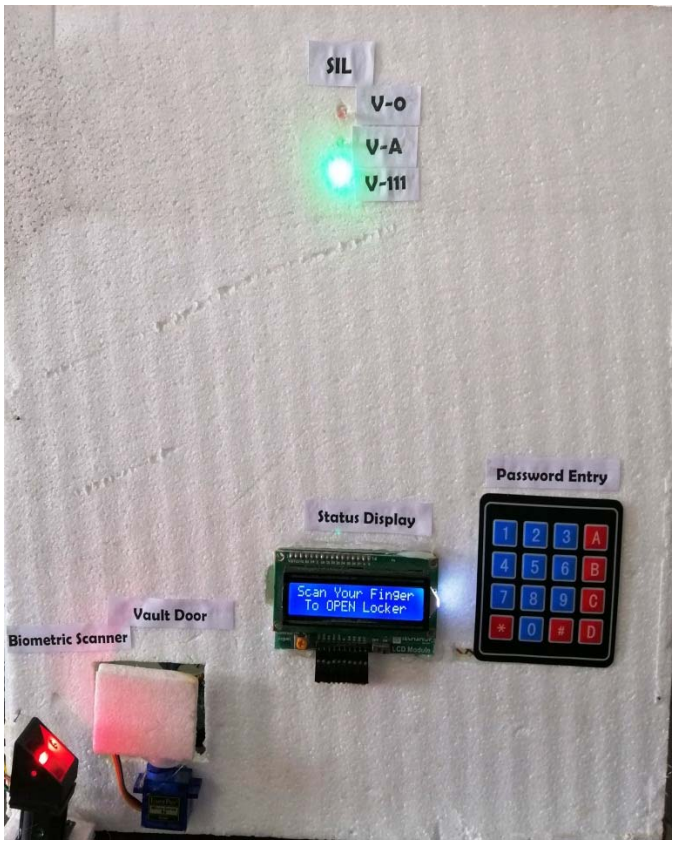


Fig. 5(b). The vault, in "SECURED" state

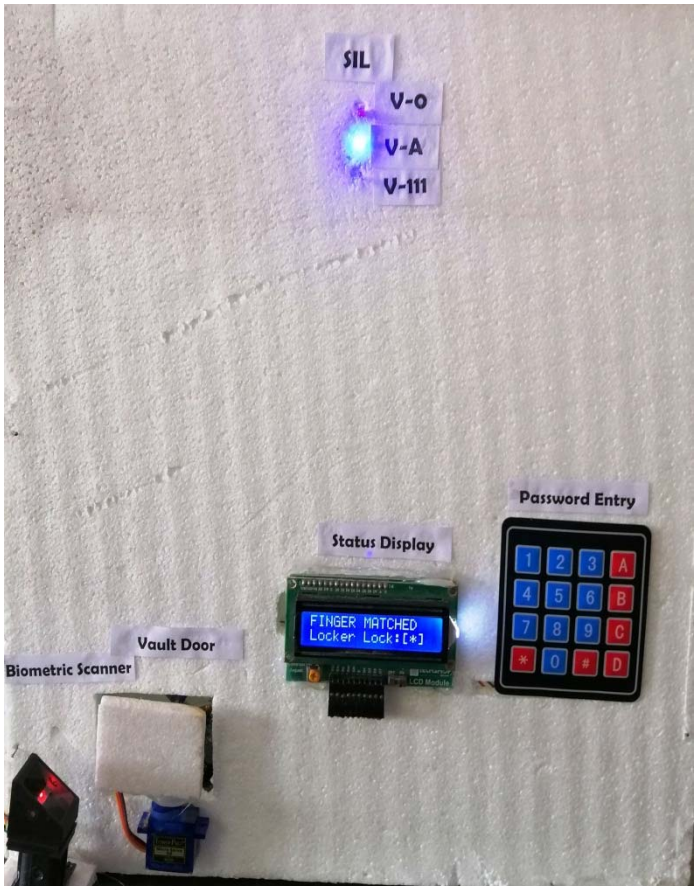


Fig. 5(c). The vault, in "UNDER USE" state

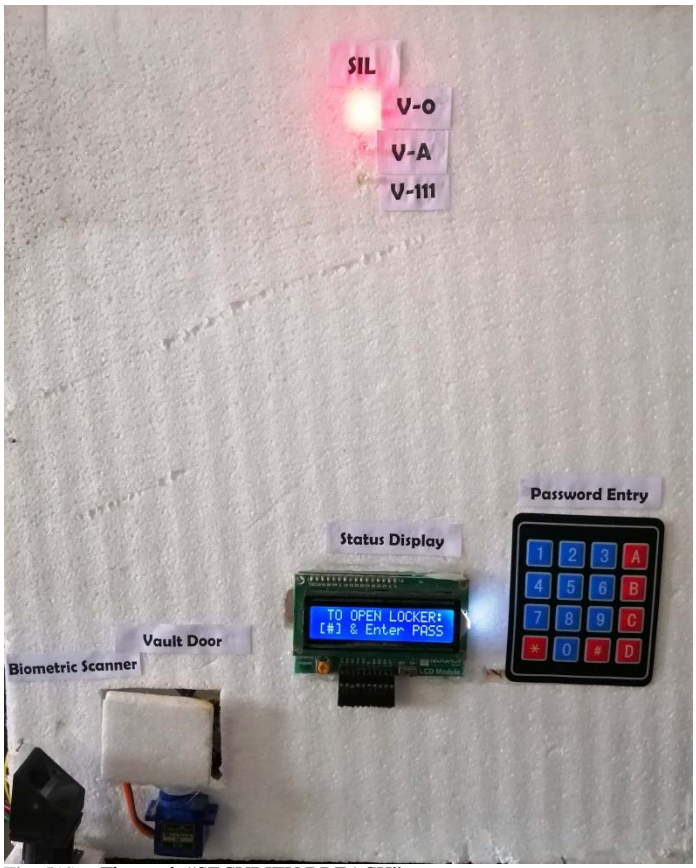


Fig. 5(d). The vault "SECURITY BREACH" state

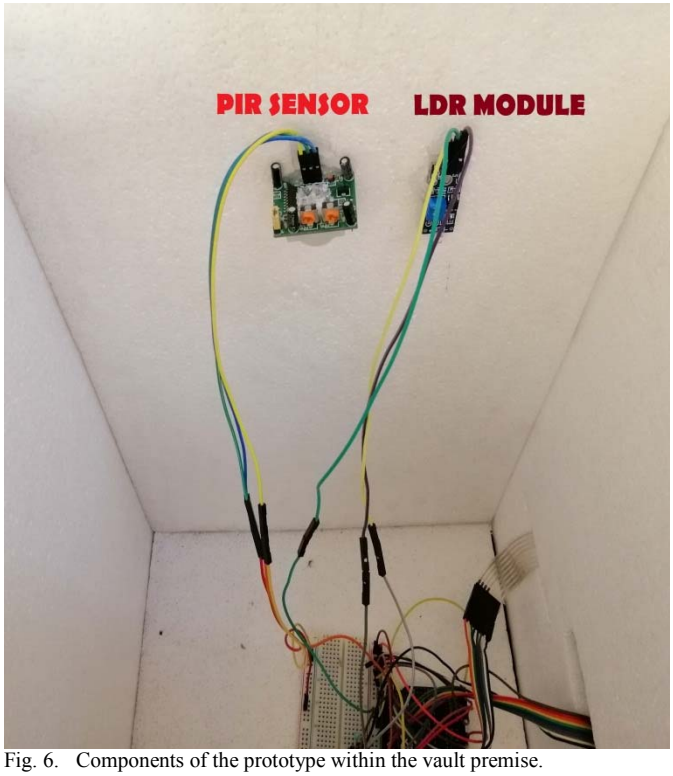


Fig. 6. Components of the prototype within the vault premise.

VI. Conclusion

Secured vaults are a basic requirement in various industrial scopes of financial, educational, management institutes and many more. In This paper, the design and implementation of a prototype of an automated vault door locking system is presented which warrants double layer of security. It ensures the proper user of the vault by securing the door with numeric password and biometric authentication. It monitors the conditions of operation of the vault from both the inside and the outside by employing several sensors which are continuously feeding information to the controller of the proposed system to confirm the robustness in terms of rightful access and security of the contents within the vault. The entire system can be easily managed with all the status updates being reeled by the controller to the administrators eliminating the unforced reasons of human errors. Implementation at a reasonable budget can also be mentioned as another beneficial aspect of the system. The observations of its operation are very encouraging to provide an easy solution to a sophisticated and unavoidable problem faced by various institutions which is quite user friendly. Furthermore, accessing the sensor data for remote monitoring, administering and handling the breach circumstances are being considered for future implementation using the vast dimensions added by IoT and cyber-physical system models. In that spectrum with enormous amount of capacity to process and store big data, an enhanced state-of-art vault door locking system is being researched and developed by the authors.

References

[1] Crystalynne D. Cortez, Jaswinder S. Badwal, Jocelyn R. Hipolito, Ditche Jane C. Astillero, Melvie S. Dela Cruz, and Jaira C. Inalao. “Development of Microcontroller-Based Biometric Locker System with Short Message Service” Lecture Notes on Software Engineering, Vol. 4, No. 2, May 2016

[2] A. Aditya Shankar, P.R.K.Sastry, A. L.Vishnu Ram, A.Vamsidhar, “Finger Print Based Door Locking System” International Journal Of Engineering And Computer Science ISSN: 2319-7242Volume 4 Issue 3 March 2015, Page No. 10810-10814.

[3] Sheikh Izzal Azid, Sushil Kumar ; “ Analysis and Performance of a Low Cost SMS Based Home Security System” International Journal of Smart Home Vol. 5, No. 3, July, 2011

[4] J. Tan, Q. Wang, C. Wu, Y. Zhou, J. Hou and Z. Yuan, "Research of Intelligent Home Security Surveillance System Based on ZigBee," 2008 International Symposium on Intelligent Information Technology Application Workshops(IITAW), vol. 00, no. , pp. 554-557, 2008.

[5] Mohd Adib B. Sarijari , Rozeha A. Rashid , Mohd Rozaini Abd Rahim , Nur Hija Mahalin “ Wireless Home Security and Automation System Utilizing ZigBee based Multi-hop Communication”

[6] Md. Kamal Hossain, Prodip Biswas, Mohammed Mynuddin, Sayidul Morsalin, “Design and Implementation of Smart Home Security System” International Journal of Modern Embedded System (IJMES) ISSN: 2320-9003(Online), Volume No.-2, Issue No.-6, December 2014.

[7] C. Malarvizhi, V. Kalaipoonguzhali and J. Anitha “Microcontroller Atmega 328P and GSM Based Advanced Home Security System” International Journal of Smart Home Vol. 11, No. 6 (2017), pp. 11-20

[8] Saurabh Kul, Amey Mundle, Vaibhav Jivrak, Prof. Mrs. Vidya Dhamdhere, “ Arduino Based Security System” International Engineering Research Journal (IERJ) Volume 2 Issue 2 Page 671-674, 2016, ISSN 2395-1621.

The Third Tab (3T): *Assistant for Visually Impaired People*

ATmega based automated blind cane

Sohal Islam
Dept. of EEE,
AUST
Dhaka, Bangladesh
sohal857@gmail.com

Muhammad Jakaria Rahimi
Assistant Professor, Dept. of EEE
AUST
Dhaka, Bangladesh
mjrahimi@gmail.com

Kazi Arham Kabir
Dept. of EEE,
AUST
Dhaka, Bangladesh
Arhamkabir0179@gmail.com

***Abstract*— Among many of human disabilities, blindness is a very common and unendurable one. Different techniques have been established to define the degree of vision defect and visual deficiency. To minimize the sufferings of visually impaired people various research works have been carried out in recent years. But still, there are a lot of aspects yet to be explored in this field. Our motive of the research work is also to lessen the sufferings by developing a cheap and easy to use the device to assist the visually impaired persons Hence we have developed a microcontroller based smart cane to assist the blind The smart cane can detect all obstacles about 180 degrees range around the direction of moving of a blind person. As soon as an obstacle is detected, it generates a voice message to warn the blind instantaneously. The device re-charging facility is the most interesting feature of our research work. An arrangement of raw magnets produces mechanical rotation which in turn produces sufficient backup power from a small DC motor. In addition to this, we have used GSM module with a panic button for getting the current location of crisis situations.**

Keywords—Cane, ATmega-2650, Ultrasonic Sensor, Speaker, GSM, Moisture Sensor, Magnets, Generator, Rechargeable Battery.

I. INTRODUCTION

Blindness or visual impairment is a major issue that has been affecting many people around the world. According to the World Health Organization (WHO), an estimated 285 million people live with vision Impairment: 39 million are blind and 246 million have moderate to severe vision impairment [1]. People who suffer from a severe reduction in vision that can't be corrected with conventional means, such as refractive correction or medication and reduces a person's ability to function at certain or all tasks. Hugh Herr said,

"Humans are not disabled. A person can never be broken. Our built environment, our technologies are broken and disabled. We the people need not accept our limitations, but can transcend disability through technological innovation."

There is a large number of problems associated with simple canes or guide dogs [2]. A simple cane offers a limited preview range in which the user has to be very careful and saunter. Use of attendant dogs is also not appropriate for all the visually impaired people because they require training and coordination with dogs and also it is not reliable. But nowadays we can see, there are so many techniques, and devices have been developed, which help these persons to move freely even in a dynamic environment [3][4]. Still, these researches existed some restrictions, as some of them can't be utilized for indoor space [5]-[10]. The cane consists of a circuit board containing an Atmega-2650, inputs from sensors, speakers for voice commands, GSM, magnets, and generator as a source of power generation and a rechargeable battery. It can only be operated by the possessors as the Microcontroller is code protected. It also can serve blind people to detect obstacles, road holes by giving voice commands.

After surveying the walking speeds for visually impaired people [11][12], Clark-Carter, et al. presents the average walking speeds for visually impaired is 1.61 m/sec [11] and Nakamura shows that the walking speeds for late blind and congenitally blind are 1.11 and 0.86 m/sec, respectively [12].

In our design, we have incorporated some new features that were not been used previously. It is equipped with battery recharging capability. It can detect obstacles in a wider range. It also provides sufficient information about the obstacle height to avoid any collision. Through its panic button feature, it can promptly send its current location information to another person to get external help quickly whenever necessary. While ensuring all these essential features the cost is not increased much and it is still very user-friendly.

The paper is organized with the following sections. In section II we have observed previous works on this type of projects. The proposed method is presented in details in section III. Our measured output accuracy is presented in section IV under experimental accuracy subheading. And finally, in the Conclusion part we have summarized our works and purposes.

II. RELATED WORKS

After World War II, people started to develop electronics devices to help disabled peoples. The 1st step towards smart stick was detecting obstacles in front of visually impaired people. Nitish Sukhija proposed A Smart Stick for Blind Man, which contains Ultrasonic Sensors, IR Sensors, and Atmega8 microcontroller unit. The model follows a specified path [13]. Mithiles Kumar proposed a system which was A Low-Cost Smart Stick for Blind and Partially Sighted People. They used ultrasonic sensors, infrared sensors, water sensor, fire sensor and light (LDR) for sensing the atmosphere around a blind person, and it gives output in the form of speech instructor, vibrator, and buzzer, so both auditory and tactile feedback is provided [14]. It has a dual feedback system and easy to use. Ashraf Anwar proposed A Smart Stick for assisting visually disabled people which consists of ultrasonic, IR, and light (LDR) sensor, a microcontroller (Arduino Uno R3), buzzers and vibrators [15]. It has Voice based navigation system with fire and light sensor which assist visually impaired people more effectively. L. Dong proposed the system based on the ultrasonic sensor distance measurement method which was called DBG Crutch Based MSensors [16]. The drive of this scheme is to help blind persons in spotting and ducking the obstacles in front, left and right of the person. The system uses a package of three ultrasonic sensors. The whole system is controlled by the STC15F2K60S2 microcontroller. Other systems used light instead of ultrasounds to detect obstacles.

Some used laser, such as the C5 laser cane [17], and, later, the Teletext [18]. Other used unfocused light, usually in the infrared range; this was the case of the Tom Pouce [18]. Devices of this type generally rely on a single sensor to provide one-dimensional range information: the user orients the device in the direction she or he wants to scan and gets a distance measurement in the result, usually in the form of vibrations sound frequency or haptic feedback. The problems associated with these devices are the necessary training requires to use them [18], their cognitive load and their limited robustness. Radio Frequency Identification Walking Stick (RFIWS) was designed in [19] to help blind persons circumnavigating on their path. This system aids in sensing and calculating the estimated distance between the path border and the blind being. RFID is used to handover and accept signal through radio wave medium. An idea of a smart wheelchair named 'HARUNOBU-6' is also a blind navigation system developed by Mori et al. [20]. Here differential GPS mechanism is used as its basic principle.

The above-discussed sticks don't give a solution to pothole and stairs detection, and also if a device runs out of battery, he won't be able to find out his way without recharging. So in this paper, we are proposing the solutions for these problems along with water detection, ultrasonic sensors, magnets & generator.

III. PROPOSED METHOD

The block diagram of the proposed method shown in **Fig. 1** consist of ATmega-2650 Microcontroller, Seven organized Ultrasonic Sensors (Six for obstacle detection and one for hole detection), Moisture detector (for detecting mud and water), Amplifier (for magnifying voice output), Headphone (as a hearing device), Vibrator (to notify through vibration), Rechargeable Battery (to power up the device), Magnets And Generator (to recharge the battery if needed), GSM And Panic Button (to notify the lover ones in crisis).

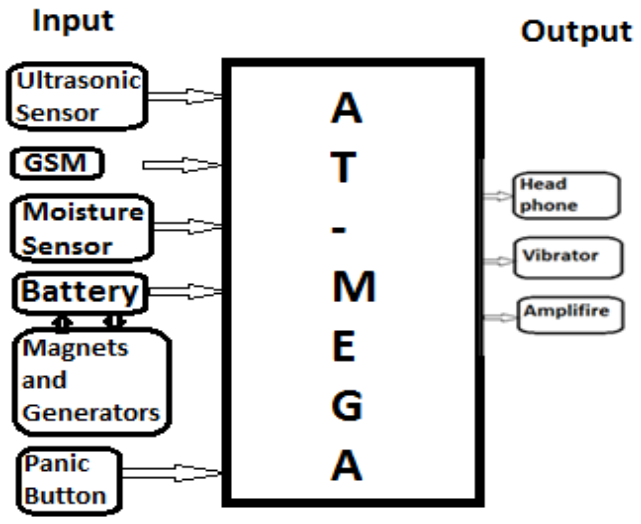


Fig. 1: Block Diagram of Proposed Cane

Here Ultrasonic Sensors, Moisture Detector, GSM and the panic button will work as the input of Microcontroller. Then the Microcontroller will process data as given instruction and will give output through Amplifier, Vibrator, and Headphone. Here Magnets and Generator will work as the device to power up the battery. After the construction of **Fig. 1,** we get a smart cane which looks like **Fig. 2**

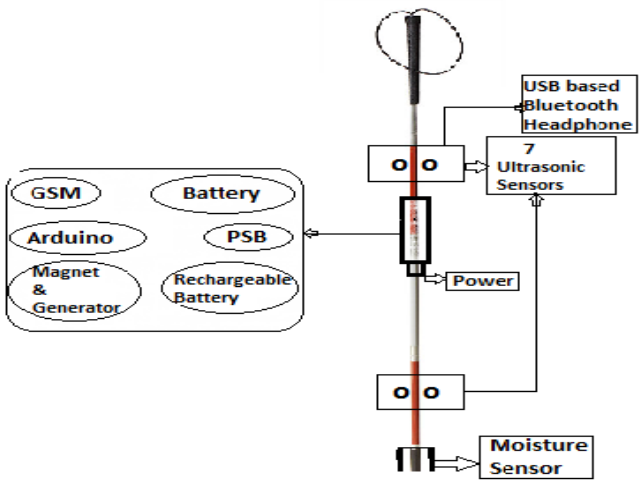


Fig. 2: Proposed Smart cane.

Obstacle Detection Procedure:

The input of the cane consists of seven ultrasonic sensors which are capable of detecting obstacles in front (about 180 degrees) of it at a range of up to 30 cm (Maximum-Efficiency). The inputs are interfaced to the Microcontroller, which determines if an obstacle is too close to the cane and hammers the output. The output consists of a vibration which provides a response and amplifier which is attached to the headphone for audio feedback.

In this process, there are too many assumptions firstly above knee level obstacle detection system in which three ultrasonic sensors (Right Upper, Middle-Upper, and Left Upper) are placed on the cane at about 3 ft. from the ground. The other one is ankle level obstacle detection system in which three ultrasonic sensors (Middle Lower, Right Lower and Left Lower) is placed on the stick at about 0.4 ft. above the ground level. The following **Table-1(a, b, c & d)** shows the variation of measurement for different assumptions.

Table-1(a): Measurement for multiple assumptions

Sensor / Case	Right upper (cm)	Left upper (cm)	Middle upper (cm)	Right lower (cm)	Left lower (cm)	Middle lower (cm)	lower (cm)
Hole	>30	>30	>30	>30	>30	>30	>30
Stair	>30	>30	>30	>30	>30	<30	<30
Front (knee level)	>30	>30	<30	>30	>30	>30	<30
Front (Ankle level)	>30	>30	>30	>30	>30	<30	<30

Table-1(b): Measurement for multiple assumptions

Sensor / Case	Right upper (cm)	Left upper (cm)	Middle upper (cm)	Right lower (cm)	Left lower (cm)	Middle lower (cm)	Lower (cm)
Above 3ft. heighted object (Front)	>30	>30	<30	>30	>30	<30	<30
Above 3ft. heighted object (Right)	<30	>30	>30	<30	>30	>30	<30
Above 3ft. heighted object (Left)	>30	<30	>30	>30	<30	>30	<30

Table-1(c): Measurement for multiple assumptions

Sensor / Case	Right upper (cm)	Left upper (cm)	Middle upper (cm)	Right lower (cm)	Left lower (cm)	Middle lower (cm)	Lower (cm)
Right (knee level)	<30	>30	>30	>30	>30	>30	<30
Right (Ankle level)	>30	>30	>30	<30	>30	>30	<30
Left (knee level)	>30	<30	>30	>30	>30	>30	<30
Left (Ankle level)	>30	>30	>30	>30	<30	>30	<30

Table-1(d): Measurement for multiple assumptions

Sensor / Case	Right upper (cm)	Left upper (cm)	Middle upper (cm)	Right lower (cm)	Left lower (cm)	Middle lower (cm)	Lower (cm)
Hole & Object (Front)	>30	>30	<30	>30	>30	<30	>30
Hole & Object (Right)	<30	>30	>30	<30	>30	>30	>30
Hole & Object (Left)	>30	<30	>30	>30	<30	>30	>30

Pot-hole detection Procedure:

In this process, we planned to detect the holes. For that, an ultrasonic sensor is placed on the bottom of the cane and at a right angle using extra support which will look like **Fig. 3.** If the sensor senses the value more than 30 cm in distance, it will provide audio and tactile feedback through headphone and vibrator.

Table-2: Practical VS Theoretical value.

Theoretical Value	Practical Value
57 cm	55.6 cm
Our accuracy of measurement is **97.5%**	

We have measured both practical and theoretical values. The theoretical calculation is shown in **Measurement and Fig. 4.** Comparing these values we got some slight deflection between them.

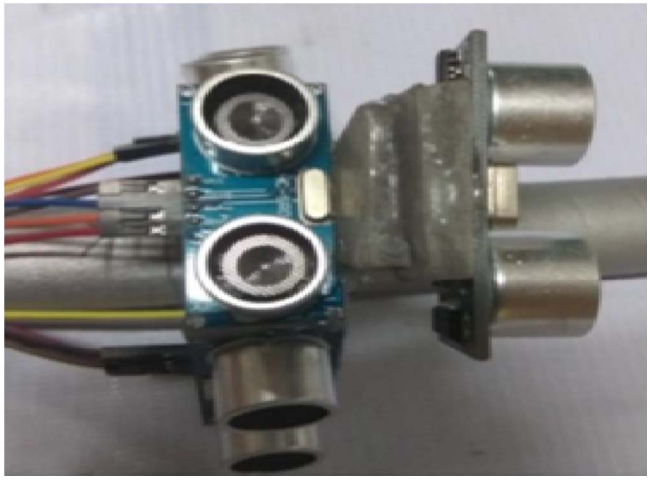


Fig. 3: Holes Detection

Measurement:

We can measure the holes from safe distance. Theoretical measurement:

$$\sin 30 = x/105 \qquad \cos 30 = z/5$$
$$\Rightarrow x = 52.5 \text{ cm.} \qquad \Rightarrow z \simeq 4.5 \text{ cm}$$
$$\text{So, } S = (x+z) = 57\text{cm.}$$

We got the above calculation from **Fig. 4**

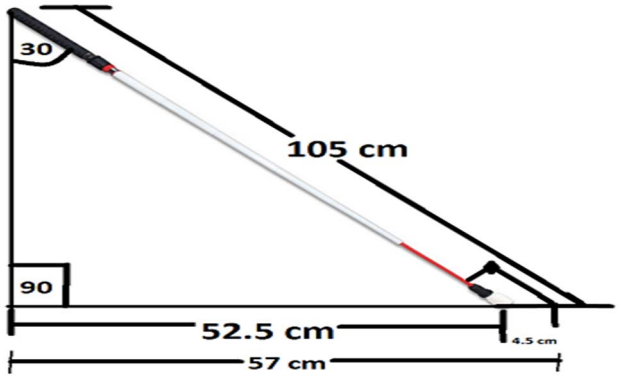


Fig. 4: Measurement of holes detection.

Moisture Detection:

We've placed the moisture detector at the bottom of the cane to sense water or mud. The sensing range of the detector is almost 3.7 cm. If the sensor detects water level or mud it alerts the user providing audio feedback at the same time the moisture sensor can declare the property of the soil such as electrical resistance, dielectric const. or interaction with neutrons as a proxy for the moisture content. It can also detect environmental factors like soil type temp. (10 -30 degree Celsius) or electric conductivity. **Fig. 5** shows a demo of moisture detection.

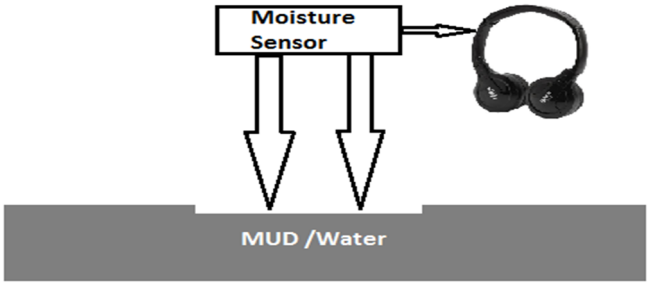


Fig. 5: Demo of Moisture Detection

GSM System and Panic Button:

Actually, we're using a function named Panic Button which is directly connected with GSM module. If anyone pushes the panic button in crisis GSM Module will turn on and send the location of the person to his/her loved ones. Information comes through a massage which is shown in **Fig. 6.**

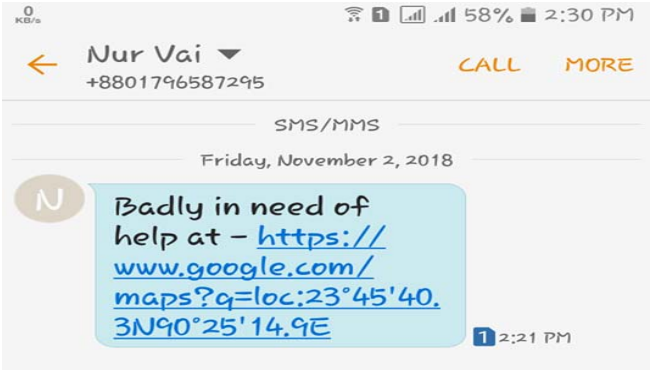


Fig. 6: Massage from the device when needed.

In the above figure, we can see the message of a Visually Impaired person with his/her exact location. The message was sent from the cane because of activating the Panic Button.

Backup Power Generation:

We're using a rechargeable battery which gives us 5v output and it is directly connected to the **Vin** of ATmega-2650. As we know the battery isn't sufficient to provide long term service (10-12 hours). So, we've thought about a backup plan to recharge our power source. We've used the magnetic field and the attraction and repulsion between two raw magnets. For this magnetic characteristics, we are able to rotate our generator which will create Mechanical Energy. It will be automatically converted into Electrical Energy.

When the battery of the cane is lower than 50%, User of the cane will be notified through the vibrator. Then the user has to press a button which will turn on the reproduction system. After this process, we got a measurement of voltage production which varies with the speed (rpm) and size of the generator and magnets respectively (such as 3.78, 4.9, 5.53, 6.3, etc). We've converted the gained voltage into a constant 5V by taking help from POLOLU S7V7FS module [21]. The module schematic is taken from POLOLU website [22] and shown in **Fig. 7(a).**

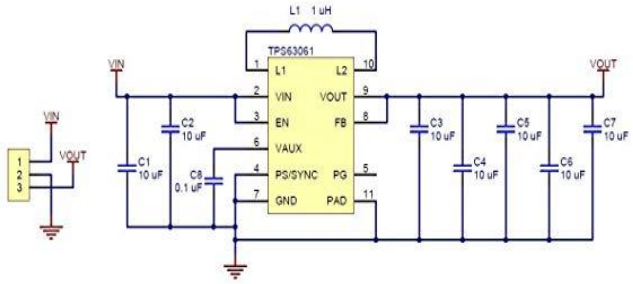


Fig. 7(a): Step up/down voltage regulator.

We've compared two types of a motor having different revolution per minute (100 & 300 rpm) and got different initial output voltages. Table (3&4) shows the measurement of the recharging process.

Table-3: Measurement of the recharging process.

Magnet Level	Initial Output (100rpm)	Initial Output (300rpm)	Voltage Regulator.	Final Output. (Voltage)
LOW	**4.9 V**	**3.6 V**	**S7V7F5**	**5**
MEDIUM	**6.8 V**	**5.2 V**	**S7V7F5**	**5**
HIGH	**9.2 V**	**6.8 V**	**S7V7F5**	**5**

Table-4: Measurement of the recharging process.

Magnet Level.	Output Current (100 rpm)	Output Current (300 rpm)
LOW	**3.60 mA**	**2.10 mA**
MEDIUM	**4.20 mA**	**2.50 mA**
HIGH	**5.66 mA**	**3.20 mA**

The whole tiny system looks like Fig. 7(b) in real life which becomes the ultimate back up [23] function for our power source.

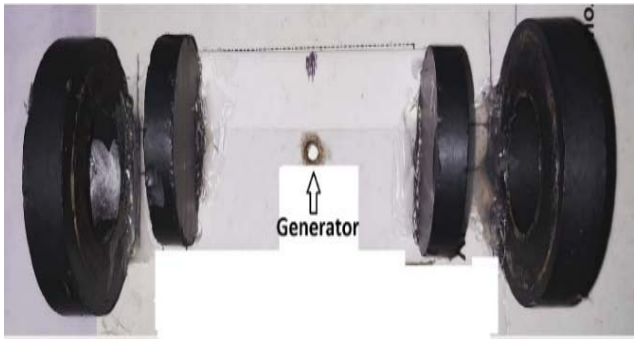


Fig. 7(b): System to recharge the Power Source.

- GENERAL DESCRIPTION OF SENSORS

Ultrasonic Sensor:

An Ultrasonic sensor is a device that can measure the distance by using sound waves at a frequency of 40 Hz. We have used HC-SR04 in our cane which helps us to determine the possible distance between the stick and any objects. The distance that one can possibly measure with the non-contact object is about from 2cm (0.8 inches) to 3 meters (approx. 118 inches). The module consists of the control circuit, ultrasonic receiver, and transmitters. It has four pins VCC, GND, Trigger, and Echo. The basic working principle [24] is: (1) It uses IO trigger for at least ten microseconds high signals; (2) The Ultrasonic sensor module repeatedly sends eight 40 kHz wave and sense if there is a signal comes back; (3) If the signal comes back in high level, time of high output IO duration is the time from sending a wave to receiving it back.

Distance = (high level time * speed of sound (344m/s))/2 [25].

Table-5: Working range of HC-SR04 [26].

Working Voltage	DC 5V
Working Current	15 mA
Working Frequency	40 Hz
Max Range	4 m
Min Range	2 cm
Measuring Angle	15 degree
Trigger Input Signal	10 us TTL pulse
Echo Output Signal	Input TTL lever signal & the range in proportion.
Dimension	45*20*15 mm

Moisture Detection:

Moisture detector is one kind of a sensor which determines the ratio of hydrogen and oxygen and measure the moisture of the surface if there is any water on the ground or not. This concerns visually impaired people to detect water and mud. It gives visually impaired people the confidence to move freely.

Global System for Mobile communication:

Global System for Mobile communication (GSM) is a second-generation cellular network usually used by mobile devices or digital devices which includes microcontrollers. GSM (SIM900A kit) have used for the communication purpose. The onboard resources in it are Serial port circuit (with protection) Antenna interface circuit (SMA bend female port) SIM card circuit (flip SIM slot) 4*3.5 fixture hole 4pcs SIM900A serial port output terminal. It has improved spectrum efficiency, international roaming, compatibility with ISDN (Integrated Services Digital

Network), SIM phone book management and short message service (SMS).

IV. EXPERIMENTAL ACCURACY

We've compared the theoretical and practical values of random distances using the Ultrasonic Sensors that were used in the cane. We saw that the values were too close but not 100% accurate. We made a graph after taking several values from random distances which are shown in **Fig. 8**. From the graph, we can see there are some floating errors (%) in between one to three percent. So,

Accuracy = (97 to 99) %.

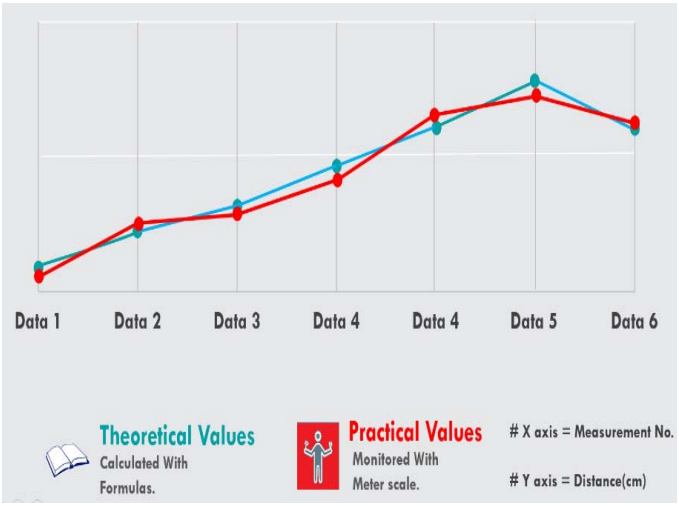


Fig. 8: Comparison between Theoretical and Practical Values

V. Conclusion

Our proposed cane removes most of the hurdles during movement of visually impaired people all over the world. Especially the blind people in the underdeveloped and developing countries will be benefited most as the cane is very cost effective and easy to use. In this system, we've used the most available components to reduce cost and complexity. People can use or repair it whenever needed because of the availability of components. The cane is equipped with a rechargeable feature which ensures its power efficiency. The cane is designed to achieve maximum durability with a compact & power efficient structure. Maintenance and repairing procedure is also kept simple. The panic button feature not only relieves the tension of the blind person but also his relatives by promptly ensuring his location in case of any undesired crisis moment. In the future, this work can be upgraded into a more compact, lightweight and user-friendly cane. This will surely contribute a lot to make life easy for visually impaired people.

References

[1] B. Sveny, G. Kovcs, and Z. T. Kardkovcs, "Blind guide - a virtual eye for guiding indoor and outdoor movement," in *5th IEEE Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 343–347, Nov 2014.

[2] I. Ulrich and J. Borenstein, "The GuideCane-applying mobile robot technologies to assist the visually impaired," in IEEE Transactions onSystems, Man, and Cybernetics - Part A: Systems and Humans, vol. 31, no. 2, pp. 131-136, Mar 2001.

[3] L. Tian, Y. Tian, and C. Yi, "Detecting good quality frames in videos captured by a wearable camera for blind navigation," in *IEEE International Conference on Bioinformatics and Biomedicine*, pp. 334–337, Dec 2013.

[4] Ayat A. Nada, Mahmoud A. Fakhr and Ahmed F. Seddik "Assistive Infrared Sensor Based Smart Stick for Blind People", Science and Information Conference (SAI), 2015, 10.1109/SAI.2015.7237289.

[5] Y. Hirahara, Y. Sakurai, Y. Siizu, K. Yanashima, K. Magatani, "Development of the Navigation System for the Visually Impaired by Using White Cane," in *2006 Proc. the 28th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society,* pp. 4893-4896.

[6] D. Jain, A. Jain, R. Paul, A. Komarika, B. Kalarishnan "A path-guided audio based indoor navigation system for persons with visual impairment," in *2013 Proc. Int. ACM SIGACCESS Conf. on Computers and Accessibility,* Article No. 33.

[7] Y. Shiizu, Y. Hirahara, K. Yanashima, K. Magatani, "The development of a white cane which navigates the visually impaired," in *2007 Proc. Annual Int. Conf. IEEE Engineering in Medicine and Biology Society,* pp. 5005-5008 .

[8] T. Seto, K. Magatani, "A Navigation System for the Visually Impaired Using Colored Navigation Lines and RFID Tags," in *2009 Proc. Annual Int. Conf. IEEE Engineering in Medicine and Biology Society,* pp. 831-834.

[9] A. J. Fukasawa, K. Magatani, "A navigation system for the visually impaired an intelligent white cane," in *2012 Proc. Annual Int. Conf. IEEE Engineering in Medicine and Biology Society,* pp. 4760-4763.

[10] T. H. Riehle, S. M. Anderson, P. A. Lichter, N. A. Giudice, S. I. Sheikh, R. J. Knuesel, D. T. Kollmann, D. S. Hedin, "Indoor Magnetic Navigation for the Blind," in *2013 Proc. Annual Int. Conf. IEEE Engineering in Medicine and Biology Society,* pp. 5187-5190.

[11] D. D. Clark-Carter, A. D. Heyes, and C. I. Howarth, "The efficiency and walking speed of visually impaired people," *Ergonomics*, vol. 29, no. 6, 1986, pp. 779-789.

[12] T. Nakamura, "Quantitative analysis of gait in the visually impaired," *Disability and Rehabilitation*, vol. 19, no. 5, 1997, pp. 194-197.

[13] N. Sukhija, S. Taksali, M. Jain, and R. Kumawat, "Smart stick for blind man," *International Journal of Electronic and Electrical Engineering*, vol. 7, no. 6, pp. 631–638, 2014.

[14] M. Kumar, F. Kabir, and S. Roy, "Low cost smart stick for blind and partially sighted people," *International Journal of Advanced Engineering and Management*, vol. 2, no. 3, pp. 65–68, 2017.

[15] S. A. Ashraf Anwar, "A smart stick for assisting blind people," *Journal of Computer Engineering*, vol. 19, no. 3, pp. 86–90, 2017.

[16] Y. Yi and L. Dong, "A design of blind-guide crutch based on multisensors," in *12th IEEE International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pp. 2288–2292, 2015.

[17] J. M. Benjamkn, N,A. and A.F. Schepis " A Laser Cane for the blind", Proc. Of the San Diego Biomedical Symposium, 1973.

[18] R. Farcy, R. Leroux, A. Jucha, R. Damaschini, C. Gregoire, A. Zogaghi, "Electronic travel aids anda electronic orientation aids for blind people: technical, rehabilitation and everyday life points of view", Conference & Workshop on assistive technologies for people with vision & hearing impairments, Kufstein, Austria, 17-21 July 2006.

[19] M. F. Saaid, I. Ismail, and M. Z. H. Noor, "Radio frequency identification walking stick (rfiws): A device for the blind," in *5th IEEE International Colloquium onSignal Processing & Its Applications,*, pp. 250–253, 2009.

[20] Mori H., and Kotani S., "Robotic Travel Aid for the Blind: HARUNOBU-6," European Conference on Disability, Virtual Reality, and Assistive Technology, Sovde, Sweden, 1998.

[21] https://www.pololu.com/product/2119 accessed at 5 September 2018.

[22] https://www.pololu.com/picture accessed at 5 September 2018.

[23] https://www.youtube.com/watch?v=OQXcfxh3w6g&t=30s accessed at 11 December 2018.

[24] K. S. K. Lokesh.A, Manjunath.T, "Electronic stick along with android smartphone's to the aid of blindly disabled individuals," *International Journal of Recent Trends in Engineering & Research*, vol. 2, no. 5, 2016.

[25] H.Sharma, M. Tripathi, A. Kumar, M. S. Gaur, "Embedded Assistive Stick for Visually Impaired Persons," in 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), July 10-12, 2018, IISC, Bengaluru, Bengaluru, India.

[26] https://cdn.sparkfun.com/datasheets/Sensors/Proximity/HCSR04.pdf accessed at 15 December 2018.

A Case Study on a Proposed Adaptive and Energy Efficient Street Lighting System for Chittagong City

Saif Hannan[1]
Electronic and Telecommunication Engineering
(Assistant Professor)
International Islamic University Chittagong
Chittagong, Bangladesh
saifhannan@yahoo.com

Goutom Barua Milton[2]
Electronic and Telecommunication Engineering
(Student)
International Islamic University Chittagong
Chittagong, Bangladesh
milton.iiuc@ymail.com

Md. Humayun Kabir[3]
Electronic and Telecommunication Engineering
(Student)
International Islamic University Chittagong
Chittagong, Bangladesh
mdhkrrabby@gmail.com

Mohammed Jashim Uddin[4]
Electronic and Telecommunication Engineering
(Assistant Professor)
International Islamic University Chittagong
Chittagong, Bangladesh
jashimcuet@yahoo.com

***Abstract*— Light emitting diodes are the commonly used street lighting solution nowadays. As saving energy in all aspects is the primary concern in the 21st century, researches are going on to make LED lights much energy efficient like, to make it adaptive with traffic flow and daylight intensity and some other factors. We have proposed a model using some sensors to design an adaptive and efficient lighting system with LEDs. Moreover, we have collected some real-time data for analyzing the traffic rate in three dominant type of roads in Chittagong city and found dimming time slots for the lights. Finally, we have calculated the electricity consumption cost with our model considering 10% consumption of power at dimming hours and 100% consumption at the rest of the times; which saves 88.26% of the lighting cost in comparison with conventional sodium lamps per year.**

Keywords—Adaptive Lighting System, Dimming Circuit, Luminous Intensity, Light Emitting Diode

I. INTRODUCTION

The street lighting system is one of the essential parts of the basic structure of a city. A city may have a different type of streets depending upon some basic rules like the number of lanes depending on vehicle speed and class, ease of entrance and exit, connection with other roads, flyovers, overpass, underpass etc. There are three types of routes: major arterial, minor arterial and collector roads. All the roads must have at least two street lights at the same place focusing on two opposite lanes placed on the road divider. For proper lighting at night, the light posts are usually set at 25/40/50 meters intervals depending upon the type of road. Sodium lamps (100 watts) are the most commonly used lights in a city. But it consumes much electricity for lighting the streets. Light emitting diodes (LED) are being used from the past few years instead of sodium lamps, halogen lamps, fluorescent lamps, light bulbs etc., to reduce the electricity consumption. An LED light may only consume 30 watts, which significantly reduces the cost by 70%. Moreover, it has a longer lifetime and low maintenance cost and appropriate lumens of brightness.

As the production of electricity is costly and saving from system loss of power is a crying issue; we have proposed a system of adaptive LED lights, which will switch on or off the lights depending on daylight intensity as well as will adjust brightness of light depending upon detection of vehicles or pedestrian passing by the light posts. If the system detects a moving/stopped object like a car or a human by a sensor placed for both side of the road, the light above the road will flash with 100% brightness; otherwise it will be dimming with only 10% brightness using less amount of power and hence will save a significant amount of energy.

II. LITERATURE REVIEW

Light emitting diodes (LED) are efficient, long-lasting and proper luminary for 21-century street lighting. In comparison with other lighting systems like HPSV, LED street lights were proved much efficient for luminosity regarding specific standards [1], [2]. Although LEDs can save significant power, research is still going on to make it adaptive with environmental situation and needs, for saving much more energy. A situation-based traffic adaptive street light control was proposed by T. Novak and K. Poll hammer [3], which used two different traffic adaptive dimming strategies and could save power from 35% to 50%. Some other techniques were applied to reduce power consumption based on astronomical clock and microcontroller [4] and local controllers, motion sensors, video-camera and electronic devices for video processing etc. for counting traffic [5] and controlling light intensity, with saving 65% power. Such techniques required employing some electronic and electrical devices which may cost much money. To make the system more reliant and less costly, some dimming methods were proposed over the years. The breakthrough in dimming traditional lighting systems was first introduced by G. Velasco-Quesada et al. [6], where multiple-tapped autotransformer and high-frequency switch converter dimmer system technologies were used. But for LEDs, the intensity controller system was proposed [7] based on the detection of humans and vehicles in the vicinity to the sensors used for the system. Pulse width modulation (PWM) was also employed at 25 KHz frequency for intensity control [8]. Further studies were done to control LED streetlights like multi-sensing [9],

the smart module (SM) and control module (CM) [10], different sensors with IoT [11] etc.

All the works described above, and any other papers didn't show how much money they can save, quantitatively. In this paper, we proposed an adaptive street lighting system based on ultrasonic sensors and LDRs for each LED light pole. The sensors will switch on or off the lights and control the intensity of the lights depending on the detection of moving objects near the sensors. We also have studied how vehicular movement in different types of roads can be considered for the calculation of power saving.

III. Methodology

Firstly, we have designed a model for automatic and adaptive street lighting system as shown in figure 3. We have shown that the system works perfectly with dimming conditions.
Secondly, we have arranged some specific data of vehicular movement during lighting hours (i.e., at night) in three different type of streets in Chittagong city area, in Bangladesh. We have collected data for consecutive seven days and taken mean value for a day. Then We have used those data to calculate dimming time for the proposed LED light.

A. System Design

The development of the system starts with the system block diagram as shown in figure 1.

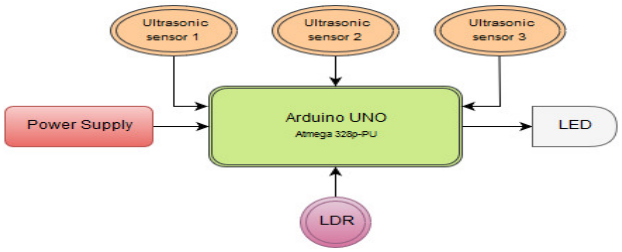


Fig. 1. Block Diagram of the system

We have declared three principal components as input for the model; these are Ultrasonic sensors, input power and LDR (light dependent resistor). Another one element that is proclaimed as the output of this system is LED. All the processes, like switching on/off the street lights (LEDs) and dimming LEDs in the absence of vehicles and humans are managed and controlled by the Arduino.

B. Hardware Specifications

Two main components are used as the input for the proposed method. LDR has been used to track out the daylight level. Its responses to daylight or any changes are measured in volt. Ultrasonic sensors have been used for tracking the vehicles and humans. When any car or human pass by the sensor range, LED will give 100% brightness.

Arduino UNO Atmega 328p-PU has been selected [12] and used for controlling the system. LED was connected to the Arduino and its intensity has been controlled by intensity module. Two levels of brightness were loaded as the command in the program for Arduino: for power saving, 10% power for lighting up in case of no traffic or passers-by and 100% power when vehicles and/or humans are passing by the road/lane.

The LDR starts the process of the proposed system. The process begins when the LDR detects no light and then LED lamps will be switched on. This system is programmed with two distinct levels of lighting. It will be switched off when the sunlight is sufficiently bright and switched on at night time when there is no sunlight. But when there is no traffic and less pedestrian, LED lamps will be dimming with 10% brightness only and will be lightened with 100% brightness when any vehicle or human pass by the lane and sensed by the sensors. The whole operation will be as below.

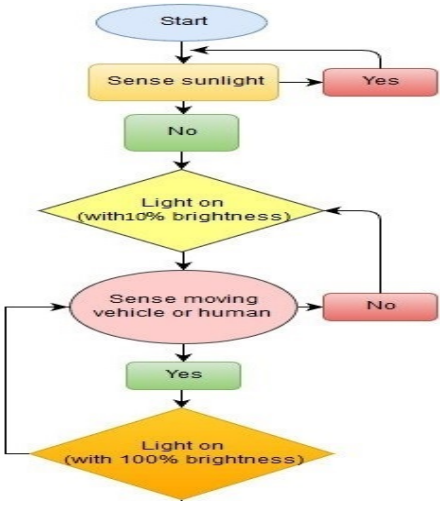


Fig. 2. The flowchart of the operation of the proposed design

IV. Implementation of The System

We have designed a prototype of the system with real-time experimentation and found it useful for our purpose, as depicted in figure 3.

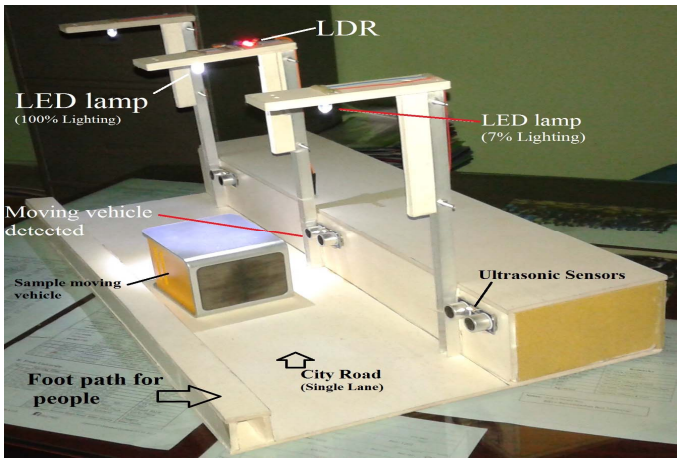


Fig. 3. Experimental setup of the system

It is clear from the figure that, LED lamps will only light up for 100% brightness when the system detects a moving vehicle or a human. Otherwise, it will just light up by 10% brightness during the absence of the car on the road and human on the footpath.

V. Data Analysis

We have collected the rate of moving vehicles in three different types of roads like Major Arterial (3 lanes on both sides), Minor Arterial (2 lanes on both sides) and Collector Roads (2 or 1 narrow lane) during dark hours (from dusk to dawn) of the day. Then we categorised the data for peak (6:00 pm to 12:00 pm) and off-peak (12:01 am to 6:00 am) times for the sake of power saving during the off-peak time of traffic.

A. Vehicle movement rate in different types of roads

Traffic for different kinds of roads has been listed in table 1. We recorded the average (from 7 days data) of vehicle movement during the dusk to dawn for a whole day for all three streets, for calculation. We listed how many vehicles moved during this time. We observed that so many cars ran on the Major Arterial from dusk to dawn. The Minor Arterial also had vehicle pressure during this time. But we found that Collector roads had a little bit pressure up to 12 pm as compared to other routes. After 12 pm, there was not so much vehicle movement during 12.01am to 6 am-midnight in the city area. The graph in figure 4 depicts the calculation.

TABLE I. Vehicle Movement Rate in Different Roads during Dusk to Dawn

Time Period	Average no of vehicles counted on		
	Major Arterial	***Minor Arterial***	***Collector***
6.00-7.00pm	7452	5004	268
7.00-8.00pm	8332	5464	240
8.00-9.00pm	7336	5380	236
9.00-10.00pm	7144	4576	204
10.00-11.00pm	6868	2052	208
11.00-12.00pm	2728	1300	168
12.00-1.00am	1584	696	80
1.00-2.00am	816	508	12
2.00-3.00am	624	272	4
3.00-4.00am	516	172	4
4.00-5.00am	580	104	0
5.00-6.00am	788	304	44

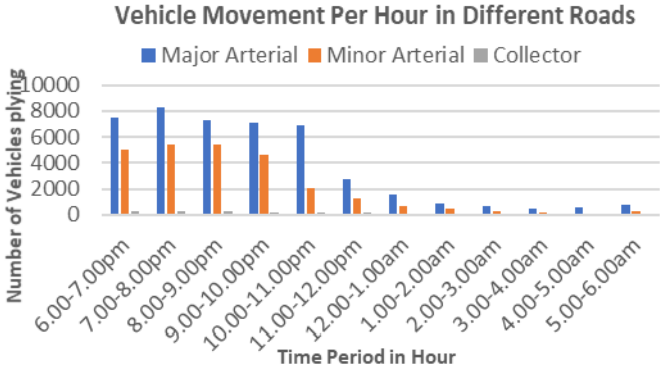


Fig. 4. Traffic movement of different roads from dusk to dawn in a day

Figure.4 shows that much traffic moves in the Major Arterial area and less traffic movement in the Collector area. From the figure, we can also see that after peak time (that means after 12 am) traffic movement in all roads decreases significantly.

B. Power Consumption

We have calculated total power consumption in all three types of roads. In present time Sodium lamps are being mostly used in all these roads. We have considered two lights for both sides of each kind of road. We also have considered the distance between two lamp posts for Major Arterial, Minor Arterial and Collector roads as 50 meters, 40 meters and 25 meters respectively.

TABLE II. Power Consumption by Sodium Light irrespective to Vehicle Movement [for 1 kilo meter]

Time Period	Power Consumption by Sodium Light in kWh		
	Major Arterial	***Minor Arterial***	***Collector***
6.00-7.00pm	4	5	8
7.00-8.00pm	4	5	8
8.00-9.00pm	4	5	8
9.00-10.00pm	4	5	8
10.00-11.00pm	4	5	8
11.00-12.00pm	4	5	8
12.00-1.00am	4	5	8
1.00-2.00am	4	5	8
2.00-3.00am	4	5	8
3.00-4.00am	4	5	8
4.00-5.00am	4	5	8
5.00-6.00am	4	5	8

According to the vehicle density per hour, Sodium Street light operates 100% brightness with constant power consumption value for each road. This is shown in table 2.

If we can replace the Sodium lamps with light-emitting diodes (LEDs), we will get the scenario, as per calculation shown in table 3 and table 4. We have programmed the system for 100% brightness of LED lights for 5 seconds, after detection of a moving vehicle or a human.

As the number of vehicles is greater than 1 for every 5 seconds in present time, the LED Streetlight operates at 100% brightness (without dimming), The ultrasonic sensors have been used to sense moving vehicles and people using the road. The recorded power consumption is shown in Figure 5 as per table 4. In this figure, the highest power consumption value is recorded 1.2 kWh from 6:00 pm to 2:00 am in Major Arterial and from 6:00 pm to 12.00 am in the Minor Arterial roads. For rest of the time, the recorded data start to decrease. Peak time is the rush hour for traffic movement as we have already seen in figure 5. So, both sodium and LED lamps consume the highest level of their power. There is no scope for dimming

the light (our proposed LED) due to the high rate of traffic movement.

TABLE III. BRIGHTNESS OF LED STREET LIGHT ACCORDING TO VEHICLE MOVEMENT RATE IN DIFFERENT ROADS DURING DUSK TO DAWN

Time Period	The brightness of LED Street Light according to Vehicle Movement		
	Major Arterial	*Minor Arterial*	*Collector*
6.00-7.00pm	100%	100%	10%
7.00-8.00pm	100%	100%	10%
8.00-9.00pm	100%	100%	10%
9.00-10.00pm	100%	100%	10%
10.00-11.00pm	100%	100%	10%
11.00-12.00pm	100%	100%	10%
12.00-1.00am	100%	10%	10%
1.00-2.00am	100%	10%	10%
2.00-3.00am	10%	10%	10%
3.00-4.00am	10%	10%	10%
4.00-5.00am	10%	10%	10%
5.00-6.00am	100%	10%	10%

Table 3 shows the brightness of LED Streetlight according to vehicle movement. In this table, we see that in Major Arterial, LED street light brightness is 10% from 2:00 am to 5.00 am, in the Minor Arterial from 12.00 am to 6.00 am, and in Collector Road LED light brightness is always 10% (from 6:00 pm to 6:00 am). The power consumption by LED lights with dimming technique is shown in table 4.

TABLE IV. POWER CONSUMPTION BY LED LIGHT (DIMMING 90%) ACCORDING TO VEHICLE MOVEMENT (FOR 1 KM ROAD)

Time Period	Power Consumption by LED Light in kWh		
	Major Arterial	*Minor Arterial*	*Collector*
6.00-7.00pm	1.2	1.5	0.24
7.00-8.00pm	1.2	1.5	0.24
8.00-9.00pm	1.2	1.5	0.24
9.00-10.00pm	1.2	1.5	0.24
10.00-11.00pm	1.2	1.5	0.24
11.00-12.00pm	1.2	1.5	0.24
12.00-1.00am	1.2	0.15	0.24
1.00-2.00am	1.2	0.15	0.24
2.00-3.00am	0.12	0.15	0.24
3.00-4.00am	0.12	0.15	0.24
4.00-5.00am	0.12	0.15	0.24
5.00-6.00am	1.2	0.15	0.24

***[Red Value Indicate Dimming]**

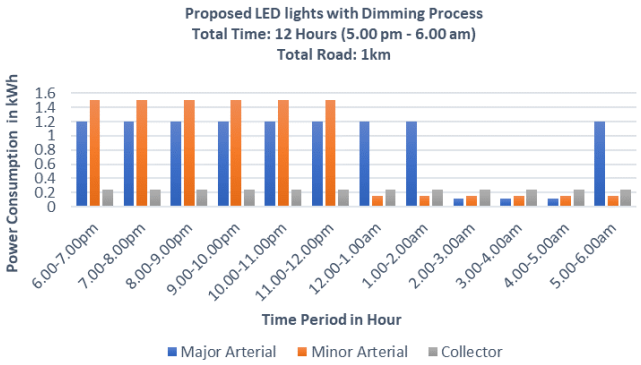


Fig. 5. Power Consumption by LED by Dimming Process

From the table 4 and figure 5, we can see that, the power consumption with the dimming operation is recorded at 2:00 to 5.00 am on Major Arterial, 12.00 am to 6.00 am on the Minor Arterial and 6:00 pm to 6.00 am in Collector roads. While the recorded data start to decrease and at the last moment it will be increased. LED light consumes less power in off-peak time because we have proposed that when traffic movement is relatively low, it (LED) will be dimming. A 30W LED light will consume only 3W, which is 10% while dimming. So, in off-peak time, power consumption will be much less.

TABLE V. COMPARISON OF POWER CONSUMPTION PER DAY PER KM

Power Consumption Duration (5.00 am – 6.00 pm)	**100W Sodium lamps**	**Proposed LED**
As per tables -2 & 4	204 kWh	23.94 kWh
Dimming Condition	Not applicable	Applied

From the table 5, we have seen that in total 12 hours, 100w sodium light consume 204 kWh per day and 30w LED light consume 23.94 kWh per day. This is because; from table 4 we can see that, the maximum power consumed by LED lamps is 1.2 kWh and the minimum is 0.24 kWh. Hence, our proposed model will save electricity cost significantly.

TABLE VI. COMPARATIVE COST ANALYSIS

Expenditure	**For calculated 1KM (in BDT)**		**For Chittagong city (3500 Sodium lamps) [13] (in BDT)**	
	100w Sodium Lamp	LED (90% dimming)	100w Sodium Lamp	LED (90% dimming)
Per day	1,570.80/=	184.34/=	32,340/=	3,795.26/=
Per month	47,124/=	5,530.14/=	9,70,200/=	1,13,857/=
Per year	5,65,488/=	66,362/=	1,16,42,400/=	13,66,285/=

VI. DISCUSSION AND CONCLUSION

From table 6, we can see how much money is needed for sodium and our proposed LED lighting systems. Now, 3,500 sodium lamps are used for Chittagong city streets [13]. We have seen that per day 32,340/= taka is needed as the expense

for sodium lighting for the whole of Chittagong city whereas using our LED light will cost 3,795/= taka only as per latest billing rate in Bangladesh [14]. Also, per month charge of sodium lighting and LED has a vast difference. Sodium lighting needs 9,70,200/= taka and LED needs 1,13,857/= taka only. Consequently, the yearly expense behind Sodium is 1,16,42,400/= taka and LED 13,66,285/= taka. So, a significant difference is apparent between using our proposed adaptive LED lights and existing sodium lights. Our system will save almost 88.26 % cost of the total consumed power of the present street lighting system of the Chittagong City in Bangladesh, as well as it will reduce the cost for any other city in the world.

REFERENCES

[1] F. Li, D. Chen, X. Song and Y. Chen, "LEDs: A Promising Energy-Saving Light Source for Road Lighting," *2009 Asia-Pacific Power and Energy Engineering Conference*, Wuhan, 2009, pp. 1-3.

[2] R. M. Ramli, Y. Z. Arief and P. D. A. Aziz, "Application of LED technology into public road lighting in Malaysia for replacing the high-pressure sodium vapour lighting," *2015 International Conference on Sustainable Energy Engineering and Application (ICSEEA)*, Bandung, 2015, pp. 76-81.

[3] T. Novak and K. Pollhammer, "Traffic-Adaptive Control of LED-Based Streetlights" in *IEEE Industrial Electronics Magazine*, vol. 9, no. 2, pp. 48-50, June 2015.

[4] M. Joshi, R. Madri, S. Sonawane, A. Gunjal and D. N. Sonawane, "Time Based Intensity Control for Energy Optimization Used for Street Lighting," *2013 Texas Instruments India Educators' Conference*, Bangalore, 2013, pp. 211-215.

[5] G. Gagliardi *et al*., "A smart city adaptive lighting system," *2018 Third International Conference on Fog and Mobile Edge Computing (FMEC)*, Barcelona, 2018, pp. 258-263.

[6] G. Velasco-Quesada, M. Roman-Lumbreras and A. Conesa-Roca, "Comparison of Central Dimmer Systems Based on Multiple-Tapped Autotransformer and High-Frequency Switching Converter," in *IEEE Transactions on Industrial Electronics*, vol. 59, no. 4, pp. 1841-1848, April 2012.

[7] D. R. Khade, N. V. Gajane, S. N. Gawade and R. A. Metri, "Intensity controller of LED street lights," *2017 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, Kollam, 2017, pp. 1-4.

[8] D. Huaqiu, C. Yong and X. Huaqiang, "An Intelligent Control of LED Street Lamp Based on Illumination Sensing," *2015 Fifth International Conference on Communication Systems and Network Technologies*, Gwalior, 2015, pp. 1189-1192.

[9] Y. Wu, C. Shi, X. Zhang and W. Yang, "Design of new intelligent street light control system," *IEEE ICCA 2010*, Xiamen, 2010, pp. 1423-1427.

[10] M. F. Pinto, T. R. F. Mendonça, C. A. Duque and H. A. C. Braga, "Street lighting system for power quality monitoring and energy-efficient illumination control," *2016 IEEE 25th International Symposium on Industrial Electronics (ISIE)*, Santa Clara, CA, 2016, pp. 34-39.

[11] P. P. F. Dheena, G. S. Raj, G. Dutt and S. V. Jinny, "IOT based smart street light management system," *2017 IEEE International Conference on Circuits and Systems (ICCS)*, Thiruvananthapuram, 2017, pp. 368-371.

[12] Arduino "what is Arduino" Arduino Guide Instruction Available in http://arduion.cc/en/guide/introduction [Accessed: 20 September 2018]

[13] Department of Engineering-Electrical, "Chattogram City Corporation," 2018. http://www.ccc.org.bd/electrical. [Accessed: 2018- 09- 20].

[14] Commercial: Tariff: 2018. http://www.bpdb.gov.bd/bpdb/index.php?option=com_content&view=article&id=35&Itemid=33. [Accessed: 2018- 09- 20].

Detection of Safe Angle for Water Vehicles due to Vulnerable Swing

Md. Shahidul Islam*, Sarfaraz Newaz†
*†Computer Science and Engineering Discipline
Khulna University, Khulna, Bangladesh
†Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
Email:*shahidcseku@gmail.com, †snewaz@cseku.ac.bd

***Abstract*—Steering of water vehicles is mainly dependent on the aero and hydrodynamic forces. The adverse situation caused by any of the natural disasters inhibits the movement of vehicles. During the hostile situation, the most important thing is safety. In this paper, a method to compute the optimal safe position for water vehicles during instability caused by the vulnerable swing is proposed. A 3-axis accelerometer is used to obtain the roll, pitch, and yaw angles, and using these, we derive a function which gives us the instantaneous instability of the vehicle. If the instability exceeds a predefined threshold, our proposed method calculates the safe angle to reach the optimal safe position. This method is computationally efficient and is suitable for on-board real-time implementation as well as autonomous marine robotic vehicles.**

***Index Terms*—Safe angle, Y-Factor, tilt angle.**

I. Introduction

During the summer and rainy season, along with the increasing water level in the river and ocean all over the world, a vast quantity of waves are generated by the rain or stormy weather. Due to the rapid rotation of the waves, steersman cannot understand which direction the boat needs to be headed. The safe position of a waterways vehicle is the state, in which the vehicle is most stable in terms of wind, vulnerable swing, and whirlpool – which may cause the vehicle to sink into water or displace its heading. To remain safe in the hostile situation caused by any of the natural disasters, it is necessary to find the safe position. The safe angle can be defined as the angle between the current position of the vehicle and the safe position. Thus, measurement of the safe angle is the precondition to reach to the safe position. In this paper, we propose a method of getting the safe angle due to vulnerable swing.

To do this job, at first, we identify the tilt, which exceeds the threshold values. Here, threshold values are the pre-calculated tilt angles in which the boat is in danger. It varies from boat to boat. Tilt angles are calculated from a 3-axis accelerometer. To set the threshold values, we first run an experiment on the boat. Then we determine the safe angle depending upon the threshold, the boat factors, and tilt angle. Our device is capable of measuring the safe angle dynamically.

Our contribution to this paper:

- we calculate the roll and pitch angle of different types of boats and established a relationship between them.
- we perform real boat experiment using our device.
- we propose a complete hardware and software architecture for determination of safe angle for any type of water vehicle in dynamic and unknown environment.

The remainder of this paper is organized as follows: Section II highlights on an overview of related work. Section III elaborates the principle of the measurements. Section IV and V describe the implementation of the proposed technology. Section VI delineate the experimentation and results. Finally, Section VII concludes this paper.

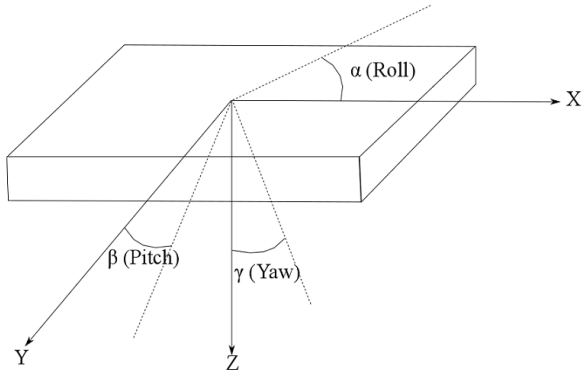


Fig. 1: Tilt angles from 3-axis accelerometer.

II. Related Works

In recent years, there were published different works based on tilt angle calculation. Among them,

Yosi et al. proposed a method to calculate the tilt and heading from an Inertial Measurement Unit (IMU) that have 3-axis gyroscopes, 3-axis accelerometers, and 3-axis magnetometers [1]. They have measured tilt values using Euler angles, the pitch (α), roll (β), and yaw (γ). By using sensor fusion algorithm they have shown that the output is likely similar to the actual value, however, the one without sensor fusion has a growth error value every step [1].

Saoud et al. present a method for calculating an optimal sailing angle that maximizes the speed of a sailboat ensuring that the heeling angle remains bounded by a pre-defined value [2]. They have shown that using this method, the heeling angle is implicitly limited to $\pm\pi/4$.

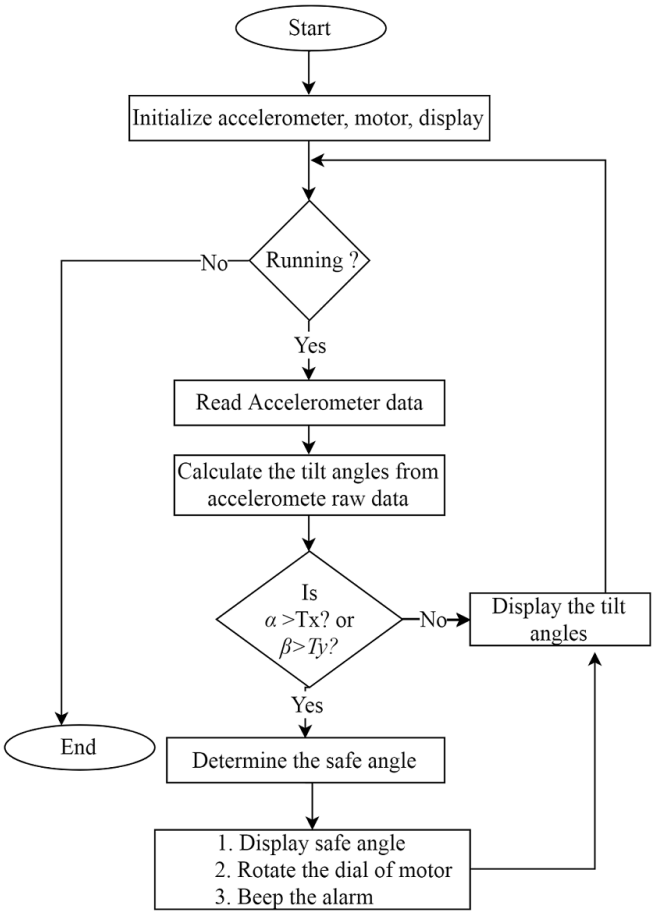


Fig. 2: System architecture of the proposed method.

Surendran et al. discussed two ways of measuring ship stability - using the available linear ship motion theory and the nonlinear theoretical model. They have obtained the solutions using the fourth order Runge-Kutta (RK) method [3]. Mainly they discussed the importance of roll angle in the stability analysis of a ship.

Furthermore, Newaz et al. described a method to generate an alarm system for waterways vehicle due to experiencing vulnerable swing [4]. They have measured tilt values using Euler angle principle. Although they have calculated threshold values, yet they did not mention the strength of boat. Moreover, they did not consider the shape of the boat. Therefore, the system can only generate an alarm in the dangerous situation, however, cannot determine the safe angle.

III. Methods

A. Measuring tilt angle

In this work, the ADXL345 accelerometer – which is a MEMS (Micro-Electromechanical System) based accelerometer, has been used to calculate the tilt angle. It is capable to provide high-precision output up to ±16g. To calculate the tilt of an object using earth's gravity, a ±1g accelerometer is enough. The accelerometer experiences acceleration in the range from -1g to +1g through 180° of tilt.

$$1g = -9.8m/s^2 \tag{1}$$

TABLE I: Output of the accelerometer for different orientation of the device.

Device position	Accelerometer reading		
	A_x	A_y	A_z
X_b Down	+g	0	0
X_b UP	-g	0	0
Y_b Down	0	+g	0
Y_b Down	0	-g	0
Z_b Down	0	0	+g
Z_b Up	0	0	-g

In Table. I, A_x, A_y, and A_z are the acceleration in x, y, and z axes respectively measured by the accelerometer. The changes in orientation of the device are described by rotations in roll α, pitch β, and yaw γ about the x, y, and z axes respectively [5]. The roll, pitch, and yaw rotation matrices are:

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\alpha) & sin(\alpha) \\ 0 & -sin(\alpha) & cos(\alpha) \end{bmatrix} \tag{2}$$

$$R_x(\beta) = \begin{bmatrix} cos(\beta) & 0 & -sin(\beta) \\ 0 & 1 & 0 \\ sin(\beta) & 0 & cos(\beta \end{bmatrix} \tag{3}$$

$$R_x(\gamma) = \begin{bmatrix} cos(\gamma) & sin(\gamma) & 0 \\ -sin(\gamma) & cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

These matrices transform the earth's gravitational field vector g by angles α in roll, β in pitch, and γ in yaw about the x, y, and z axes respectively. Hence, equations driven for calculating the angles from the accelerometer outputs are as follows [6]:

$$\alpha = \arctan\left(\frac{A_x}{\sqrt{A_y^2 + A_z^2}}\right) \tag{5}$$

$$\beta = \arctan\left(\frac{A_y}{\sqrt{A_x^2 + A_z^2}}\right) \tag{6}$$

$$\gamma = \arctan\left(\frac{A_z}{\sqrt{A_x^2 + A_y^2}}\right) \tag{7}$$

The coordinate system of the accelerometer which will be used in this document are shown in Fig. 1. Here-

- the x-axis is aligned along the width of the device.
- the y-axis is aligned along the length of the device.
- the z-axis points downwards so that it is aligned with gravity and also aligned at right angles to both the x and y axes.

B. Measuring boat-factors

Due to their different shapes, the capability of floating and tolerance of vulnerable swing varies from one boat to another. When a ship is at sea, due to the influence of waves, there cause a variety of motions, such as rolling, yaw, pitch, sway, heave, and so on. Among the 3 tilt angles − roll, pitch, and yaw the roll angle has a great importance on its stability and safety. The significance of yaw angle is negligible [2]. Hence we are considering only roll and pitch angles. To understand the impact of roll and pitch on the boat we need to find the relationship between them.

The length of a boat is usually termed as LOA (Length overall). On the other hand, the width of it is known as BOC (Beam on Centerline). The ratio of LOA/BOC is used to estimate the stability of a boat. The lower the ratio, the greater the boat's stability. For example, a boat with LOA/BOC = 1 is square or circular in shape and has the highest stability. Usually, the LAO/BOA for small sailboats are from 2:1 to 5:1 and for large ships the ratio is 20:1 [7]. Due to the architectural shape of the boats, the bow is more stable than beam during the swing while on the water. However, due to different draft or draught of a boat's hull, the LAO/BOC ratio does not always work although if we consider draft in the calculation.

Roll motion is considered to be the most important in the stability analysis of a water vehicle. The factors that dominate roll angles of different water vehicles are [3]:

- the ratio of the period of rolling and the encountering period of the wave
- the shape of the hull, total weight, and floatage
- the wave steepness, H/L, where H and L are the height and length of the wave
- the damping effect of the underwater parts of the hull
- the encountering angle and speed of the wave

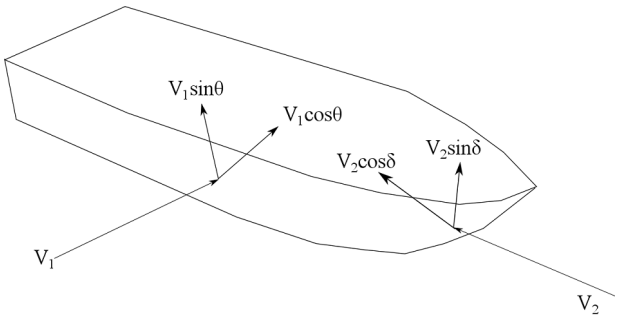


Fig. 3: Y-$Factor$.

Lemma 1: Tilt angle is proportional to the sine component of the wave force acting on a boat.

Proof: Assume that, currently the boat position is at θ degree to the wave force. When the wave, V strikes on the boat, the wave force acts on the boat with two components: $F_x = Vcos\theta$ and $F_y = Vsin\theta$. The F_x cause the thrust which is responsible to push the boat forward and F_y is the cause of tilting. Therefore, tilting depends only on F_y. Mathematically,

$$Tilt \propto Vsin\theta \tag{8}$$

From eqn. 8, the wave V_1 with θ degree will tilt the boat with $V_1 sin\theta$ force. Similarly, the wave V_2 with δ degree will tilt the boat with $V_2 sin\delta$ force. If $V_1 = V_2$ and $\theta = \delta$, then we get, (tilt by $V_1 sin\theta$) = k×(tilt by $V_2 sin\delta$). Where k is a proportional constant. The tilt angle caused by $V_1 sin\theta$ is α and $V_2 sin\delta$ is β. Hence, $k = \alpha/\beta$. We named k as Y-$Factor$. Therefore, the Y-$Factor$ is defined as the ratio of rolling α and pitch angle β of the boat.

$$Y\text{-}Factor = \frac{\alpha}{\beta} \tag{9}$$

The above equation gives the Y-$Factor$ of a boat. For any specific boat, before measuring the safe angle, we need to determine the Y-$Factor$ of it. Additionally, the actual F_y is Y-Factor times greater than the apparent F_y. Therefore,

$$F_y = Y\text{-}Factor \times \beta \tag{10}$$

C. Measuring threshold values

We will use two thresholds- T_x, the threshold for x-axis and T_y, the threshold for y-axis. We will measure the maximum threshold values, $T_{x_{max}}$ and $T_{y_{max}}$ for the boats in which angle the boat will sink. Then we can use any appropriate threshold values $1 \leq T_x \leq T_{x_{max}}$ and $1 \leq T_y \leq T_{y_{max}}$. Note that the threshold values are for both positive and negative tilt angles. Therefore, we take absolute values as input.

D. Measuring safe angle

Safe angle θ is defined as the angle between the current position of the boat and the position in which the boat is in the highest stable in the situation of vulnerable swing.

To simplify the calculation we assume 2 cases. They are:

Assumption 1: The wave velocity is assumed to be always parallel to the water surface, moreover is also assumed to be flat.

Assumption 2: Tilt angle along y-axis is not-null ($|\beta| > 0$) for all time.

If two forces $\vec{F_1}$ and $\vec{F_2}$ acting on an object, then the angle of the resultant force with respect to $\vec{F_1}$ is given by the following equation [8],

$$\theta = arctan\left(\frac{\vec{F_2}}{\vec{F_1}}\right) \tag{11}$$

Applying on eqn. 11 we get,

$$\theta = arctan\left(\frac{\vec{F_\alpha}}{\vec{F_\beta}}\right) = \arctan\left(\frac{|\alpha|}{Y\text{-}Factor \times |\beta|}\right) \tag{12}$$

Where $\vec{F_\alpha}$ is the force that creates roll angle α and $\vec{F_\beta}$ is the force that creates pitch angle β.

If $|\alpha| \geq T_x$ or $|\beta| \geq T_y$, then we step to find the safe angle. The safe angle is calculated between 0 to 90 degrees. Then the direction of rotation is calculated. Hence, the range of safe angle remains 0 to ± 90 degrees. The direction of the angle is determined by Table II.

TABLE II: Safe angle direction.

X	Y	Angle direction
$\alpha \geq 0$	$\beta \geq 0$	Clockwise
$\alpha \geq 0$	$\beta < 0$	Anti-Clockwise
$\alpha < 0$	$\beta \geq 0$	Anti-Clockwise
$\alpha < 0$	$\beta < 0$	Clockwise

IV. Hardware Implementation

The main board used for this system is Arduino Uno 3 which contains an ATmega 328 microcontroller, 2 MB built-in ROM. All other components are connected with the Arduino Uno 3 board. The adxl345 accelerometer gives the acceleration data to the main board. The main board process data to get the tilt angle. Under certain conditions, the Uno board display the Safe Angle to the display. It also displays the regular tilt angle. An indicator dial connected to the servo motor to display the safe angle in an analog way. If the boat is not in safe position, then a continuous alarm will be played in buzzer. Fig. 4 shows the block diagram of our proposed device. In Fig. 5, a snapshot of our developed device is shown. Fig. 5a is the internal circuitry of the device and Fig. 5b is the top view of our device.

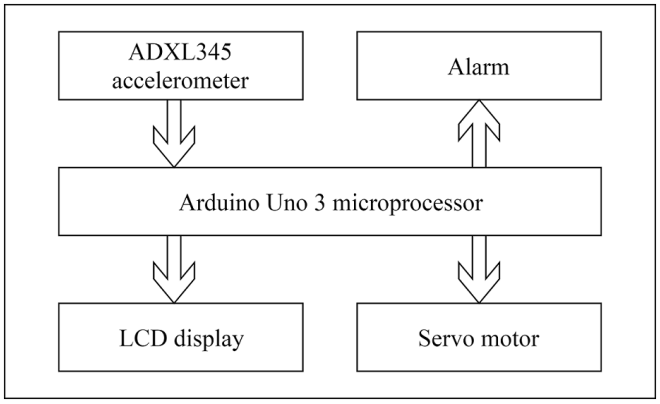
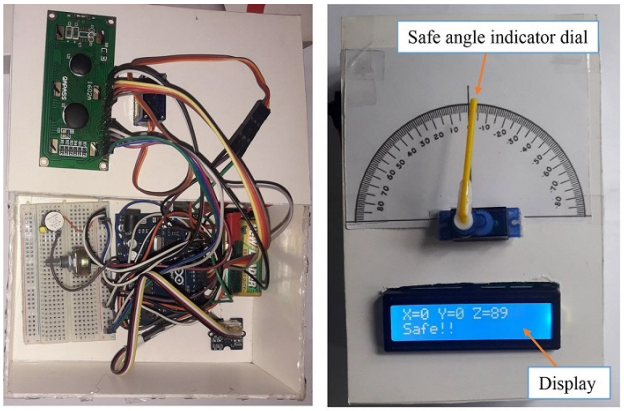


Fig. 4: Block diagram of the device.

V. Software Implementation

We implement the software of the proposed system in C++. At first, the components of the system initialized. Next, the Arduino reads the accelerometer data and process to determine the tilt angles. If $|\alpha|$ is greater than or equal to T_x or $|\beta|$ is greater than or equal to T_y then calculate the safe angle. The safe angle is normalized between 0 to ± 90 degrees and displayed in the LCD as well as indicated in spines of the servo motor. Otherwise, display the tilt angle and place the indicator in the zero degrees. The process repeats unless the device turned off. The algorithm for our developed system is given in Algorithm 1.



(a) (b)

Fig. 5: Our developed device. (a) Internal circuitry (b) Snapshot

Algorithm 1 Algorithm for the developed system.

Input: Y-Factor, Roll (α), Pitch (β), and Yaw (γ) angles
Output: Safe angle (θ)
Initialize LCD Display, Servo Motor, Accelerometer (adxl345), and Buzzer
while (true) **do**
    Read data (A_x, A_y, A_z) from Accelerometer
    Process data to get tilt angles in degree using equations 5, 6, and 7
    Show tilt angles on LCD Display
    if ($|\alpha| \geq T_x$ or $|\beta| \geq T_y$) **then**
        Calculate safe angle (θ) using equation 12
        Show safe angle on LCD Display
        Show safe angle using the indicator dial attached with the servo motor
        Sound alarm
        Blink light
    else
        Clear safe angle from display and print "Safe"
        Place the indicator dial attached to the motor at zero degrees
    end if
end while

VI. Experimental Results

A. Y-Factor Result

We have tested in unwavering water on the stationary boat by creating artificial waves. The device was installed parallel to the boat. Table III shows the 10 readings of roll and pitch angle as x_{α_1} and y_{β_1} respectively for port and starboard and 10 others x_{α_2} and y_{β_2} for stern and bow.

We perform several tests on different types of boats and compare their $Y\text{-}Factor$ in Table IV. The detailed data are displayed in Fig. 8. We observe from Fig. 8a that, the artificial wave force is applied along the x-axis of boat no. 1. Therefore, the tilt angle with respect to x-axis α_1 vary from 10 to 30 degrees and tilt angle with respect to y-axis β_1 is nearly zero

TABLE III: Data for measuring $Y\text{-}Factor$.

	x_{α_1}		x_{β_1}		y_{α_2}		y_{β_2}	
Sl.	Port	Star board	Port	Star board	Stern	Bow	Stern	Bow
1	22	19	1	0	1	1	10	10
2	23	18	1	1	1	1	10	10
3	23	20	1	1	1	1	10	9
4	23	20	0	1	1	1	12	9
5	25	19	0	1	1	1	11	10
6	26	20	0	0	0	2	10	10
7	27	21	0	1	1	1	10	9
8	27	21	0	1	1	3	10	10
9	27	20	0	1	1	1	10	10
10	28	17	0	1	1	1	10	7
Total	251	195	3	8	9	13	103	94
Avg.	25.1	19.5	0.3	0.8	0.9	1.3	10.3	9.4

degrees. After that, when the force is applied along the y-axis of boat no. 1 then, the boat experience a little swing along the x-axis (α_2) and the tilt with respect to y-axis β_2 remains under 10 degrees. These two different situations raised while both the wave force applied on the boat has the same value. Rest of the observations also admit the same property except for Fig. 8e as the boat no. 5 have no change in x-axis and y-axis. From these experimental values we can say that, the LOA/BOC ratio does not always works since they differ in their height.

TABLE IV: $Y\text{-}Factor$ Comparison for several types of boats.

Sl.	LOA (ft)	BOC (ft)	X (θ°)	Y (θ°)	$Y\text{-}Factor$
1	9	2	21.83	9.27	2.35
2	12	3	19.5	7.99	2.44
3	12	3	23.85	9.57	2.49
4	5	2	31.15	4.44	7.02
5	0.66	0.66	41.32	41.03	0.99

B. Threshold values

We have tested on the same five types of boat from $Y\text{-}Factor$. The roll angle α just before drowning the boat is the maximum threshold $T_{x_{max}}$ and the pitch angle β is the maximum threshold $T_{y_{max}}$. The threshold values are given in Table V.

TABLE V: Maximum threshold values for several types of boats.

Sl.	LOA (ft)	BOC (ft)	$T_{x_{max}}$ (θ°)	$T_{y_{max}}$ (θ°)
1	9	2	30	44
2	12	3	34	45
3	12	3	34	46
4	5	2	35	40
5	0.66	0.66	47	47

C. Safe Angle Result

Table IV gives the $Y\text{-}Factor$ of different types of boats that we experimented. From Table V, we get the maximum thresholds that can be set to a boat. We have run over experiment on a real boat (boat no. 1) with T_x = 10, T_y= 5, and Y_f = 2.35 and another boat (boat no. 4) with T_x = 20, T_y = 20, and Y_f = 1. The boats are assumed to be safe state if the tilt angles are below the thresholds values ($\alpha < 15$, $\beta < 10$ for boat no. 1 and $\alpha < 20$, $\beta < 20$ for boat no. 4). In this state, the boat is in the safe position and at low risk. Hence, we have nothing to do with the boat. If the tilt angles exceed the threshold values then the device calculates the safe angle form the tilt information. We rotate our boat accordingly with our device's calculation to reach the safe position. After reached the safe position, the boat was displayed as safe on the display. Since the wave was not unidirectional therefore the boat always faced some tilts. Yet we can see that, it was the minor tilt angles after we rotate the boat.
The result for the boat no. 1 shown in Fig. 6. The result for the boat no. 4 shown in Fig. 7. Both Fig. 6 and Fig. 7 shows the 100 experimental reading. The device gives the tilt angles and the safe angle. According to the device suggestion when we rotate our boat, we can see in both Fig. 6 and Fig. 7, the safe angle values minimizes and the boat achieves more stability. As there always some wave in the river, therefore, our device continuously display the safe angle with a 0.5 sec interval. It is our observation that the safe angle between 0 and ± 5 degrees is practically safe.

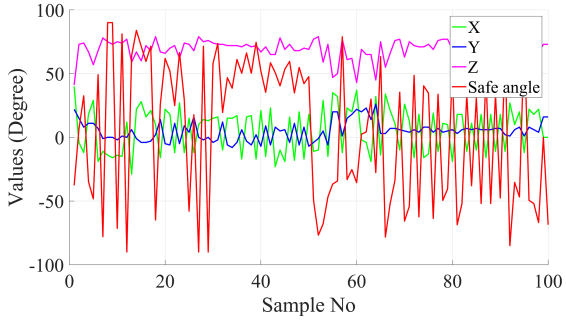


Fig. 6: Experimental reading from boat no. 1.

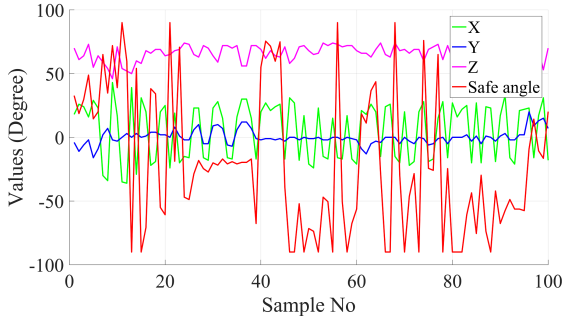


Fig. 7: Experimental reading from boat no. 4.

VII. Conclusion

In this paper, a method has been designed to determine the safe angle for water vehicles in order to remain safe during the vulnerable swing caused by the natural disaster. Due to the mechanical imbalance, the analog output may vary from ± 2 degrees. Our calculated values for $Y\text{-}Factors$ are

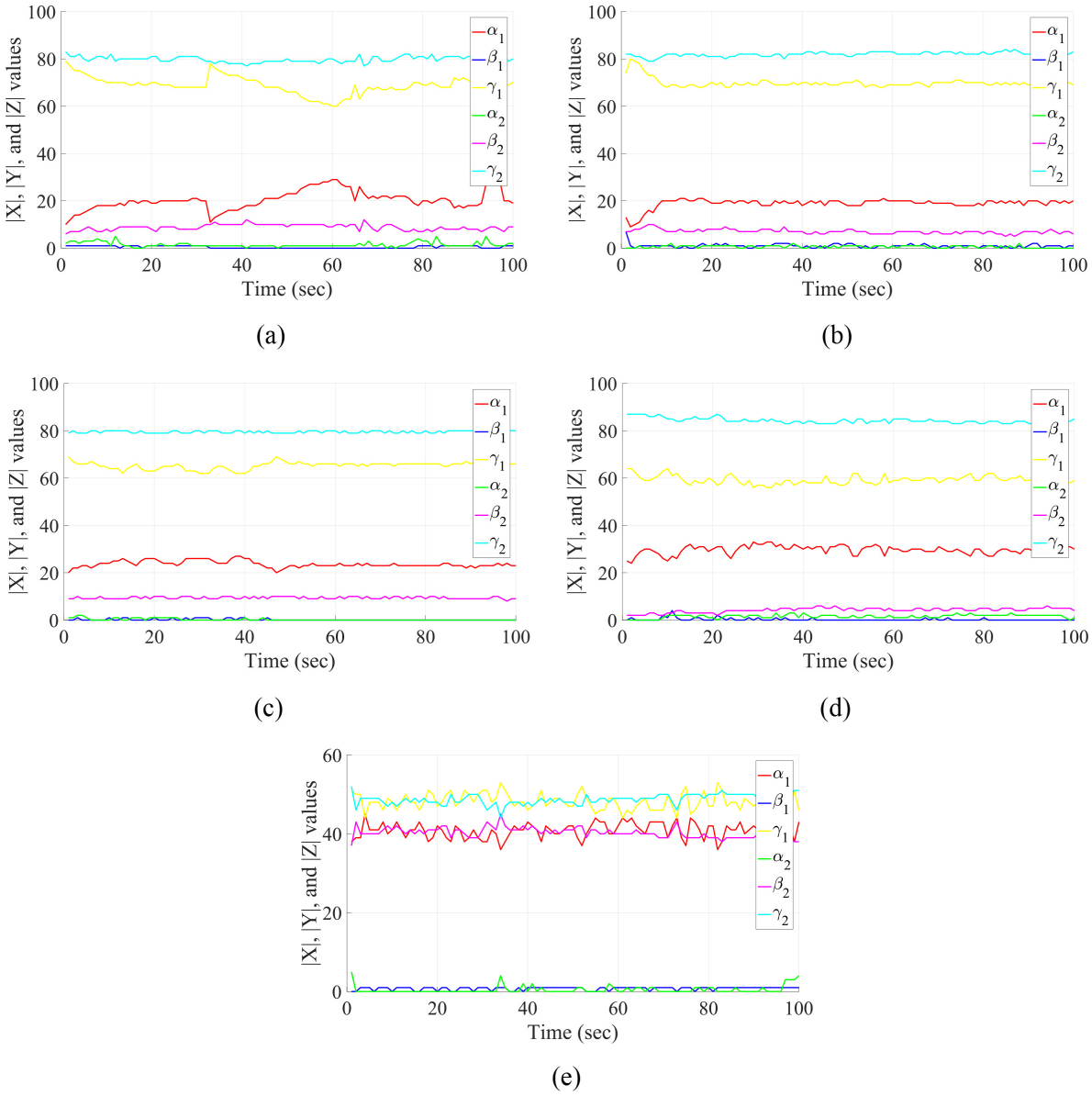


Fig. 8: $Y\text{-}Factor$ result of 5 boats. Tilt angles when the artificial wave force applied on: (a) Boat-1 (b) Boat-2 (c) Boat-3 (d) Boat-4 (e) Boat-5 . α_1, β_1, γ_1 are the tilt angles when wave force applied on x-axis and α_2, β_2, γ_2 for y-axis.

not universal to all the boats with the same dimension. It depends on the area of the hull, water line, height of the boat, materials of the boat, etc. Therefore, before using the device, the $Y\text{-}Factor$ of the boat must be calculated explicitly.

The average maximum threshold values are taken from several test case just before drowning the boat. The experiment was done artificially. In practical, the maximum threshold or the tilt angles while drowning may depend on the nature of the wave.

REFERENCES

[1] S. Yosi, N. Agung, and S. Unang, "Tilt and heading measurement using sensor fusion from inertial measurement unit," in *Control, Electronics, Renewable Energy and Communications (ICCEREC), 2015 International Conference on*, pp. 193–197, IEEE, 2015.

[2] H. Saoud, M.-D. Hua, F. Plumet, and F. B. Amar, "Optimal sail angle computation for an autonomous sailboat robot," in *Decision and Control (CDC), 2015 IEEE 54th Annual Conference on*, pp. 807–813, IEEE, 2015.

[3] S. Surendran and J. V. R. Reddy, "Numerical simulation of ship stability for dynamic environment," *Ocean Engineering*, vol. 30, no. 10, pp. 1305–1317, 2003.

[4] S. Newaz, S. S. A. Romeo, and A. A. Al Islam, "Towards devising a solution for generating an alarm in a vehicle on waterways experiencing vulnerable swing," in *Networking, Systems and Security (NSysS), 2017 International Conference on*, pp. 195–200, IEEE, 2017.

[5] M. Pedley, "Tilt sensing using a three-axis accelerometer," *Freescale semiconductor application note*, vol. 1, pp. 2012–2013, 2013.

[6] C. J. Fisher, "Using an accelerometer for inclination sensing," *AN-1057, Application note, Analog Devices*, 2010.

[7] W. B. Hayler and J. M. Keever, *American merchant seaman's manual: for seamen by seamen*. Cornell Maritime Press, 2003.

[8] R. A. Serway and J. W. Jewett, *Physics for scientists and engineers with modern physics*. Cengage learning, 2018.

A Comparative Study of Walking Assistance Tools Developed for the Visually Impaired People

Mahfuza Khanom, Muhammad Sheikh Sadi, and Md. Milon Islam
Department of Computer Science and Engineering
Khulna University of Engineering & Technology, Khulna-9203, Bangladesh
mahfuza7508@gmail.com, sadi@cse.kuet.ac.bd, and milonislam@cse.kuet.ac.bd

***Abstract*—The number of visually impaired people is growing with the fast increase of population. The visually impaired people face much difficulty in their day-to-day lives due to losing their vision. This paper analyzed the performances of the raspberry pi and Arduino for the obstacle detected to assist the visually impaired people. We have used raspberry pi, Arduino and ultrasonic sensor as hardware components in this paper. The position of the obstacles is varied from the sensor and the data is measured corresponding to actual distance. The average accuracy and error rate are obtained 98.47% and 1.53% from the raspberry pi. In addition, Arduino achieved the average accuracy and error rate of 95.22% and 4.78%. Raspberry pi obtained more accuracy and lower error rate than Arduino.**

Keywords- Raspberry Pi; Ultrasonic Sensor; Arduino UNO; Visually Impaired; Obstacle Detection.

I. INTRODUCTION

As per World Health Organization (WHO) statistics [1], there are approximately 253 million visually impaired people. Among them, 36 million people are blind and about 217 million people have low vision. The people who are visually impaired need help mostly to continue their daily activities. Most importantly they are to know the position of the obstacle and other interference [2] when they are in new or unknown places for their safe navigation. Sometimes, the visually impaired people use dogs as their guide for their navigation. However, this is expensive as well as difficult to maintain. The use of a stick or cane [3] is an alternative option for the visually impaired people to detect an object without any prior idea.

Modern science is trying to reduce the difficulties of visually impaired people. Some of the tools have already been developed to fulfil the demand of the blind users. Navguide [4] provides the surroundings information and detect the knee-level and head level obstacles to the users. Smart walking stick [5] detect obstacles, pit, and water on the road to help the blind. Electronic Mobility Cane used Atmega328 microcontroller for obstacle detection and iWalk [6] detects the obstacles using Arduino having the accuracy of about 90%. A wearable head mounted device [7] detects the obstacle in left, right and front direction using ultrasonic sensor and the Arduino.

To aid the visually impaired people, previously we have proposed few approaches as shown in [8], [9] and [17]. However, these approaches are able to detect the obstacles in front of the visually impaired people. These tools are sensor based and able to detect the obstacles and generate an alarm signal to notify the users. The existing literature reflects that most of the walking assistance devices are developed using microcontrollers (Arduino, Raspberry Pi etc.). However, an extensive comparative study to evaluate the performance of these technologies is somehow missing. Hence, we have analyzed detection accuracy and error rate for both raspberry pi and Arduino in our study.

The rest of the paper is organized as follows. The related work that covers the recent developments in this area is described in Section II. The major hardware components that are used to develop the system are described in Section III. The system design is illustrated briefly in Section IV. Section V outlines the experimental outcomes of the system. Section VI concludes the paper.

II. OVERVIEW OF THE STATE OF ART

Many research groups from different parts of the world have been developing different systems to help the visually impaired people with their daily lives. The work associated with this field is outlined briefly as follows.

Mulky et al. [10] developed an autonomous scooter that can aid the visually impaired people to navigate individually and securely in unknown environment. The system provides the current location, maps, and nearby obstacle's location to the users. The system used stereo camera and raspberry pi as sensing and processing device respectively. However, the scooter provides short-range vision-based mapping. Sareeka et al. [11] developed a wearable device named as pseudoEye that captures the texts from the environment and converts it to voice signal which is delivered to users. The system used a mini camera, a push button, raspberry pi and a SoC (System on Chip) attached to a cap. The raspberry pi captures the data from the environment using camera and extracts texts from the images and generates voice signal corresponding to texts. A portable power bank is used to power on the raspberry pi. However, the system can recognize few numbers of texts and fails to detect handwritten character. Xiao et al. [12] proposed a low cost system to aid the visually impaired people using raspberry pi. The system used magnetic compass and gyroscope to measure the forward direction of the users. In addition, the users can provide input their address through a microphone. The device is light weight and wearable at waist. The overall production cost of the proposed system is approximately $138. The system provides 3-D sound as a feedback signal. However, the system cannot recognize landmarks and the position as well as the velocity of the users is not addressed in here. S. Khade and Y. Dandawate [13] developed an obstacle detection framework for the visually impaired people in unknown surroundings using Raspberry Pi. The system perceived the obstacles in the ROI without the help of cameras. The hardware used in this

scheme is 'Raspberry Pi 2-B', a wearable camera with 5MP resolution and the overall system is simulated in MATLAB as well as Python language. The authors generated the database by visiting a school for the blinds. The database is retrieved by capturing the video of the blind head, and approximately 25 videos are taken for the research purpose. The database is converted into a video frame initially, and the motion vectors are used for the obstacle detection. The segmentation is done using background subtraction procedure. Finally, the Region of Interest (ROI) is created to detect the obstacles. The system is cost-effective and has lesser weight, but the illumination and brightness situations are constant. However, in the real-world scenario, these properties change frequently. So, the prototype is not compatible with the real world. The system can detect the obstacles only, but is not able to categorize them.

Aljahdali et al. [14] proposed a smart assistive walker device that decreases the risks of accidents of the visually impaired people. The hardware parts of the system are sensor, controller, output and the software part of the system are "Map" and "Find Me" Apps. The sensors collect the real time data and when the system detects an obstacle, the controller sends it via IoT signal to the output to avoid collision. However, the system is not able to measure the distance from the user to obstacle and unable to alert the authority. Khan et al. [15] developed a mini hand-carrying device that can be used as an alternative of smart stick/cane. The system can detect the obstacles, manholes as well as potholes and notify the users by alarming buzzer. The system used Arduino UNO, two sonar sensors, two buzzers, 9v battery and a voltage converter. However, the performance of the system has not been assessed by the blind users. Sharma et al. [16] developed a smart stick for the visually impaired people in order to provide independent mobility. The system can detect the ground level and knee level obstacles, potholes and staircases. It provides vibration signal with the presence of hindrances. The system is comprised of three sonar sensors, a moisture sensor, two vibration motors, two buzzers and Arduino microcontroller. However, the detection range of pothole is approximately 30 cm which is too small to avoid risks of accidents.

Kamal et al. [17] developed a walking assistant for the visually impaired people to aid them in navigation by avoiding obstacles around them. The system consists of two modules such as wearable spectacle module and smartphone module. The wearable spectacle module perceives an object midpoint and receives the copy of the surface for processing. The smartphone module obtains the copy sent from the wearable spectacle module and then processes the copy and finally computes the smoothness of the surface. The developed module is small, wearable and lightweight. Pissaloux et al. [18] proposed a system that assists users in walking with obstacle avoidance, orientation, real physical displacement, etc. It is based on the real gist. The system consists of TactiPad for obstacle detection and tactile gist for display. The system is beneficial for the visually impaired people in their mobility in an unfamiliar environment, but the system is comparatively bulky. Ponnada et al. [19] proposed a manhole and staircase detection system for the visually impaired people that can help the individuals to avoid the manhole and staircase in the way of walking. The system used ultrasonic sensor for the detection of manhole and the staircase and computer vision-based technique to detect the hindrances. The SURF algorithm extracts the features from the real time images and Bivariate Gaussian Mixture Model (BGMM) detects the staircase and manhole. The main components used in the scheme are Arduino Kit, ultrasonic sensor, vibrator, HC-05 Bluetooth and camera. The proposed device is small and lightweight and will be effective for the blinds and deaf people as the system will give notification about the current position and give direction to them about the stairs. However, the system may faultier in the dark or in a place where there is noise. The system fails to categorize the moving obstacles in case of staircase detection.

III. MAJOR HARDWARE COMPONENTS

The major hardware components which are used to analyze the performances of raspberry pi and Arduino are outlined as follows.

A. Raspberry Pi

Raspberry pi is mini sized single board computer that is used to conduct any kinds of real time research. It has 17 GPIO pins that can easily connect with electronic devices (LEDs, switches etc.). The raspberry pi B+ model contains four USB ports, an audio jack port, two ports for camera interfacing, a HDMI port and an Ethernet port. The logic level of raspberry pi is approximately 3.3v. It is compatible with operating system Raspbian. It supports Python, Java, C, and C++ as programming languages and has a 700MHz single core processor.

B. Arduino UNO

Arduino is referred as a platform where software and hardware both can be embedded through Arduino board. It is compatible with operating systems like Windows, Linux, Mac OS and built with microcontrollers as well as microcontroller kits. Arduino can detect the set of instructions and perform the tasks through these microcontrollers. Nowadays, it is serving as a brain of thousands of projects. It consists of digital pins, analog pins, power pins.

C. Ultrasonic Sensor

Several sensors are used to construct different types of assistants to provide various services. Among the different sensors, ultrasonic sensors are widely used. The sensor can measure the distance of the obstacles in front of it in a certain range. The range can be 3cm to 3m and can cover up to 15 degrees. It consists of only 4 pins. The pins are Vcc, trigger (transmitter), echo (receiver) and ground. VCC pin is used to power-up the sensor. It accepts up-to +5V and ground pin is connected to ground. It sends ultrasonic pulse continuously through trigger pin. When the signal is reflected by any obstacles the Echo pin receives the signal and a control circuit measures the distance of the obstacle.

IV. SYSTEM DESIGN

In the study, an ultrasonic sonar sensor is used for obstacle detection to assist the visually impaired. The distance from the obstacle to the individuals is determined. The ultrasonic sonar sensor sends a sound wave at a specific frequency towards the obstacle and receives it back when it is reflected

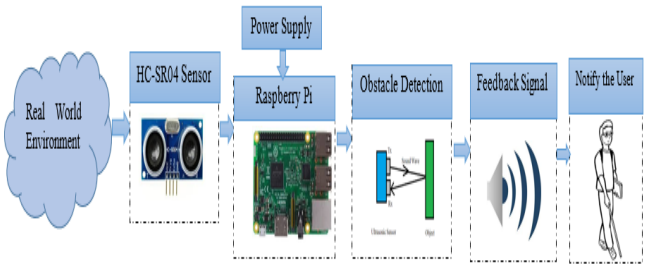


Fig. 1. A typical example of developed walking assistance tool (using raspberry pi).

by the obstacle. It archives the travel time of sound wave for forth and back. By using the travel time, sensor calculates the distance between the user and the obstacle.

$$Distance = (speed\ of\ sound \times elapsed\ time)\ /2 \quad (1)$$

In the study, the real-world data will be collected using ultrasonic sensors. The collected data will be processed using raspberry pi and Arduino for obstacle detection. A typical example of developed assisting tool (using raspberry pi) is illustrated in Fig. 1. The system generates feedback signal with the presence of obstacles and send it to the users for notification.

The connections of components are mapped with the raspberry pi and Arduino which is depicted in Table I. From the Table I, it is shown that the Trigger and Echo pins of the sensor are connected to D_9 and D_{10} pin of the Arduino. In case of raspberry pi, the Trigger and Echo pins of the sensor are connected to GPIO 17 and GPIO 18 pin. The GND and VCC pins of the sensor are connected to GND and VCC pin of raspberry pi and the Arduino.

V. Experimental Results Analysis

The experiment was performed in a virtual environment to measure the performances of the raspberry pi and Arduino. We have created five regions for positioning the obstacles. The regions are as follows.

- Region 1: (0-100cm)
- Region 2: (11-20cm)
- Region 3: (21-50cm)
- Region 4: (51-100cm)
- Region 5: (101-150cm)

The obstaclaes are placed at above described regions and the sonar sensors are used to measure the distances of the obstacles.

The data that are gathered from sensor using raspberry pi and Arduino are depicted in Table II and Table III.

TABLE I. MAPPING BETWEEN RASPBERRY PI, ARDUINO AND SENSOR

Raspberry Pin	Arduino Pin	Connected Device
GPIO 17	D9	Trigger pin of sonar sensor
GPIO 18	D10	Echo pin of sonar sensor
5V	5V	Common +5v for all device
GND	GND	Common ground(-) for all device

The comparison between the actual distance and measured distance for Arduino and raspberry pi is illustrated in Fig. 2 and Fig. 3.

Fig. 2 and Fig. 3 reveal that the measured distance and actual distance is nearly identical at the beginning when the obstacles are very close to users. These representation shows the distortion of the measured distance from the actual distance. It is shown that the distortion is not severe and the measured distance is acceptable. The distortion of Arduino is comparatively higher than the raspberry pi.

TABLE II. REAL TIME DATA COLLECTED FROM ULTRASONIC SENSOR USING RASPBERRY PI

Actual Distance (cm)	Measured Distance (cm)	Error Rate (%)
10	10.11	1.13
20	19.83	0.86
50	48.44	3.12
100	98.05	1.95
150	150.89	0.59

TABLE III. REAL TIME DATA COLLECTED FROM ULTRASONIC SENSOR USING ARDUINO

Actual Distance (cm)	Measured Distance (cm)	Error Rate (%)
10	9	10
20	19	5
50	48.65	2.7
100	96.12	3.88
150	146.49	2.34

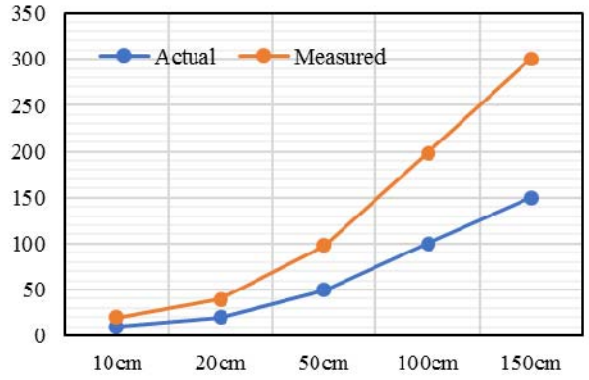


Fig. 2. Comparison with actual and measured distance in Raspberry Pi.

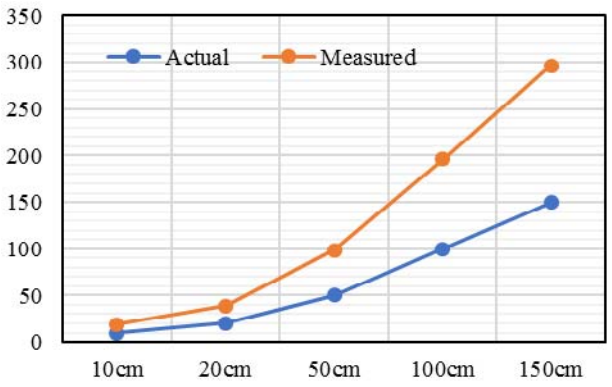


Fig. 3. Comparison with actual and measured distance in Arduino.

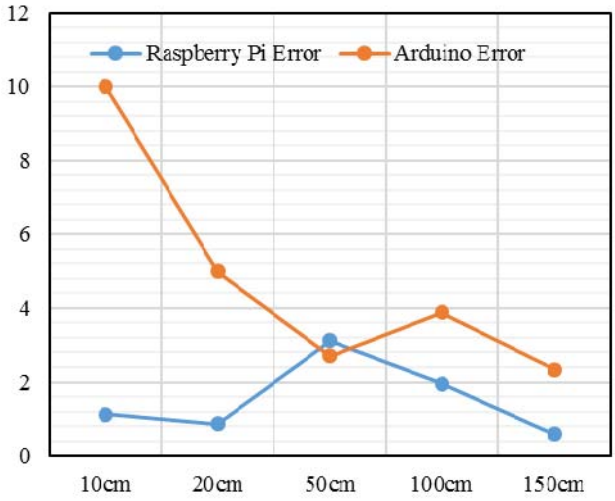


Fig. 4. Error rate for raspberry pi and Arduino for the obstacle detection using ultrasonic sensor.

The performances of obstacle detection are assessed by calculating the errors at different regions. In our study, the error represents the Mean Absolute Error (MAE). The errors at diverse regions is illustrated in Fig. 4. The error rate at region 1, region 3 and region 5 are about 1.13%, 3.12% and 0.59% respectively in case of Arduino. In addition, the error rate of 10%, 2.7% and 2.34% are achieved from region 1, region 3 and region 5 respectively. From the Fig. 4, it can be shown that raspberry pi obtained the lowest error rate that Arduino for each case except the region 3. The lower the error rate depicts the higher accuracy. The average error rate is 1.53% and 4.78% is achieved from raspberry pi and Arduino respectively. Therefore, we can conclude that the raspberry pi performs better than Arduino for the obstacle detection using ultrasonic sensor.

The performance and cost comparison of raspberry pi and Arduino are demonstrated in Table IV. From the Table IV, it can be shown that raspberry pi performs better in case of obstacle detection but the price of the raspberry pi is much higher than the Arduino.

TABLE IV. PERFORMANCE, COST AND WEIGHT COMPARISON

	Raspberry Pi	**Arduino**
Accuracy (%)	98.47	95.22
Error Rate (%)	1.53	4.78
Cost	$35	$5
Weight	368gm	400gm

TABLE V. PERFORMANCE ANALYSIS OF THE PROPOSED SYSTEM AND EXISTING SYSTEMS

Authors	**Cost**	**Weight**	**Accuracy**
Olanrewaju et al. [6]	N/M	N/M	90%
Sadi et al. [8]	$10	160gm	96%
Bhatlawande et al. [20]	$280	0.503 kg	85%
Zahir et al. [21]	$30	N/M	N/M
Proposed System	**$42**	**368gm**	**98.47%**

***Not Mentioned: N/M**

Table V illustrates the performance analysis between proposed system and existing systems. Olanrewaju et al. [6] developed iWalk that obtained accuracy of 90%. Sadi et al. [8] proposed a system that cost and weight are about $10, 160 gm respectively and achieved accuracy of 96%. Bhatlawande et al. [20] developed an EMC having the cost and weight $280, 0.503kg respectively with accuracy of 85%. Zahir et al. [21] developed a system with cost of $30. The cost and weight of the proposed system are about $42 and 368 gm respectively and obtained accuracy of 98.47% in case of raspberry pi.

VI. Conclusion

The study evaluated the performance of the walking assistance technologies developed by raspberry pi and Arduino. The system that is able to detect the obstacles accurately assists the visually impaired people in mobility. The raspberry pi based system is more perfect than an Arduino based system for obstacle detection. The raspberry pi based system achieved the accuracy of 98.47% and error rate of 1.53% whereas for Arduino based system accuracy and error rate were achieved as 95.22% and 4.78% respectively. The system is the preliminary study before developing a walking assistant for the visually impaired people. The paper suggests that raspberry pi is the most significant processing device to develop any obstacle detection system. The limitation of raspberry pi based system is that it is very expensive. In future, this work may be integrated with a stick that can support the visually impaired individuals in safe navigation.

References

[1] M. M. Islam, M. Sheikh Sadi, K. Z. Zamli and M. M. Ahmed, "Developing Walking Assistants for Visually Impaired People: A Review," in *IEEE Sensors Journal*, vol. 19, no. 8, pp. 2814-2828, Apr., 2019.

[2] A. Riazi, F. Riazi, R. Yoosfi and F. Bahmeei, "Outdoor difficulties experienced by a group of visually impaired Iranian people", in *Journal of current ophthalmology,* vol. 28, no. 2, pp. 85-90, Jun. 2016.

[3] M. A. Ikbal, F. Rahman and M. H. Kabir, "Microcontroller Based Smart Walking Stick for Visually Impaired People," *2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT)*, Dhaka, Bangladesh, 2018, pp. 255-259.

[4] K. Patil, Q. Jawadwala and F. C. Shu, "Design and Construction of Electronic Aid for Visually Impaired People," in *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 2, pp. 172-182, Apr. 2018.

[5] G.Gayathri , M . Vishnupriya , R. Nandhini and M.M Banupriya , " Smart walking stick for visually impaired," *International Journal Of Enginering And Computer Science(IJECS)*, vol.3, No.3, pp.4057-4061, 2014.

[6] R. F. Olanrewaju, M. L. A. M. Radzi and M. Rehab, "iWalk: Intelligent walking stick for visually impaired subjects," in *Proc. IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, Putrajaya, pp. 1-4, 2017.

[7] M. Wang, R. Li, R. Zhang, G. Li and D. Zhang, "A Wearable SSVEP-Based BCI System for Quadcopter Control Using Head-Mounted Device," in *IEEE Access*, vol. 6, pp. 26789-26798, 2018.

[8] M. S. Sadi, S. Mahmud, M. M. Kamal and A. I. Bayazid, " Automated walk-in assistant for the blinds," in *Proc. International Conference on Electrical Engineering and Information & Communication Technology*, Dhaka, pp. 1-4, 2014.

[9] M. M. Islam and M. S. Sadi, "Path Hole Detection to Assist the Visually Impaired People in Navigation," in *Proc. International Conference on Electrical Engineering and Information & Communication Technology*, Dhaka, in press, 2018.

[10] R. S. Mulky, S. Koganti, S. Shahi and K. Liu, "Autonomous Scooter Navigation for People with Mobility Challenges," in *Proc. IEEE International Conference on Cognitive Computing (ICCC)*, San Francisco, CA, pp. 87-90, 2018.

[11] A. G. Sareeka, K. Kirthika, M. R. Gowthame and V. Sucharitha, "pseudoEye — Mobility assistance for visually impaired using image recognition," in *Proc. 2nd International Conference on Inventive Systems and Control (ICISC)*, Coimbatore, pp. 174-178, 2018.

[12] J. Xiao, K. Ramdath, M. Losilevish, D. Sigh, A. Tsakas, "A Low Cost Outdoor Assistive Navigation System for Blind People," in *Proc. 8th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, Melbourne, Australia, pp. 828–833, 2013.

[13] S. Khade and Y. Dandawate, "Hardware Implementation of Obstacle Detection for Assisting Visually Impaired People in an Unfamiliar Environment by Using Raspberry Pi," in *Proc. International Conference on Smart Trends for Information Technology and Computer Communications,* Springer, Singapur, pp.889-895, 2016.

[14] M. Aljahdali, R. Abokhamees, A. Bensenouci, T. Brahimi and M. Bensenouci, "IoT based assistive walker device for frail & visually impaired people," in *Proc. 15th Learning and Technology Conference (L&T)*, Jeddah, pp. 171-177, 2018.

[15] N. S. Khan, S. Kundu, S. Al Ahsan, M. Sarker and M. N. Islam, "An Assistive System of Walking for Visually Impaired," in *Proc. International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, Rajshahi, pp. 1-4, 2018.

[16] H. Sharma, M. Tripathi, A. Kumar and M. S. Gaur, "Embedded Assistive Stick for Visually Impaired Persons," in *Proc. 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Bangalore, pp. 1-6, 2018.

[17] M. M. Kamal, A. I. Bayazid, M. S. Sadi, M. M. Islam and N. Hasan, "Towards developing walking assistants for the visually impaired people," in *Proc. IEEE Region 10 Humanitarian Technology Conference (R10-HTC),* Dhaka, pp. 238-241, 2017.

[18] E. E. Pissaloux, R. Velázquez and F. Maingreaud, "A New Framework for Cognitive Mobility of Visually Impaired Users in Using Tactile Device," in *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 6, pp. 1040-1051, Dec. 2017.

[19] S. Ponnada, S. Yarramalle and T.V. M. Rao, "A Hybrid Approach for Identification of Manhole and Staircase to Assist Visually Challenged," in *IEEE Access*, vol. 6, pp. 41013-41022, 2018.

[20] S. Bhatlawande, M. Mahadevappa, J. Mukherjee, M. Biswas, D. Das, and S. Gupta, "Design, development, and clinical evaluation of the electronic mobility cane for vision rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 6, pp. 1148–1159, Nov. 2014.

[21] E. Zahir, M. S. Hossain, M. W. Iqbal, I. Jalil and S. M. Kabir, "Implementing and testing an ultrasonic sensor based mobility aid for a visually impaired person," in *Proc. IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, Dhaka, pp. 453-456, 2017.

Performance Evaluation of a Wearable 2.45 GHz Planar Printed Meandering Monopole Textile Antenna on Flexible Substrates

Md. Mamun Ur Rashid[1]
Department of Electrical & Electronic Engineering
American International University-Bangladesh
Dhaka, Bangladesh
i.mamunrashid@gmail.com

Ashiqur Rahman[2]
Department of Electrical & Electronic Engineering
American International University-Bangladesh
Dhaka, Bangladesh
ashiqur@aiub.edu

Liton Chandra Paul[3]
Department of Electronic & Telecommunication Engineering
Pabna University of Science & Technology
Pabna, Bangladesh
litonpaulete@gmail.com

Ajay Krishno Sarkar[4]
Department of Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
sarkarajay139@gmail.com

***Abstract*—This paper presents a wearable planar printed meandering monopole textile antenna design having operating frequency of 2.4 – 2.45 GHz on 3 flexible substrate materials: Jeans, Leather and Cordura®. The printed elements have been designed by using a mender line technique to reduce the size of the antenna significantly than the proposed antennas of past literature. Antenna performance is analyzed for different bending conditions and Specific Absorption Rate (SAR) is evaluated when the antenna is in contact with human body using CST Microwave Studio. The compact design and miniature size of the antenna is suitable for easy integration in smart clothing for tracking, healthcare, security and other ISM applications.**

Keywords—textile antenna, wearable antenna, bending conditions, SAR

I. INTRODUCTION

Flexible textile antennas are currently under an extensive research focus to develop applications based on smart and wearable systems. Textile antennas are an emerging solution to the challenges of monitoring people's healthcare, protection and safety. To integrate textile antennas in clothing, antenna substrate need to be flexible, low profile and hidden as much as possible. For this reason, research efforts are continuing around the world on developing efficient and flexible antenna substrate materials and miniaturizing the size of the antenna. An extensive survey was conducted in the past to accumulate the key points for the design and development of textile antennas [1]. Comparison of different textile materials used in past research works for ISM and WLAN applications was made on this survey, such as: Cotton/Polyester (relative permittivity, $\varepsilon_r = 1.6$ & loss tangent, $\tan\delta = 0.02$) with Flectron as conductive material [2], Fleece Fabric ($\varepsilon_r = 1.04 – 1.25$) with Flectron and knitted Copper fabric as conductive material [3]-[5], Woolen Felt ($\varepsilon_r = 1.45$, $\tan\delta = 0.02$) with Silver-copper-nickel plated woven fabric as conductive material [6], Felt ($\varepsilon_r = 1.30$, $\tan\delta = 0.02$) with Zelt as conductive material [7] and Cordura® ($\varepsilon_r = 1.1 – 1.7$) with Copper tape as conductive material [8]. All of the antennas analyzed in this survey was microstrip patch antennas which have very large antenna size. S. Zhu and R. Langley also have conducted a detailed antenna substrate material survey before choosing Felt ($\varepsilon_r = 1.3$, $\tan\delta = 0.023$) over Silk, Tween, Panama, Moleskin and Perspex to design a dual-band coplanar patch antenna integrated with an electromagnetic band gap substrate having Zelt as the conducting material for WLAN applications [7]. Modern wearable on-body communication system requires more efficient small sized antennas for smart clothing and other ISM applications.

Very few research works have been conducted on the development of textile antenna using Leather and Jeans substrate. M. I. Ahmed and et al. have conducted an investigation on the characteristics of 3 different textile materials: Goch, Jeans and Leather [9]. This work measured the dielectric constant, ε_r and loss tangent, $\tan\delta$ of the 3 substrates (Goch, Jeans & Leather) using 2 different methods: (a) using a microstrip ring resonator method and (b) using Dielectric Assessment Kit (DAK). The proposed antennas are averagely 30mm x 45mm in dimension. Fabrication techniques of textile antennas such as: Copper foil, Copper mesh, Conductive spray and Copper with rubber were analyzed by J. G. Santas and et al. [10]. This paper presented 3 textile antennas (60mm x 80mm) designed on Leather substrate applying, (a) Copper tape, (b) Woven Copper thread and (c) Conductive spray. Another textile antenna on Cotton Jeans substrate was proposed in past literature [11]. However, these antennas are very large in size to consider for modern day wearable on-body communication system to integrate in smart clothing. Miniaturization of textile antenna is now a necessity and requires more research efforts. Additionally, these works lack sufficient experimental or simulation data of antenna performance under bending conditions and Specific Absorption Rate (SAR) evaluation in the presence of human body.

In this paper, we are presenting a 2.4 – 2.45 GHz Planar Printed Meandering Monopole Textile Antenna designed using 3 substrates: Jeans, Leather and Cordura® to meet the challenges of wearable on-body communication systems. The meandering monopole design of the antenna reduces the antenna size significantly and can be easily integrated in smart clothing for monitoring, ensuring security, tracking and healthcare. The antenna patch and ground plane is printed on both sides of the substrate ensuring efficient antenna performance. The antenna was tested under bending conditions and SAR was evaluated in the presence of human tissue box model using CST Microwave Studio 2018. Besides Jeans and Leather, the antenna was also designed on Cordura® Fabric substrate as it is a high-tech durable Japanese fabric made for producing bags and backpacks [12].

II. METHODOLOGY

The proposed 2.4 – 2.45 GHz Textile Antennas were designed on 3 substrates: Jeans, Leather [9] and Condura® [1]. The dielectric properties characterization of the substrates is summarized in Table 1.

TABLE I. DIELECTRIC CHARACTERIZATION OF THE SUBSTRATES

Substrate Material	Dielectric Constant (ε_r)	Loss Tangent (tanδ)	Substrate Height (mm)
Jeans	1.78	0.085	0.6
Leather	1.79	0.042	1.1
Cordura®	1.1	-	0.5

The patch elements and ground plane are made out of annealed copper (Cu). The antenna patch and ground plane are printed on both sides of the substrate and are mirrored to each other. The antenna patch and ground planes of both sides are connected to each other through a via made out of knitted copper wire. Meander line technique was used to design the printed elements of the antenna for the reduction of the size of the antenna. There are two ways to reduce the resonant frequency of a monopole by meandering the monopole element.. First one is by increasing the width of the meander line sections. Therefore the increment of the capacitance to ground reduces the resonant frequency. In order to maintain the overall length of the antenna, reduction of meander section numbers also reduces the effective self-capacitance of the monopole element. This phenomenon concurrently increases the effective self-inductance which leads to the reduction of resonant frequency [13]. Secondly, resonant frequency can be reduced by increasing the total element length while maintaining the same physical volume. The increment in total element length also increases the effective self-inductance which eventually reduces the resonant frequency [13]. Through parametric sweep optimization process using CST Microwave Studio the resonance frequency at 2.4 – 2.45 GHz was obtained. Figure 1 presents the design of both front and back side of the antenna and table 2 summarizes the optimized parameter values for each antenna element.

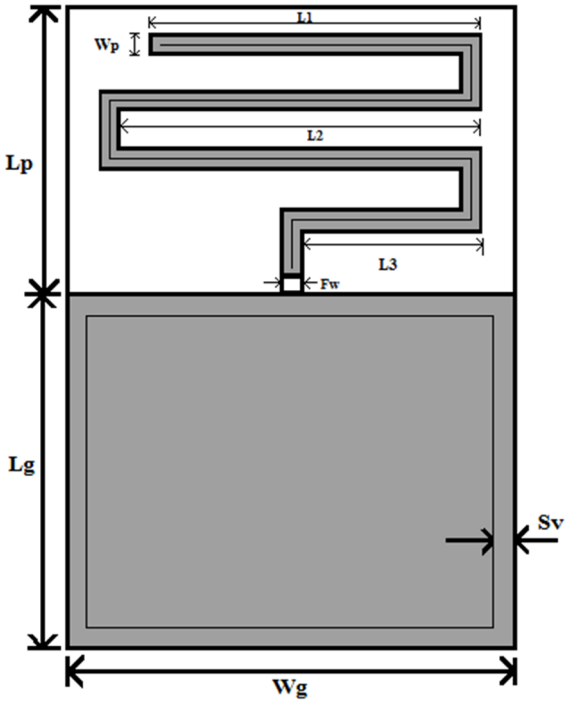
TABLE II. DESIGN PARAMETERS OF ANTENNA FOR JEANS, LEATHER & CORDURA® SUBSTRATES

Design Parameters (mm)	Substrate Materials		
	Jeans	Leather	Cordura®
Lg	23.54	23.84	23.18
Wg	16.07	16.27	15.82
Lp	11.96	12.11	11.78
Wp	0.74	0.76	0.75
Fw	0.22	0.22	0.22
Sf	0.37	0.38	0.37
S1	2.99	3.03	2.94
S2	1.12	1.14	1.10
Sv	0.37	0.38	0.37
L1	6.73	6.81	6.62
L2	13.45	13.62	13.25
L3	13.27	13.43	13.06

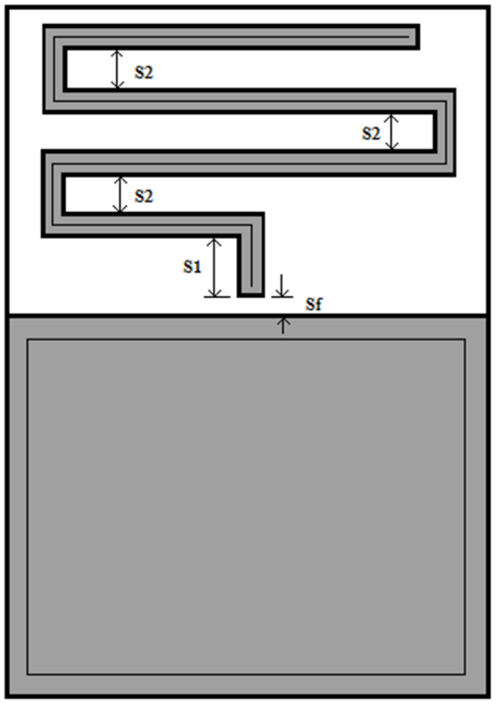
The designed 2.4 – 2.45 GHz textile antennas are fed with respect to the upper ground plane using discrete port feeding technique. CST Microwave Studio 2018 was used to analyze and measure the performance of all 3 antennas. The antenna has a linearly polarized omnidirectional pattern like a conventional dipole. The antenna has a low input resistance ≈ 20 Ω due the reduction in size by meandering the monopole [14]. The performance of the antenna under bending conditions and when placed on the human body are evaluated in terms of SAR.

III. ANTENNA DESIGN & SIMULATION

3 Planar Printed Meandering Monopole Antenna of 2.4 – 2.45 GHz operating frequency are designed in CST Microwave Studio's simulation environment on 3 substrates: Jeans, Leather and Cordura®. Discrete port feeding technique was used during the simulation and port input impedance was 20Ω. The designed antennas in CST Microwave Studio are presented in figure 2.

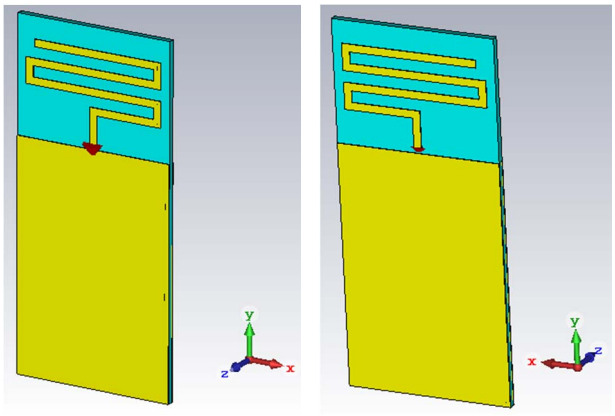


(a) Front Side of the antenna



(b) Back side of the antenna

Fig 1. Dimensions of the antenna patch and ground plane

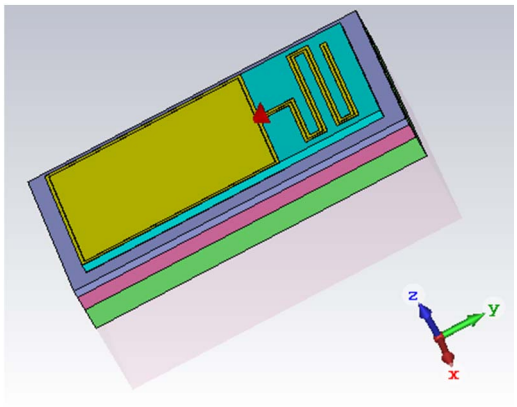


(a) Front (b) Back

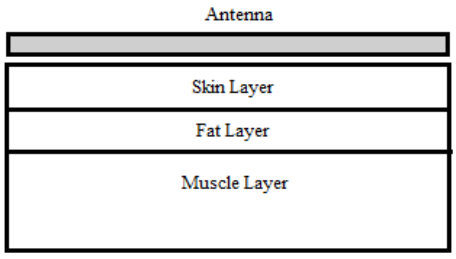
Fig 2. Antenna Design (Jeans Substrate) in CST Microwave Studio

IV. Human Tissue Model & Tissue Properties

The 2.4 – 2.45 GHz antennas were simulated over a human tissue model consisting of 3 layers of tissue: Skin, Fat and Muscle which is shown in figure 3. The permittivity, electrical conductivity, density and dimensions of all 3 layers of the tissue were considered from the Tissue Properties Database of the Foundation for Research on Information Technologies in Society (IT'IS Foundation) [15]. The tissue properties of all 3 considered tissue layers at 2.4 GHz are summarized in table 3.



(a) Antenna over human tissue box model in simulation environment



(b) Format of the human tissue box model with Antenna

Fig 3. Setup for SAR evaluation of the antenna using a human tissue box model

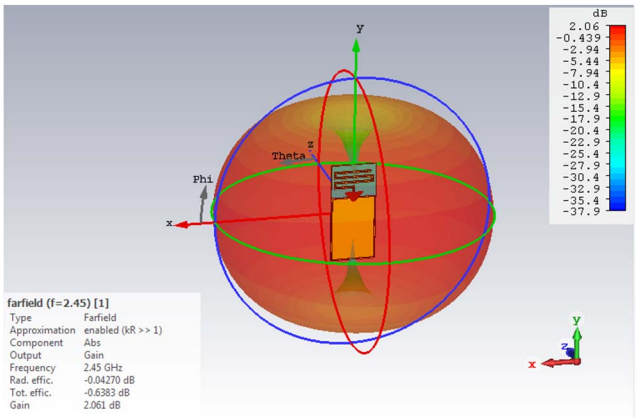
TABLE III. Tissue Properties at 2.4 GHz

Tissue Name	Skin	Fat	Muscle
Permittivity	3.80E+1	1.08E+1	5.27E+1
Electrical Conductivity (S/m)	1.46	2.68E-1	1.74
Density (kg/m^3)	1109	911	1090
Dimension (mm)	20 x 40 x 2	20 x 40 x 4	20 x 40 x 15

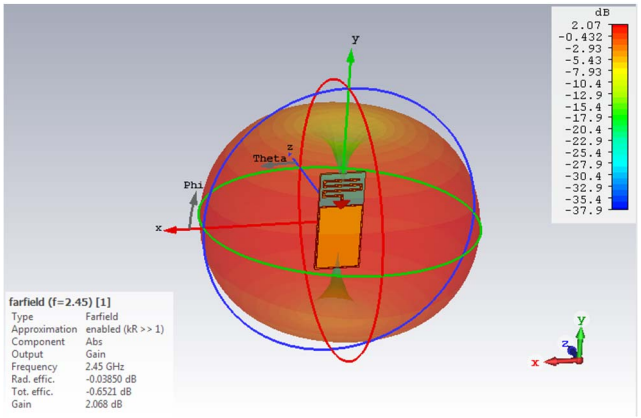
V. Results & Discussion

A. Antenna Performance Evaluation

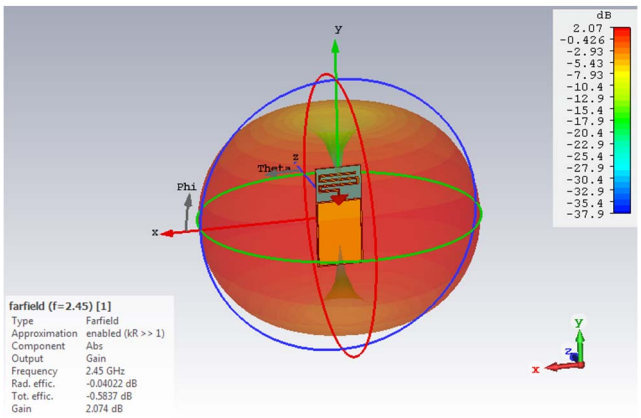
All 3 antennas made out of Jeans, Leather & Cordura® substrate are simulated using CST Microwave Studio. The antennas exhibited uniform and linearly polarized omnidirectional radiation pattern presented in figure 4. The simulated directivity, gain (IEEE), E-Field (Emax), H-Field (Hmax), power flow, surface current and Voltage Standing Wave Ratio (VSWR) of the antennas are summarized in table 4. The reflections coefficients S11 of all 3 antennas are presented in figure 5. The center frequencies of the antennas are between 2.4 – 2.45 GHz and average bandwidth of the antennas were 0.18 GHz ranging from 2.36 GHz to 2.54 GHz.



(a) Jeans



(b) Leather



(c) Cordura®

Fig 4. Linearly Polarized Omnidirectional Radiation Pattern of the Antennas

TABLE IV. SIMULATED DATA OBTAINED FROM THE SIMULATION OF 3 TEXTILE ANTENNAS OF 2.45 GHZ: JEANS, LEATHER & CORDURA®

	Jeans	Leather	Cordura®
Directivity (dBi)	2.104	2.106	2.114
Gain (IEEE) (dB)	2.061	2.068	2.074
E-Field (Emax) (dBV/m)	16.23	16.22	16.30
H-Field (Hmax) (dBA/m)	-35.29	-35.30	-35.22
Power Flow (Pmax) (dBW/m^2)	-12.54	-12.55	-12.47
Surface Current (A/m)	211.623	177.716	206.958
VSWR	1.159	1.171	1.215

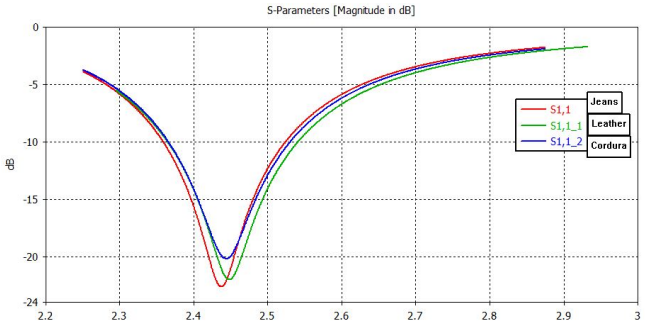
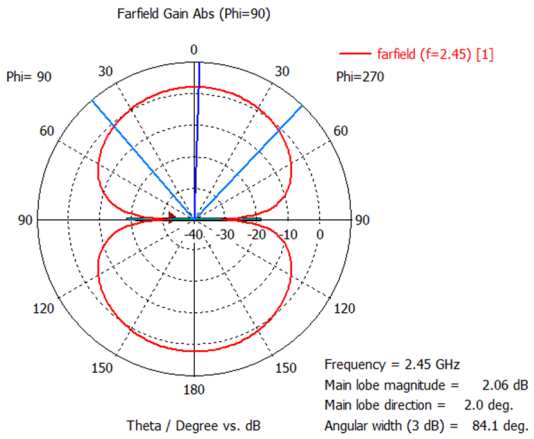
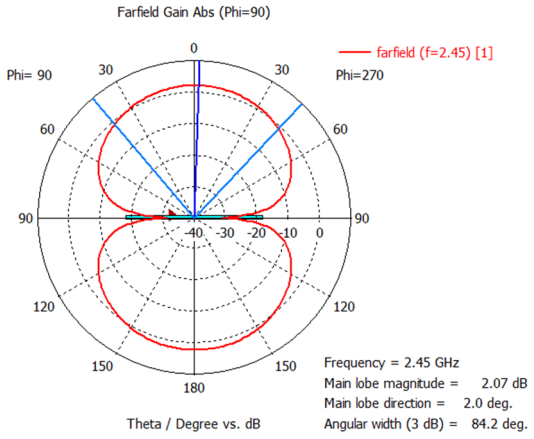


Fig 5. Reflection Coefficients (S11) of all 3 antennas: Jeans, Leather & Cordura®

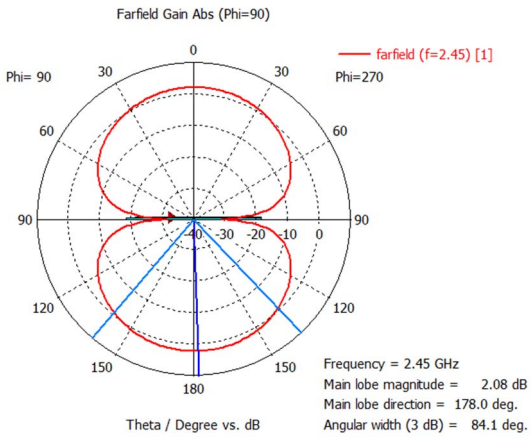
The farfield gain plots of the 3 antennas demonstrates the impedance matching obtained in figure 6. The input impedance of the antennas was 20Ω.



(a) Jeans



(b) Leather

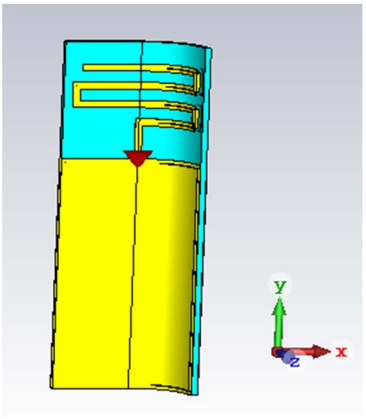


(c) Cordura®

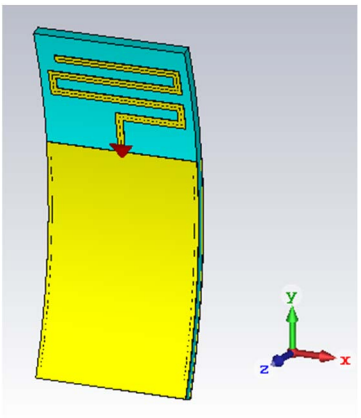
Fig 6. Gain Plot of all three 2.45 GHz textile antennas

B. *Performance Evaluation Under Bending Conditions*

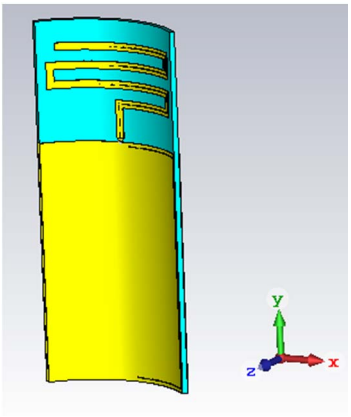
To Evaluate the antennas under bending conditions different types of bending were tested on the antennas. For the purpose of this investigation Jeans substrate antenna were single sided bended at an angle of 45 degrees, the Leather substrate antenna was bended horizontally along a 50mm radius cylinder and the Cordura® substrate antenna was bended vertically along a 50mm radius cylinder. The bended textile antennas are presented in figure 7. After the bending the center frequency of the antennas shifted by 0.02 GHz which is very negligible (figure 8) and the other performance parameters remained the same. The results signify the durability of the textile antennas.



(a) Jeans

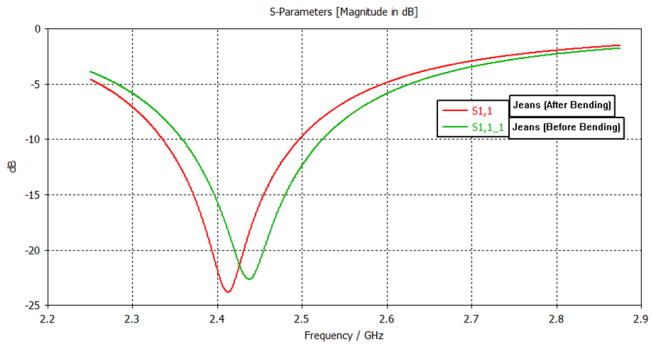


(b) Leather

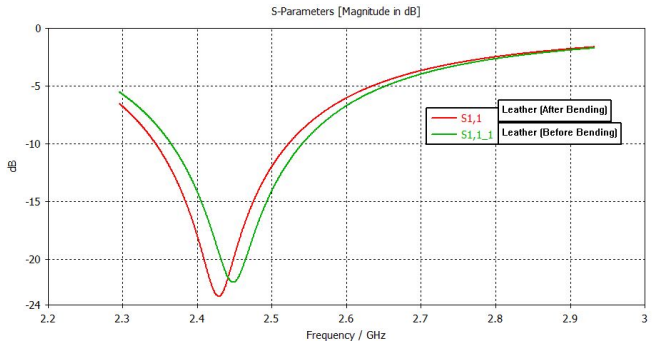


(c) Cordura®

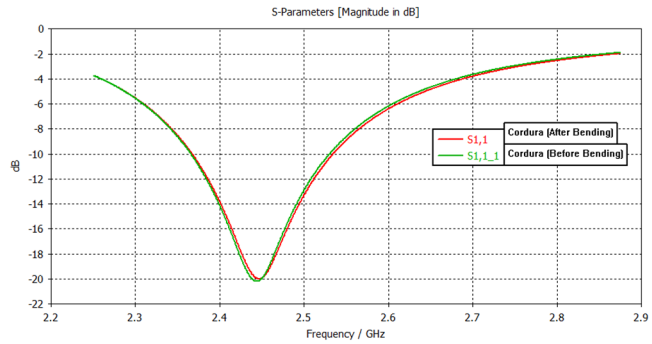
Fig 7. Performance Evaluation of Antenna Under Bending Conditions



(a) Jeans



(b) Leather



(c) Cordura®

Fig 8. Center Frequency Shifting After Bending of the Antenna Substrates

C. SAR Evaluation

The 2.4 – 2.45 GHz planar printed meandering monopole textile antennas designed using Jeans, Leather and Cordura® substrates are simulated over a human tissue box. The SAR distribution is presented in figure 9.

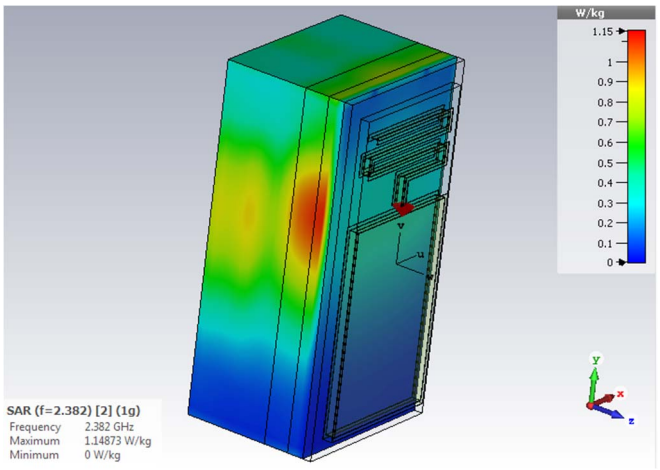


Fig 9. SAR Distribution

The simulated maximum SAR was 1.14873 W/kg for 1g of tissue which satisfies the IEEE and ICNIRP standards of SAR of 1g of tissue which is 1.6 W/kg [16].

VI. Conclusion

This paper presents three wearable 2.4 – 2.45 GHz planar printed meandering monopole textile antennas designed using Jeans, Leather and Cordura® substrate. Meanderline design technique can be utilized to reduce the size of the antenna maintaining efficient performance and durability. Modern era research is concentrated on technologies based on micro and nano scale. Miniaturization of textile antennas is needed to support the evolution of wearable smart on-body communication systems. The proposed antennas of previous literature proposed very large microstrip patch based designed antennas which are incompatible with smart clothing. The proposed antenna dimension in this paper is only 16 mm x 35 mm and the antenna exhibits satisfying performance under different bending conditions and in contact with human body. Further research can be done by fabricating the antenna and testing the performance in practical environment.

References

[1] Salvado, R., Loss, C., Gonçalves, R., & Pinho, P., "Textile materials for the design of wearable antennas: a survey." in Sensors (Basel, Switzerland), 12(11), 15841-57, 2012.

[2] M. L. Scarpello, I. Kazani, C. Hertleer, H. Rogier and D. Vande Ginste, "Stability and Efficiency of Screen-Printed Wearable and Washable Antennas," in IEEE Antennas and Wireless Propagation Letters, vol. 11, pp. 838-841, 2012.

[3] C. Hertleer, H. Rogier, L. Vallozzi and L. Van Langenhove, "A Textile Antenna for Off-Body Communication Integrated Into Protective Clothing for Firefighters," in IEEE Transactions on Antennas and Propagation, vol. 57, no. 4, pp. 919-925, April 2009.

[4] C. Hertleer, H. Rogier, L. Vallozzi and F. Declercq, "A Textile Antenna based on High-Performance Fabrics," The Second European Conference on Antennas and Propagation, EuCAP 2007, Edinburgh, 2007, pp. 1-5.

[5] P. Salonen and L. Hurme, "A novel fabric WLAN antenna for wearable applications," IEEE Antennas and Propagation Society International Symposium. Digest. Held in conjunction with: USNC/CNC/URSI North American Radio Sci. Meeting (Cat. No.03CH37450), Columbus, OH, 2003, pp. 700-703 vol.2.

[6] I. Locher, M. Klemm, T. Kirstein and G. Trster, "Design and Characterization of Purely Textile Patch Antennas," in IEEE Transactions on Advanced Packaging, vol. 29, no. 4, pp. 777-788, Nov. 2006.

[7] S. Zhu and R. Langley, "Dual-Band Wearable Textile Antenna on an EBG Substrate," in IEEE Transactions on Antennas and Propagation, vol. 57, no. 4, pp. 926-935, April 2009.

[8] P. Salonen, Y. Rahmat-Samii, M. Schaffrath and M. Kivikoski, "Effect of textile materials on wearable antenna performance: a case study of GPS antennas," IEEE Antennas and Propagation Society Symposium, 2004., Monterey, CA, USA, 2004, pp. 459-462 Vol.1.

[9] M. I. Ahmed, M. F. Ahmed, A. A. Shaalan, "Investigation and Comparison of 2.4 GHz Wearable Antennas on Three Textile Substrates and Its Performance Characteristics", in Open Journal of Antennas and Propagation, vol. 5, no. 3, pp. 110-120, September 2017.

[10] J. G. Santas, A. Alomainy and Y. Hao, "Textile Antennas for On-Body Communications: Techniques and Properties," The Second European Conference on Antennas and Propagation, EuCAP 2007, Edinburgh, 2007, pp. 1-4.

[11] Bivas Roy, Partha Pratim Sarkar & Santosh Chowdhury, "Wideband rectangular wearable jeans antenna", in Microwave and Optical Technology Letters, vol. 55, no. 6, pp. 1270-1273, 27 March 2013.

[12] Cordura.com. (2018). CORDURA® - Welcome. [online] Available at: https://www.cordura.com/en [Accessed 15 Nov. 2018].

[13] S. R. Best, "A discussion on the properties of electrically small self-resonant wire antennas," IEEE Antennas and Propagation Magazine, vol. 46, no. 6, pp. 9-22, Dec. 2004.

[14] R. Wallace, "Design note DN024 Monopole PCB antenna with single or dual band option,"[Online]. Available:

http://www.ti.com/lit/an/swra227e/swra227e.pdf. Accessed: Nov. 1, 2018.

[15] C. Gabriel. Compilation of the Dielectric Properties of Body Tissues at RF and Microwave Frequencies, Report N.AL/OE-TR- 1996-0037, Occupational and environmental health directorate, Radiofrequency Radiation Division, Brooks Air Force Base, Texas (USA), 1996.

[16] P. Stavroulakis, Biological effects of electromagnetic fields, 2003, Springer.

Intelligent Path-finder for the blind

A. Z. M. Tahmidul Kabir
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
tahmidulkabir@gmail.com

Nirmol Deb Nath
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
debnathnirmol@gmail.com

Mohitosh Pramanik
Department of Computer Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
mohitoshpm@gmail.com

Tawsif Ibne Alam
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
tawsif.alam@aiub.edu

***Abstract*— Eyesight is one of the most useful in living organism, visually impaired people do not consist of that sense. They are unable to see the beauty of nature except by using the imagination, facing several problems in their daily life continuously in executing certain movements such as walking, feeling something properly, seeing the world and the surrounding people. All the problems faced by the visually impaired people cannot be solved totally, as in some cases there is no option to give their eyesight back. This project, "Intelligent Path-Finder" will tend to make life easier and less problematic. Visually impaired people use general stick for their movement but the sticks are not capable of detecting the obstacles or other objects in front of them automatically; as a result, they have to poke the obstacles to understand the type. To overcome these problems, the "Intelligent Path-Finder" is a remarkable device which can detect various types of obstacles from knee level to ground level, for example liquid, uneven surfaces. In the dark, a special feature is used to prevent any accident by glowing led light by sensing darkness that can protect the user from moving vehicles and some other difficulties by warning normal people. If the user is in trouble, he can send location by pressing an emergency switch, and his family can also know his location by sending SMS. This device consists of a microcontroller, Ultrasonic sensor, Light dependent resistor etc. The combination of all this equipment together will be making this device relatively accurate, user-friendly, robust and cheap. The device is not a third eye for the blind but it is really helpful to them by making their daily life easier than the previous one.**

***Keywords*— Ultrasonic sensor, Light dependent resistor (LDR), Global positioning system (GPS) and Global System for Mobile communications (GSM), Blind people.**

I. INTRODUCTION

Visual impairment is known as partial or complete loss of sight. Blindness may be caused by injury, by lesions of the brain or optic nerve, by disease of the cornea or retina, or by pathological changes originating in systemic disorders [1]. There are a lot of visually impaired people all over the world. According to the report of World Health Organization (WHO) on October 11, 2018, it is estimated that there are 1.3 billion people who live with some form of visual impairment [2]. There are 188.5 million people with mild vision impairment, 217 million people with moderate to severe vision impairment and 36 million people are blind [2]. 82% of the totally blind people are aged fifty or more [3].

According, to The International Agency for the Prevention of Blindness (IAPB), the total number of blind people in 2015 is 253 million [4]. From, this huge number of people, 89% of visually impaired live in low and middle-income countries [4].

Blind people are helpless in their daily life, in terms of mobility. They are not capable of going anywhere without the help of other people or devices. Sometimes, they take help from trained dogs, but they are very costly. As a result, blind people use ordinary sticks to walk from one place to another by touching the obstacles or poking them. It is also unsafe for them to walk with the ordinary stick because in most cases, they do not understand the type of object and the distance of the object from them. There is also possibility to collide with obstacles and in most cases, causing an accident, with severe injury and sometimes even causes death. In the circumstances stated above, there are many devices, available in the market for assisting visually impaired people, but the cost, accuracy and reliability of those devices are still up for debate.

To overcome the situation, this study led to a prototype of an "Intelligent Path-Finder" with high efficiency and user-friendly features in low budget. The device uses properties of sound to detect obstacles, especially uneven surfaces

and liquid, with high reliability. LDR is integrated in this device to detect the darkness. GPS and GSM are also integrated to track the location continuously and for sending data to a phone number. The user can choose not to share his/her location by simply turning off the tracker if he/she wishes.

II. Integrated Equipment

The prototype has been assembled using various types of components are integrated together. In this section, integrated components are described below in details:

TABLE I IMPORTANT FEATURES OF THESE COMPONENTS USED IN THIS PROJECT

No. of component	No. of unit	Important features for this project
Arduino Mega	1	• Main processing unit. • 30 Digital pins used out of 54 pins. • 1 Analog pin used out of 16 pins. • 5V is used to operate all of the components. • To power up the Arduino 9V is used in this project.
Ultrasonic sensor	5	• To measure the distance. • It can measure 0 to 13 feet. • To detect obstacles and uneven surface
Light-dependent resistor	1	• To detect light intensity. • To lighten up the stick in the dark period.
SIM808	1	• Detect user location. • Send SMS to user family in an emergency situation.
Vibration motor	4	• To inform the position of the obstacles using vibration.
Buzzer	1	• To inform the user by alarming if any kinds of liquids are found.
LED	10	• To concern the vehicle or normal people in a dark period.

[a] Important features of each component

III. System Description

The working principle of the system has been described with the help of system level block diagrams, to minimize the effort in understanding complex circuitry. Some simulations have also been added before final implementation of the project was done.

A. *Block diagram*

This diagram shows the input-output relationship with the main processing unit and the source. The diagram is given below:

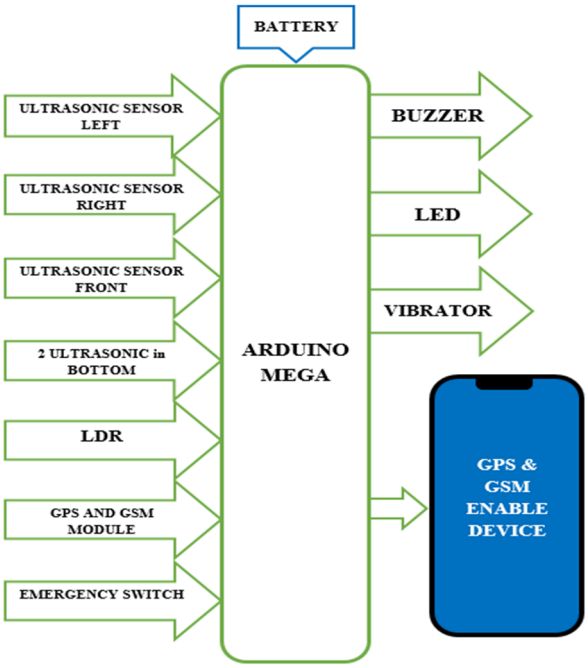


Fig.1. Block diagram of the system.

On the left side of Fig.1, inputs of the device are shown connected with the main processing unit. There are total 5 ultrasonic sensors used in this system. These sensors sense the obstacles around the user and then send signals to the microcontroller to process them. LDR is added in the left side as an input which detects the darkness of the surrounding environment and sends signals to the microcontroller. GPS-GSM module is added as an input of the device. GPS works for locating the position of the user and GSM works for sending the message to the accessed phone but both of them send signals to microcontroller for the further process. An emergency switch has been added as an input of the system which sends a signal to the microcontroller for sending a text whenever the user wants. A battery is integrated in this device as the power source. Buzzer has been added as an output of the system which provides a beep when liquid is detected in front of the user. A total of 4 vibration motors have been integrated in this device as outputs. These motors vibrate when obstacles are detected around the user and also when uneven surface is detected in the walking path of the user.

B. *Flow diagram*

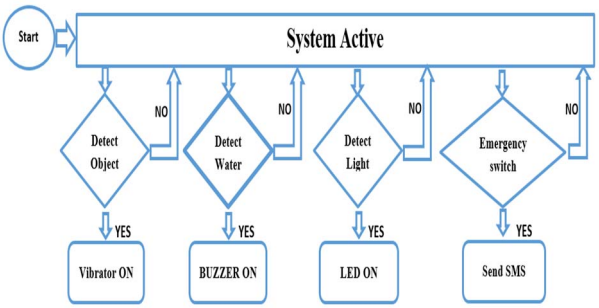


Fig.2. Flow diagram of the system.

Fig. 2 represents the flow diagram of the system. After activating the system, HC-SR04 ultrasonic sensor starts to detect obstacles in front of the device as shown in fig. 4. The sensor works with the simple formula that is given below [5]:

Distance = Speed × Time [5].

The Ultrasonic transmitter transmits an ultrasonic wave, this wave travels in air and when it gets objected by any material it gets reflected back toward the sensor and this reflected wave is observed by the Ultrasonic receiver module as shown in the picture below [5].

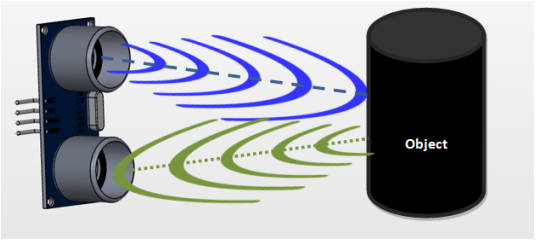


Fig.3. Obstacle detection by an Ultrasonic sensor [5].

Now, to calculate the distance using the above formulae, we should know the speed and time [5]. Since we are using the Ultrasonic wave, we know the universal speed of US wave at room conditions which is 330m/s [5]. The circuitry inbuilt on the module will calculate the time taken for the US wave to come back and turns on the echo pin high for that same particular amount of time, this way we can also know the time taken [5]. Now, using the Arduino Mega (ATmega2560) distance can be calculated simply [5]. The device can detect the liquid in its path and this system is implemented by a very simple method. Two wires are connected in this device in such a way that these are open in the bottom of the device. If any liquid is present bottom of the device then the wires become shorted and current starts flowing through the device. This will send a signal to the Arduino Mega (ATmega2560) and after processing the signal a beep is heard from the buzzer as an output of the water detection system. Darkness detection system is also introduced in this device. Light Dependent Resistor (LDR) is used as the device which can detect the darkness. A graph of fluctuating resistance of LDR with the intensity of the light is shown in fig. 4.

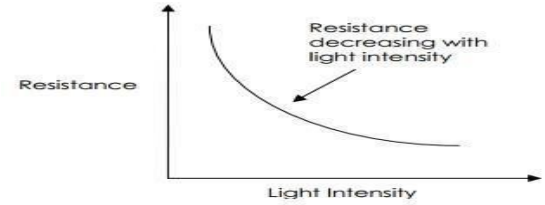


Fig.4. relation between light intensity and resistance [6].

The most common type of LDR has a resistance that falls with an increase in the light intensity falling upon the device (as shown in the image above) [6]. The resistance of an LDR may typically have the following resistances [6]:

Daylight = 5000Ω

Dark = 20000000Ω

LDR is used in such a way that when light intensity is lower, all the LED of the device are turned on and when light intensity is higher all the LED are turned off. Another special feature of tracking the location and sending a message to an accessed phone number is also introduced in this device. GPS system continuously tracks the location of the user and when the user wants, he can send the text with his location by pressing the emergency switch and GSM module works for sending the text to the accessed phone number. This system also works in another way, when someone wants to know the location of the blind people from any distance, they have to send a text to the number which is inserted in the device by writing the word "Location". After receiving the text with the word "Location" the device will give a reply automatically to the sender number with the location of the blind people

C. *Simulation*

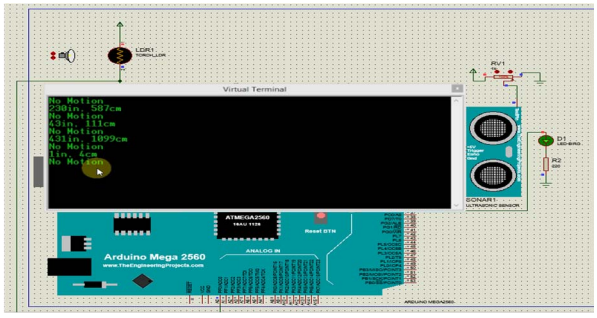


Fig.5. simulation of the system

Fig. 5 shows the simulation of the system. The figure shows that there are two sensors in this simulation, one of them is Ultrasonic Sensor and another one is LDR. The simulation was done in Proteus.

D. *Overview*

The image of the device is given below:



Fig.6. Design of the device.

Fig.6 shows the full hardware parts integrated with the device. In this figure, it is seen that there are 4 vibration motor on the top side of the device. The first vibration motor is on the front side, the second one is on the left side, the third one is on the right side and the last one is situated at a place where the user will keep his palm. A circuit box is situated in the middle portion of the device. Three ultrasonic sensors are added under the circuit box and another two ultrasonic sensors are added in the bottom side of the device.

E. *Power consumption:*

The overall power consumption of the device is shown below:

TABLE II POWER CONSUMPTION PER DAY

Component	Power Consumption Per Day
Arduino Mega	170mA*60s*60*24h =14688000mA =14688A
SIM 808(GPS & GSM)	1mA*60*60*24h =86400mA =86.4A
HC – SR04 sensor	5*15mA*60*60*24h =6480000mA =6480A
LDR sensor	0.5mA*60*60*24h =43200mA =43.2A
Vibration motor	4*17mA*60*60*24h =5875200mA =5875.2A
Buzzer	15mA*60*60*24h =1296000mA =1296A
LED	2*20mA*60*60*24h =3456000mA =3456A
TOTAL (A)	**31924.8A**
TOTAL (mA)	**31924800mA**
SOURCE POWER 5V,10000mAh	**36000000mA**

[b] Total power needs to operate this device

All the components assembled in the device has a supply voltage rating of 5V and the power source used for this device a 5V-10000mAh standalone battery pack. The power consumption table has been completed assuming all the equipment are running at the same time, continuously and after calculating the values it appeared that the device can run for 24 hours, efficiently.

F. The outlay of the device:

The expense of all the equipment used in the device is shown table-III, given below:

TABLE III COST CALCULATION

COMPONENT	Quantity	Price (USD)
Arduino Mega	1	$8.95
HC-SR04	5	$5
LDR & Buzzer	1 & 1	$0.2
Vibration motor	4	$4
LED & resistor	2 & 2	$0.15
GPS & GSM (SIM 808)	1	$19.08
Switch	1	$0.06
Syb170 mini bread board	1	$0.4
Battery	1	$12.16
TOTAL PRICE		**$50**

[c] Price list of the components

The total cost of the device is only $50. Compared to the features and usefulness of this device, the cost is very low and the device is easily affordable.

IV. FEATURES AND COMPARATIVE ADVANTAGES:

Features available in the device:

- Detects any kind of obstacles in its path up to 180°.
- The device detects liquid.
- Uneven surface detection is another special feature of this device.
- Location tracking and sending a text with the user location in two ways. When the user needs to share his location or the family members needs to know the location of the user.
- This device also detects the darkness to make others known the user.
- Highly effective, much accurate, user-friendly and affordable.

Comparative advantages of the device:

The cost of this device is lower than many devices available in the market having the same features. The name of the device which was introduced before was An Assistive System of Walking for Visually Impaired [7]. This device offered $30 cost but there were many lacking accuracies about measuring distance and there was no any system of detection uneven surface and liquid, as well as there, was no any GPS-GSM module to track the location and to send the data. Another device named Radio Frequency
Identification Walking Stick (RFIWS): A Device for the Blind [8] used RFID technology to make their device but there is no liquid detection system and location tracking system. Overall, they proposed a good solution but the user was biased to follow a specific path whereas using our device user is independent of which path to follow. Also, their proposed solution is costlier because cost increases along with path length but our device having no dependencies to which path to follow and path length, only the device cost is to be taken into account. Making it a lot cheaper in comparison. The other paper proposed a device named An Electronic Walking Stick for Blinds [9]. They used the GSM module to send a text with a warning but there was no location tracking system and the GPS-GSM module could be used together to track the location and to send a text with the exact location. Another group proposed a device-SWSVIP–Smart Walking Stick for the Visually Impaired People using Low Latency Communication [10]. They used the GPS module and mobile application to send the data via Bluetooth but the problem is that a Bluetooth enabled phone is always needed with user of the stick. The system actually made the device complex but, in our device, there is no need of any mobile phone to send the data to another phone number and this system makes the device simple and easily affordable. Another group introduced a system named Blindar: An Invisible Eye for the Blind People [11]. They proposed a system which was a combination of ultrasonic sensor, MQ2 gas sensor, Wi-Fi and GPS module but they also did not introduce any liquid detection system and they proposed the GPS system and Wi-Fi system which is

unreliable because Wi-Fi is not available everywhere especially in street to upload the data to the selected website. To overcome the drawbacks of the proposed devices earlier we have introduced the “Intelligent Path-finder” for the blind people. One of the main advantages of this device is that the location tracking in two ways. Sending a text to an accessed family member is available when the user wants at any time and the other way to get the location of the user is sending a text from any phone number from anywhere with text having a word “Location”. Liquid detection with 100% accuracy without any sensor has been achieved in this prototype, making the device less costly and this system was not introduced in any other devices earlier. Also, special advantage of this device is to make the user visible in dark and this feature also was not introduced before in any device.

V. CONCLUSION

We are living in a modern, civilized and a very developed world, constantly trying to make the life better. Though the people of the modern world are trying to make this living place more beautiful and peaceful day by day, there are many subjects which are not updated according to the developed world and one of those subjects is the cure for the visually impaired. The attempt of trying to do something for them and, as a result, many autonomous devices are introduced earlier for the visually challenged. It seems that those devices are either out of range for the blind people or they are not suitable and accurate for them. In order to really help the blind people, a device named “Intelligent Path-Finder” is already introduced in this paper. This device consists of a few sensors with a combination of microcontrollers and a power source which is already described in the system description section. As a result, it can detect any type of obstacle and uneven surface up to 180° around the user. It also detects the liquid and darkness with its inbuilt method with 100% accuracy. Despite being so accurate and advanced, the cost of the device is kept in a very low range which may be approximately $50 and the cost tends to be lower than the many devices available in the market having same or less features.

REFERENCES

[1] "The Free Dictionary," [Online]. Available: https://encyclopedia2.thefreedictionary.com/Blind+people. [Accessed 29 November 2018].

[2] "World Health Organization," [Online]. [Accessed 29 November 2018].

[3] "World Health Organization," 8 December 2017. [Online]. [Accessed 29 November 2018].

[4] "IAPB," [Online]. Available: https://www.iapb.org/vision-2020/who-facts/. [Accessed 29 November 2018].

[5] "Components101," [Online]. Available: https://components101.com/ultrasonic-sensor-working-pinout-datasheet. [Accessed 29 November 2018].

[6] "Kitronik," 12 July 2018. [Online]. Available: https://www.kitronik.co.uk/blog/how-an-ldr-light-dependent-resistor-works/. [Accessed 29 November 2018].

[7] N. S. Khan, S. Kundu, S. Al Ahsan, M. Sarker and M. N. Islam, "An Assistive System of Walking for Visually Impaired," *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, Rajshahi, 2018, pp. 1-4. doi: 10.1109/IC4ME2.2018.8465669

[8] M. F. Saaid, I. Ismail and Mohd Zikrul Hakim Noor, "Radio Frequency Identification Walking Stick (RFIWS): A device for the blind," *2009 5th International Colloquium on Signal Processing & Its Applications*, Kuala Lumpur, 2009, pp. 250-253. doi: 10.1109/CSPA.2009.5069227

[9] S. Chaurasia and K. V. N. Kavitha, "An electronic walking stick for blinds," *International Conference on Information Communication and Embedded Systems (ICICES2014)*, Chennai, 2014, pp. 1-5. doi: 10.1109/ICICES.2014.7033988

[10] A. Shaha, S. Rewari and S. Gunasekharan, "SWSVIP-Smart Walking Stick for the Visually Impaired People using Low Latency Communication," *2018 International Conference on Smart City and Emerging Technology (ICSCET)*, Mumbai, 2018, pp. 1-5. doi: 10.1109/ICSCET.2018.8537385

[11] Z. Saquib, V. Murari and S. N. Bhargav, "BlinDar: An invisible eye for the blind people making life easy for the blind with Internet of Things (IoT)," *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, Bangalore, 2017, pp. 71-75. doi: 10.1109/RTEICT.2017.8256560

A multi-controlled semi-autonomous wheelchair for old and physically challenged people

Md. Eftekhar Alam
Dept. of EEE
International Islamic University Chittagong (IIUC)
Chittagong, Bangladesh
eftekhar78@gmail.com

Mohammed Abdul Kader
Dept. of EEE
International Islamic University Chittagong (IIUC)
Chittagong, Bangladesh
kader05cuet@gmail.com

Umme Hany
Dept. of CSE
International Islamic University Chittagong (IIUC)
Chittagong, Bangladesh
honeysweety237@gmail.com

Rumana Arjuman
Dept. of CSE
International Islamic University Chittagong (IIUC)
Chittagong, Bangladesh
rumu844@gmail.com

Aisha Siddika
Dept. of CSE
International Islamic University Chittagong (IIUC)
Chittagong, Bangladesh
aishamumu22@gmail.com

Md.Zakaria Islam
Research and Development Division
Hyper Systems ltd
Dhaka, Bangladesh
zakariarai.mailbox@gmail.com

***Abstract*— Physical disability has been incessant thinking for the past several eras. Wheelchairs are the sift partner for most of the disabled people in their day-to-day life. In order to collaborate with physically disabled people, a multi-controlled semiautonomous wheelchair is developed in this paper which can carry them prudently without the assistance of any other human caretaker. The disabled people can navigate and control the speed of the wheelchair by joystick, voice command or gesture command. The system can detect the obstacle in the front side and back side of the wheelchair and can stop the movement automatically to avoid collation. It can also detect any hole when it is moving in the font direction and stops the movement to protect the disabled person from an accident. In case of an accident if wheelchair tilted extensively from its normal orientation the system automatically stops the motors, sends SMS to the relative with GPS location and sounds the buzzer to draw the attention from others. The evaluation of the performance of the system shows that it can be a very helpful friend in the daily life of physically challenged people.**

Keywords*— wheelchair; voice control; gesture control; microcontroller; joystick

I. INTRODUCTION

Disability in mobility can be either an inborn or acquired with age problem. It could also be the effect of a disease. People who have a broken bone also fall into this category of disability. Types of disabilities include various physical and mental impairments that can hamper or reduce a person's ability to carry out his day to day activities. These impairments can be termed as a disability of the person to do his or her day to day activities. Currently, around 15% of the world's population, or roughly 650 million people, live with a disability. In most of the countries, females have higher rates of disability than males. As the population ages, this figure is expected to increase. Eighty percent of persons with disabilities live in developing countries. Therefore it is a prodigious responsibility of the researcher to develop a user-friendly and cost-effective system to ensure the safe movement of the disabled people.

Wheelchairs driven by electric motors can be classified as autonomous, semiautonomous or manually guided wheelchair [5]. A lot of research work on a wheelchair has been done where the researchers proposed autonomous, semiautonomous or manually guided wheelchair to assist the comfortable and safe movement of physically challenged people. Voice, gesture, and the touch-based wheelchair is introduced in [1,2,3,4] in which the user can navigate the wheelchair by means of voice, gesture or touch switch. Computer vision based smart wheelchair is proposed in [7] where the system can detect the obstacle and track human. In [8], a smart wheelchair model is proposed which can be controlled by eye tracking. Some previous research only focused on controlling the speed and direction of wheelchair whether some proposed system also considered the detection of obstacle during movement of the wheelchair.

In our proposed system we have focused on combining all the possible features of previous work considering multiple options of movement control, safety issues, accident detection and tracking the location of the wheelchair. The major contributions of this research are listed as follows-

- The system has several options such as gesture, voice, joystick, and android application to control both the speed and direction of wheelchair.
- The system can detect the obstacle as well as a hole in the front and back side of the wheelchair.
- The 3-axis accelerometer sensor is incorporated to sense the orientation of the wheelchair to detect an accident.
- It sends SMS with location information to the relative of the disabled person and also plays sound by buzzer in an accidental case.
- By sending an SMS to the system the relative of the disabled person can know his/her current location.

II. OVERVIEW OF THE SYSTEM

The overall block diagram of the system is shown in "Fig. 1". An 8-bit microcontroller (ATmega328P) is used which is act as a brain of the system. It coordinates all the input-output devices of the system following the instructions loaded in its ROM. Sonar sensor-1 and sonar sensor-3 are used to detect the obstacle in the front and back side of wheelchair respectively, whether sonar sensor-2 is used to detect any hole in front of the wheelchair. The role of the

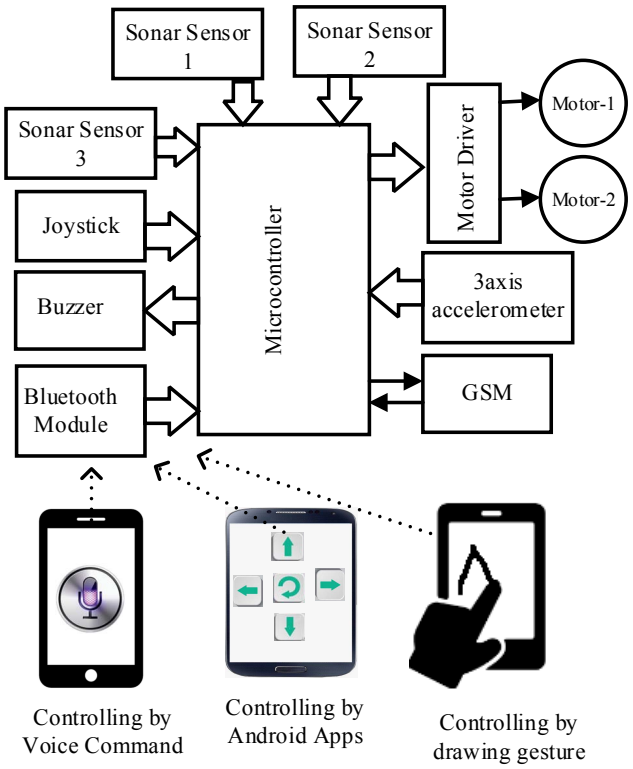


Fig. 1. Block Diagram of the System.

joystick is to control the speed and movement of the wheelchair. Controlling the movement of the wheelchair by voice, gesture command or android apps depends on a smartphone where a Bluetooth device in the system receives data from the smartphone and takes a decision. The microcontroller can detect the orientation of wheelchair with respect to the earth surface by reading data from a 3-axis accelerometer sensor. Abnormal orientation represents an accident to the system. Buzzer module is used to produce sound in case of an accident. Two dc motors move the wheelchair in forward, reverse, left or right direction. These dc motors are controlled by the microcontroller via a motor driver. The system read the location of the wheelchair through a GSM modem and also sends SMS to the relative of disable person in case of an accident. The system sends an SMS with the current location of the wheelchair through the GSM modem in the response of an SMS received from the relative of the disabled person.

III. System Design

This section describes the methodology of developing the prototype of the proposed system. The development process includes controlling speed and direction of DC motors by joystick, controlling the movement of the wheelchair by voice and gesture, detection of obstacle and accident, reading location and sending SMS by GSM.

A. Controlling speed and direction of wheelchair

The movement of the wheelchair depends on two DC motors. If both the DC motors rotate in the forward or reverse direction, the wheelchair goes in forward or reverse direction respectively. Turning left and turning right is performed by changing the speed or direction of DC motors.

In the proposed prototype, the microcontroller controls the speed and direction by L298N motor driver which is a dual H-bridge motor driver IC. It has 4 input (IN1, IN2, IN3, IN4) and two enable (ENA, ENB) pins. The microcontroller can change the direction of DC motors by changing the logic level of these four input pins of the motor driver. The speed of the motor depends on the enable pins. If the status of enable pin is logic 1, the motor will rotate with its maximum speed. But applying pulse width modulated (PWM) signal with different duty cycle to the enable pins can vary the speed of the motor. This technique is useful to control the speed of wheelchair as well as to turn it in the left or right direction. The relation of movement associated with the status of input and enable pins of the motor driver is illustrated in "Table-I".

TABLE I. Movement of Wheelchair

IN1	IN2	IN3	IN4	ENA	ENB	Movement of wheelchair	
						Direction	Speed
0	0	0	0	0	0	Stop	0
1	0	1	0	1	1	Forward	Maximum
1	0	1	0	PWM	PWM	Forward	Depends on duty cycle
0	1	0	1	1	1	Reverse	Maximum
0	1	0	1	PWM	PWM	Reverse	Depends on duty cycle
1	0	0	0	1	1	Right	Maximum
0	0	1	0	1	1	Left	Maximum
1	1	1	1	1	1	Stop	0

To control the movement of the wheelchair by joystick, microcontroller read data from joystick and set the status of input and enable pins of a motor driver following "Table-I". A joystick has two potentiometers that allow us to measure the movement of the stick in 2-D. The potentiometers are variable resistors and they act as sensors providing us with a variable voltage depending on the rotation of the device around its shaft. The change of voltages with respect to the movement of the joystick shaft is anlog that changes from 0 to 5V for the complete movement in both x or y-direction. These analog voltages are converted into digital by using the A/D converter of the microcontroller. The microcontroller

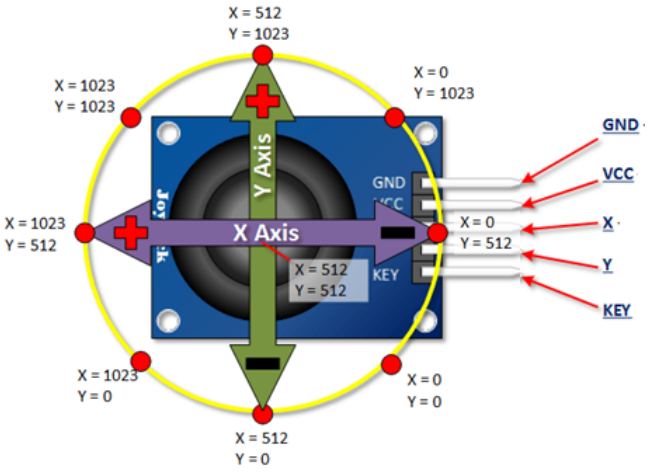


Fig. 2. Digital value of voltages in different positions of joystick shaft.

used in this project has a 10-bit A/D converter with six analog channel. We have set the reference voltage of A/D converter as 5V. Therefore, the conversion of the analog voltage from 0 to 5V results a 10-bit data which varies from 0 (00 0000 0000 in binary) to 1023 (11 1111 1111 in binary). In the normal position of joystick shaft the analog value of both x and y pins are around 2.5 V (corresponding digital data is 512). But if we move the shaft in right towards the x-axis the value reduces gradually from 2.5 V to 0 V (corresponding change in digital value: 512 to 0) whether moving left along the x-axis gradually increases the voltage from 2.5 to 5V (corresponding change in digital value: 512 to 1023). The digital value of voltages in different position of joystick shaft is shown in "Fig. 2". The relation of these digital data with motor direction and speed is shown in "Table-II".

TABLE II. MOVEMENT OF WHEELCHAIR

x-value	y-value	Motor direction and speed
400 to 600	400 to 600	Stop
601 to 1023	400 to 600	Forward direction with speed 0 to maximum.
399 to 0	400 to 600	Reverse direction with speed 0 to maximum.
400 to 600	601 to 1023	Left motor forward and right motor reverse with speed 0 to maximum.
400 to 600	399 to 0	Left motor reverse and right motor forward with speed 0 to maximum.

To control the movement of the wheelchair by voice we have used an android application named 'ARM_Voice'. It converts the human voice into text using 'google voice assistant'. The text is then transmitted to the microcontroller of the wheelchair circuit via Bluetooth. Analyzing the received text, the microcontroller controls speed and direction of DC motor. Another application named 'ARM_Gesture' is used which reads a gesture drawn in the smartphone and send different ASCII code for each different gesture to the microcontroller via Bluetooth. The ASCII code transmitted for the different gesture is listed in "Table-III". After receiving these code microcontroller sets the logic state of motor driver input pins and duty cycle of the enable pins to controls the movement of a wheelchair according to the user instruction.

TABLE III. CODE TRANSMITTED FOR DIFFERENT GESTURE

Gesture	Transmitted Code	Gesture	Transmitted Code
 Forward	0100 0001	 Fast	0100 0101
 Reverse	0100 0010	 Medium	0100 0110
 Right	0100 0011	 Slow	0100 0111
 Left	0100 0100	 Stop	0100 1000

B. Obstacle Detection

Sonar sensors are used to detect obstacle and hole in the front or back direction of wheelchair. A sonar sensor has four pins: Vcc, GND, Trigger, and Echo. The trigger pin is connected with an output pin of microcontroller where the echo pin is connected with an input pin. The microcontroller sends a low to high transition in the trigger pin of the sonar sensor. In response to the low to high transition in the trigger pin, the sonar sensor generates and transmits eight sound pulses of frequency 42 KHz. In the meantime, echo in of sonar sensor becomes HIGH. It becomes LOW if the sound transmitted from sensor reflected back to the sonar receiver. Therefore, the HIGH time of the echo pin represents the traveling time of sound to travel 2d distance, where d is the distance of the obstacle. The HIGH time of the echo pin is measured by using the internal timer of microcontroller and this time is placed in the following equation [13] to measure the distance of the obstacle

$$d=(0.0342\times t)/2 \text{ cm} \quad (1)$$

where, d is the distance and t is the time in microsecond measured by the internal timer of microcontroller.

C. Accident Detection

To detect the accident microcontroller continuously read the orientation of the wheelchair from a 3-axis

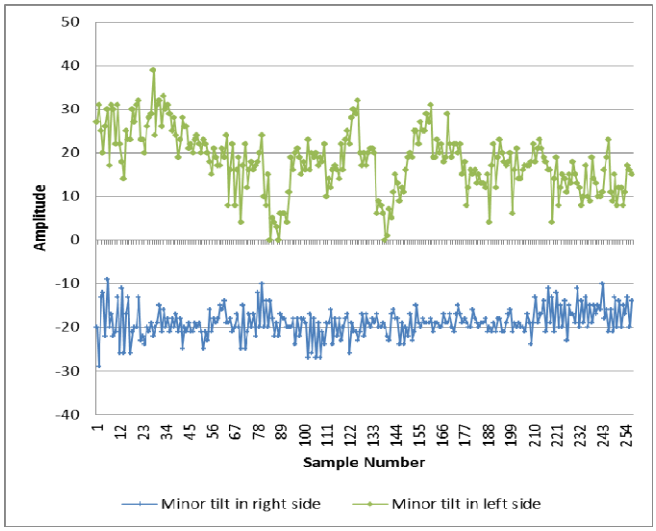


Fig. 3. X-value of accelerometer in normal orientation of wheelchair.

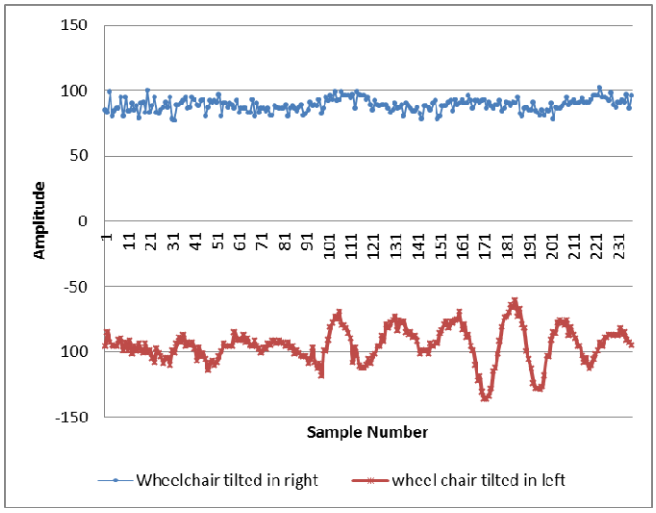


Fig. 4. X-value of accelerometer in abnormal orientation of wheelchair

accelerometer sensor (ADXL345). The accelerometer sensor provides six bytes data which carry information about the orientation (x-value, y-value and z-value) of the sensor with respect to the earth surface. The microcontroller reads the data from the accelerometer sensor through its I2C (Inter-Integrated Circuit) bus: SCL (Serial Clock) and SDA (Serial Data Lines). In normal orientation when wheelchair moves, the x-value, and y-value varies from -30 to 40 which is shown in "Fig. 3". On the other hand, if wheelchair tilted perilously or capsized due to accident x-value and y-value of ADXL345 will fall below -30 (up to -150) or rise above 40 (up to +150) which is illustrated in "Fig. 4". Therefore, the microcontroller can detect the accident examining the data of ADXL345.

D. Reading Location and Sending SMS

The GSM (SIM800) module is used to get the location and send SMS. The microcontroller communicates with the GSM through some AT commands using UART (Universal Asynchronous Receiver and Transmitter) communication protocol. The baud rate used for serial communication is 115200 bps. The list of AT commands [14] used to read location and for sending and receiving SMS are listed in "Table-IV".

TABLE IV. LIST OF AT COMMANDS

AT Commands	**Functions**
AT+CPIN	Returns READY when modem is connected with network.
AT+CMGF=1	To setup SMS as a TEXT message
AT+CMGS="Phone Number" <Enter> Message Body <Ctrl-Z>	To send a SMS
AT+CNMI=2,2,0,0,0	Activate direct reading of incoming SMS.
AT+SAPBR=3,1,"Contype","GPRS"	Set the connection type to GPRS
AT+SAPBR=3,1,"APN","www"	Enable the GPRS
AT+SAPBR=2,1	Query if the connection is setup properly, if we get back a IP address then we can proceed
AT+CIPGSMLOC=1,1	Read Location (latitude and longitude)

IV. IMPLEMENTATION AND RESULTS

In this section, algorithm of the project, the complete implementation and results have been discussed with the proper demonstration.

A. Flowchart

Flowchart of the program loaded in the microcontroller is given in "Fig. 5". The instructions (program) are written in two main functions: void setup() and void loop(). The instructions written in void setup() function are executed only one time where the instructions written in void loop() are executed repeatedly. Here, joining I2C bus as master, initializing I/O pins, sending control word to accelerometer, defining UART baud rate, etc are executed only one time and other instructions are placed in an infinite loop. In the infinite loop section, microcontroller continuously

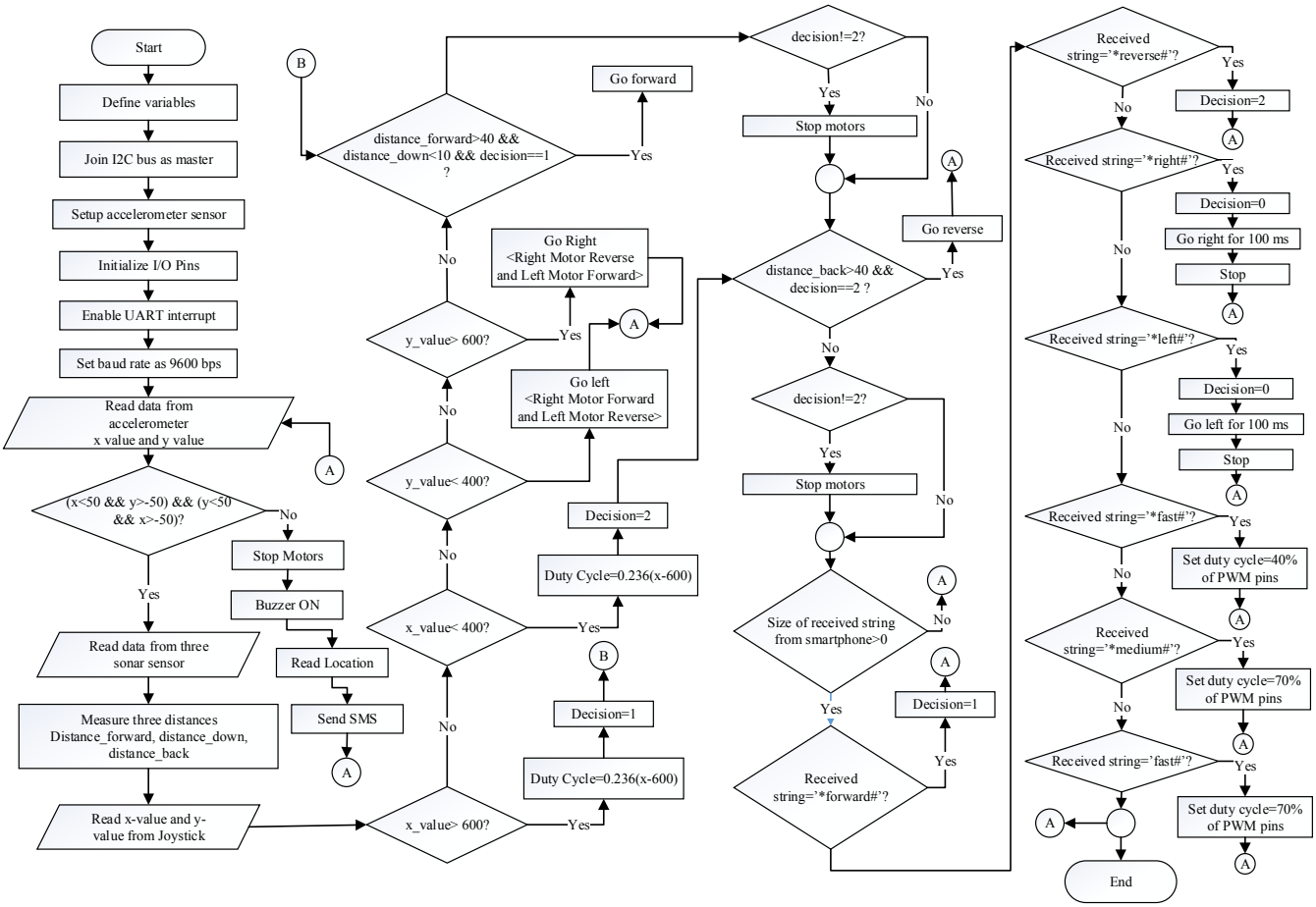


Fig. 5. Flowchart of the program loaded in microcontroller.

read the x-axis and y-axis value of 3-axis accelerometer and check the given condition. If the condition is false, microcontroller stops all motors and sends HIGH-LOW pulse to the buzzer. In case, the condition is true, the microcontroller reads all sonar sensors and calculates the distance of the obstacle from all sensors. Then it checks the given conditions, read the data from joystick and also read the user instructions transmitted from the smartphone to take a decision.

B. *Implementation*

A prototype of the system is implemented according to the methodology described above. The implemented prototype is shown in "Fig. 6". All the components used to develop the prototype and total price of the components are listed in "Table V". To implement a real system which can carry a human with a weight of 80 kg, we have to change only motors, motor driver and chassis of the developed prototype without changing the other components.

TABLE V. LIST OF COMPONENTS

Components Name	Quantity	Unit Price	Total Price
Microcontroller (ATmega328P)	1	420	420
Sonar Sensor	3	80	240
3-axis accelerometer (ADXL345)	1	890	890
GSM (SIM800L)	1	500	500
Motor Driver (L298N)	1	154	154
DC Gear Motor	2	400	800
Joystick	1	218	218
Bluetooth HC05	1	279	279
Buzzer	1	10	10
Chassis	1	1000	1000
	Total Price		4,511



Fig. 6. Implementation of speed controller circuit.

C. *Results*

To demonstrate the performance of the implemented prototype we have taken the value of different parameters such as data of joystick, the distance of obstacles (forward distance, backward distance, and down distance), and the duty cycle of PWM signals given to the enable pins of the motor driver (normalized speed of the left and right motor). All these parameters are the internal variables of the microcontroller and are recorded to an SD card during the operation of the system. Then to realize the operation of the wheelchair we have plotted the different parameters with respect to time. "Fig. 7" and "Fig. 8" represents the normalized speed (duty cycle of the PWM signals of motor driver enable pins) and direction of DC motors with the movement of joystick shaft. In "Fig. 7", the x-value of joystick has a value of 128 in normal condition and the speed of both motors is zero for this situation. But, when x-value increases from 128 to 255, the speed of motors also rises from 0 to 100% in the forward direction. On the other hand, when x-value decreases from 128 to 0, the speed of the motors is rising from 0 to 100% in the reverse direction. The variation of motor speed with the movement of the joystick shaft along the y-axis is illustrated in "Fig. 8".

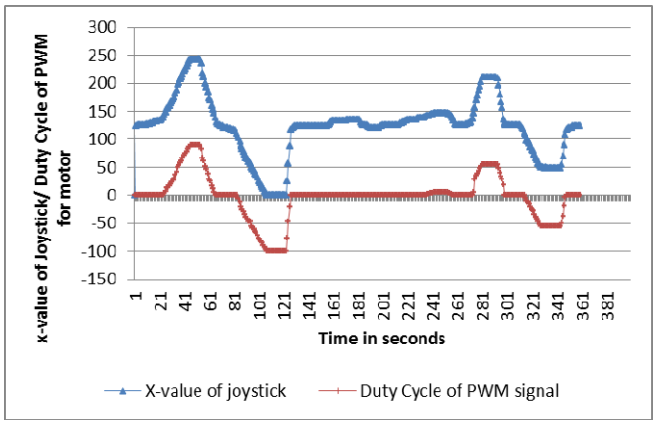


Fig. 7. Variation of motor speed with the movement of joystick shaft along x-axis.

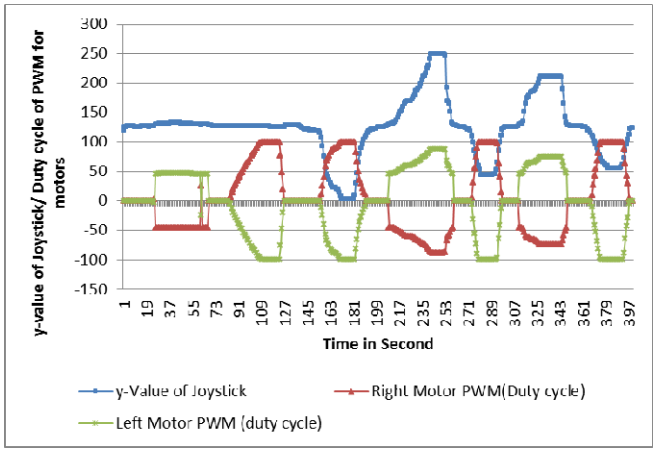


Fig. 8. Variation of motor speed with the movement of joystick shaft along y-axis.

"Fig.9 ", shows the performance of wheelchair in different voice command with the distance of obstacle in front, back and down direction. In the 8th second of operation, the system has received the voice command 'go forward', as a response of the command both motors are going forward with 100% speed. In the 23rd second, the system has received the command 'go slow' and motor speed decreases to 45%. Similarly, after 41 seconds later the system sets the

motor speed as 60% after getting the voice command 'go medium'. During the time (8th second to 56th second) as the wheelchair has moved in the forward direction, the distance of obstacle in the forward direction was coming closer. Eventually, when the distance of obstacle in forwarding direction reduced below 40 cm, the speed of the motor became zero.

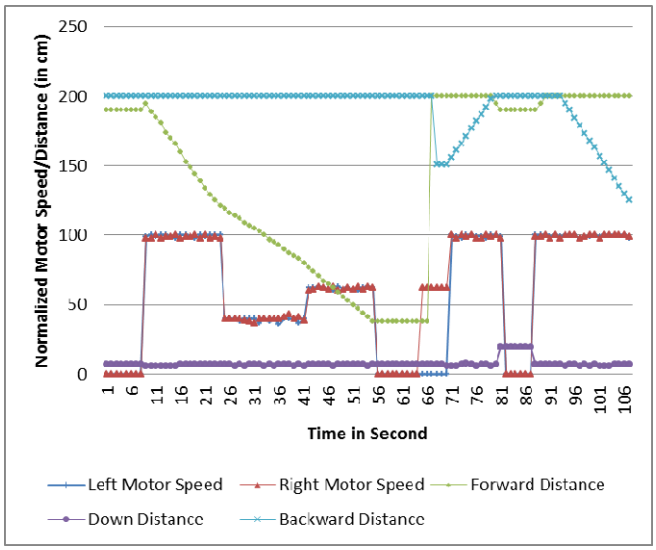


Fig. 9. Performance of the Wheelchair.

V. CONCLUSION

This project is comfortable for those people who are physically disabled. It provides more comfort to move their wheelchair in their own directions without the help of other people. This project greatly decreases the dependency of physically challenged people on their family members. The wheelchair also provides efficient risk management by obstacle detection and obstacle avoidance. During any emergencies or accident, disabled people can alert any of the people around him. He or she can easily do this by the help of the buzzer that is placed on the wheelchair. The wheelchair is designed in a cost-effective way so that if our project is commercialized the needy users in developing countries will get benefit from it. In future we can add some more useful features to the proposed system like autonomous tracking to carry disabled people from one location to another location, computer vision based object detection and IoT based health monitoring.

REFERENCES

[1] K. Sudheer, T.V.Janardhana rao, CH. Srhidevi, M.S.Madhan Mohan, Voice and Gesture Based Electric-Powered Wheelchair Using ARM 2278-5841, Vol 1, Issue 6, November 2012.

[2] Dalsaniya, Artee Kunal, and Dhanashri H. Gawali. "Smart phone based wheelchair navigation and home automation for disabled." Intelligent Systems and Control (ISCO), 2016 10th International Conference on. IEEE, 2016.

[3] Khadilkar, Shraddha Uddhav, and Narendra Wagdarikar. "Android phone controlled voice, gesture and touch screen operated smart wheelchair." 2015 International Conference on Pervasive Computing (ICPC). IEEE, 2015.

[4] Prathyusha, M., K. S. Roy, and Mahaboob Ali Shaik. "Voice and touch screen based direction and speed control of wheel chair for physically challenged using arduino." International Journal of Engineering Trends and Technology (IJETT) 4, no. 4 (2013): 1242-1244.

[5] Nakamura, Atsuhiro, et al. "Development of intelligent power wheelchair assisting for frail elderly people on daily life." ICCAS-SICE, 2009. IEEE, 2009.

[6] Leaman, Jesse, and Hung Manh La. "A comprehensive review of smart wheelchairs: past, present, and future." IEEE Transactions on Human-Machine Systems 47.4 (2017): 486-499.

[7] Utaminingrum, Fitri, et al. "Development of Computer Vision Based Obstacle Detection and Human Tracking on Smart Wheelchair for Disabled Patient." Computational and Business Intelligence, 5th International Symposium on Computational and Business Intelligence (ISCBI). 2017.

[8] Wanluk, Nutthanan, et al. "Smart wheelchair based on eye tracking." Biomedical Engineering International Conference (BMEiCON), 2016 9th. IEEE, 2016.

[9] Gomi, Takashi, and Ann Griffith. "Developing intelligent wheelchairs for the handicapped." Assistive Technology and Artificial Intelligence. Springer, Berlin, Heidelberg, 1998. 150-178.

[10] Soma, Shridevi, et al. "An Approach to Develop a Smart and Intelligent Wheelchair." 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2018.

[11] Nishimori, Masato, Takeshi Saitoh, and Ryosuke Konishi. "Voice controlled intelligent wheelchair." SICE, 2007 annual conference. IEEE, 2007.

[12] Puviarasi, R., Mritha Ramalingam, and Elanchezhian Chinnavan. "Low cost self-assistive voice controlled technology for disabled people." Int J Mod Eng Res (IJMER) 3 (2013): 21332138.

[13] Mohammed Abdul Kader et. al. " Line Following Autonomous Office Assistant Robot with PID Algorithm." 2nd Int. Conf. on Innovations in Science, Engineering and Technology (ICISET), Chittagong, Bangladesh, 2018.

[14] M. E. Alam, M. A. Kader, S. A. Proma and S. Sharma, "Development of a Voice and SMS Controlled Dot Matrix Display Based Smart Noticing System with RF Transceiver and GSM Modem," 2018 21st International Conference of Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2018, pp. 1-5. doi: 10.1109/ICCITECHN.2018.8631914

Towards Developing an Intelligent Wheelchair for People with Congenital Disabilities and Mobility Impairment

Tasneea Hossain*, Md Sabbir-Ul-Alam Sabbir†, Asma Mariam‡,
Toki Tahmid Inan, Muhammad Nazrul Islam, Khairul Mahbub, Muhammad Tawsif Sazid
Department of Computer Science and Engineering, Military Institute of Science and Technology, Dhaka-1216, Bangladesh
Email: *tasneeahossain11@gmail.com, †sualsabbir@gmail.com, ‡asmamariam512@gmail.com,

***Abstract*—Physical disability in people may be either congenital or acquired. In either case, the disabled person requires an aid for movement in both the indoor and outdoor environment. The most common form of assistance to provide support to the disabled is the wheelchair. Therefore, the aim of this article was to design and develop a prototype of the smart wheelchair to provide assistance to various types of mobility impairments. People suffering from various motor disorders are incapacitated and thus, the chair has the provision for both - voice commands as well as manual controls via a joystick. The wheelchair includes a customizable voice control system which is Bengali by default and incorporates a feature for asking help in case of necessity. Along with this, the provision of continuous monitoring is also present for ease of caregivers to locate the user when help is asked or the caregiver is unable to be physically present with the user. Other features such as obstacle detection, speed control on ramps, lighting up of the headlight in dark environment, reclining of the wheelchair have been integrated into the wheelchair as well. The prototype was demonstrated and evaluated in a controlled environment. The outcome of the experiments proved its effectiveness and efficiency for the focused users.**

***Keywords*—** Intelligent system; disability; congenital disability; mobility impairment; smart wheelchair; assistive tool

I. Introduction

Disability is a disorder that hampers the daily life of the affected people by limiting the person's movement, senses, ability to think or act. Disability can either be congenital or acquired. Congenital disabilities are usually inherited while acquired disabilities occur due to any accident or illness. According to the report of Australian National University, disabilities can be of different types i.e. vision impairment, deaf or hard of hearing, mental health conditions, intellectual disability, acquired brain injury, autism spectrum disorder, physical or motor disability [1]. Motor disability is the partial or total loss in muscular functions of the body [2]. These disabilities include arthritis, cerebral palsy, multiple-sclerosis, muscular dystrophy, acquired spinal injury (paraplegia or quadriplegia), post-polio syndrome, spina bifida [3]. All of these disabilities inhibit the normal movement of the affected people.

Nearly 15% of the world's population consists of people having disabilities and 2-4% of them have serious problems in functioning [4]. According to a U.S. Census Bureau report during the period of 2008-2012, nearly 40% of people aged 65 and above had at least one disability and of those 15.7 million people, two-thirds of them were reported to have problems in either walking or climbing [5]. In 2014, 3.8 million veterans in US had a disability that was a result of a disease or injury during their service in the military [6]. Bangladesh has a population of 163 million [1] of which nearly 10% comprises of people with disability[2]. So, a significant number of people in Bangladesh are suffering from some kind of disabilities.

Despite the fact that, one of the most common aids for movement for the physically impaired people is the wheelchair, it is still not enough to meet the needs of all types of motor disabilities. Wheelchairs are mainly designed to help people suffering from disabilities in the lower limb area. As a result, people who have issues like tremor or problems in the upper limb area are unable to operate the wheelchair properly. Along with this problem,it is almost impossible to always provide assistance to the affected patient. Smart wheelchairs are widely used in developed countries to address these problems which costs roughly between $750 to $13000 [7]. Thus they are not affordable to the disabled people of developing countries like Bangladesh. Moreover, language barrier has been a barrier in adopting technology in Bangladesh. Keeping these problems in mind, a smart solution has been devised that incorporates the disabled patient to have the maximum control of the wheelchair and can be operated with the help of voice as well as a joystick. It also includes continuous monitoring so that caregivers can check on the patient without being physically present near the patient at all times. The main aim is to provide complete freedom of movement without any external assistance. The smart wheel is also integrated with other features such as reclining back to avoid muscle cramps, asking for help via SMS by a simple click, obstacle detection, lighting up the wheelchair''s light in dark environment, and speed control to provide maximum ease to the patient.

Therefore, the objective of this research is to design a prototype of a smart wheelchair that will provide maximum assistance to the disabled so that they can move with maximum independence.

In the following sections, Section II presents the overview of the Related Works, Section III comprises of the Conceptual Design of the System, Section IV discusses the Implementation of the System. The demonstration and evaluation results are discussed in Section V. Finally, Section VI contains the Conclusion.

II. Related Works

A significant part of the world population is being affected by the motor or physical impairment. Thus a number of studies has already been carried out to assist them and to ease their struggle

[1] www.worldometers.info/world-population/bangladesh-population/
[2] www.wikipedia.org/wiki/Disability_in_Bangladesh#cite_ref-wbrep_2-0

[8]–[10]. This section briefly introduced the work related to the design. development, and evaluation of wheelchair to support the people with congenital and mobility disability.

A few earlier studies focused to the design and development of wheelchair controlled by voice-commands. For example, Meena et al. [11] developed a wheelchair for paraplegic and other physically disabled people that has feature of voice recognition in addition to the use of joystick. Similarly, the wheelchair proposed by Nishimori et al. [8] included the feature of controlling the system with three types of voice commands: the basic reaction command, the short moving reaction command, and the verification command. Again, wheelchair developed by Mazo et al. [12] and Simpson & Levine [13] also controlled by the voice commands and incorporated the features of detecting obstacles like holes, walls, etc. to reduce the number of collisions on a simple navigation task. On the other hand, a smart wheelchair was developed that used eye movements to drive the wheelchair [14]. A web camera has been used to detect the eye movement and an error handling mechanism has been used to calibrate values of threshold for the eye movement for the user.

Some research highlighted the navigation feature and its accuracy while proposing the wheelchair. The objective of such studies was to provide personalized wheelchair with dynamic routing functions and better accuracy of position [15], [16] . This chair included the real-time navigation support, that helps to move in areas having ramps , doorways with sufficient width for the user to pass. Similarly, Levine et al. [17] developed the *NavChair* that support the navigation of wheelchair for users with large range of disabilities. The *NavChair* included the ultrasonic sensors, joystick and an interface module; obstacle avoidance; door passage and automatic wall following, and was able to move through crowds and make abrupt change in directions with minimal reduction in speed. In another work, Simpson et al. [18] developed a Smart Power Assistance Module (SPAM) which was intended for physically disabled people who were also blind. Keeping vision impairment under consideration, the chair provided obstacle detection to avoid the collisions on a simple navigation task.

The customization of seating position like changing the seat height, movie the seat support bar forward or rearward, transferring from a wheelchair to a bed, etc. were considered in [19], [20] to making the use of such wheelchair comfortable. Some other studies discussed how the Artificial Intelligence was being considered to develop the wheelchair. For example, Braga et al. [21] introduced the concept for intelligent wheelchairs called *IntellWheels* that took inputs from traditional and USB joysticks, keyboard ,head movement as well as facial expressions, while Prathyusha et al. [22], [23] developed an intelligent wheelchair that used the programmable voice recognition circuit for voice commands. Hartman et al. came up with a a smart chair [24] focusing on parallel computer processing where hardware, software and sensor technology was integrated. The system is designed in such a way that the human takes the decision and the rest of its motion is controlled by the smart control technology. Considering the cost associated with a smart wheelchair Simpson et al. came up with a smart wheelchair component system [25] that could be added to existing wheelchairs with minimal modification. It was then tested across four existing wheelchair models.

To sum up, most of the mentioned wheelchairs did not contain any continuous monitoring. Again, a survey conducted by Fehr et al. [26] found the inadequacy of the electrical wheelchair since people with severe disabilities will not be able to steer power wheelchair in areas where space is limited. To find the adequacy of control interfaces, control methods such as joystick, sip-and-puff and head or chin controls were used. The existing wheelchair could not be suitable for many of the physically disabled people from Bangladeshi due to the excessive cost and voice commands (in English). Along with these, the provision of light and speed control for ramps was not handled properly. Therefore, this research concentrates on building a smart intelligent wheel chair with continuous monitoring using Bengali command. This paper handles the aforementioned limitations which in turn paved the way towards an affordable wheelchair option with multiple functionality that can provide maximum support in a affordable cost to the disabled. The proposed system in this paper tries to use the existing technology in a new manner to help physically impaired people in Bangladesh.

III. Conceptual Design of the System

The proposed system is a voice controlled smart wheelchair which supports any language but Bengali has been chosen in order to support the physically impaired people of Bangladesh. Along with the voice control, it also supports manual control. The two methods have been kept for reaching out to various types of mobility impairments. Figure 1 represents the basic work flow of the wheelchair. .

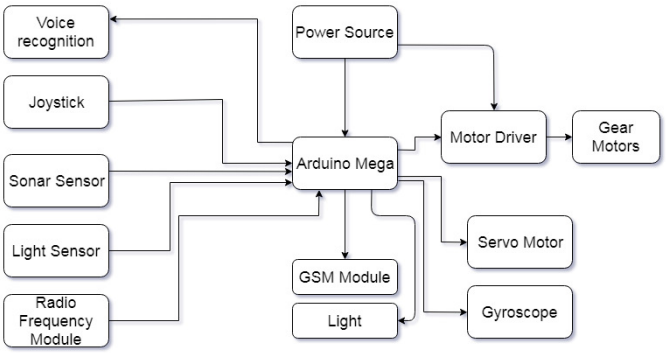
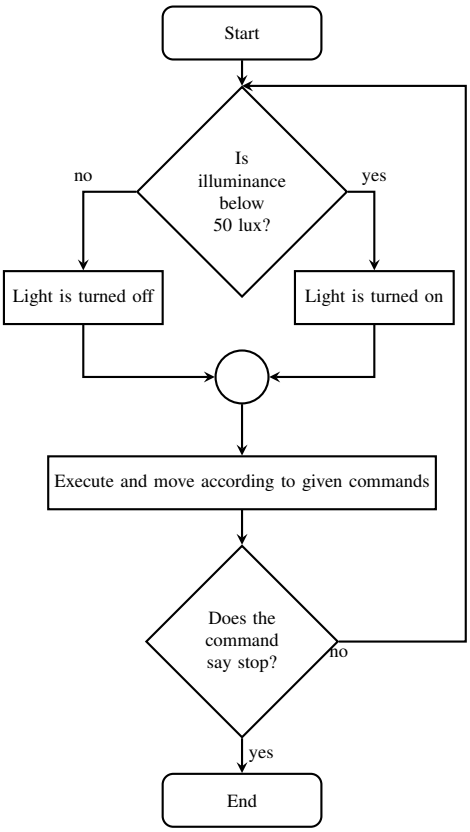


Fig. 1. Architecture of the proposed smart wheelchair

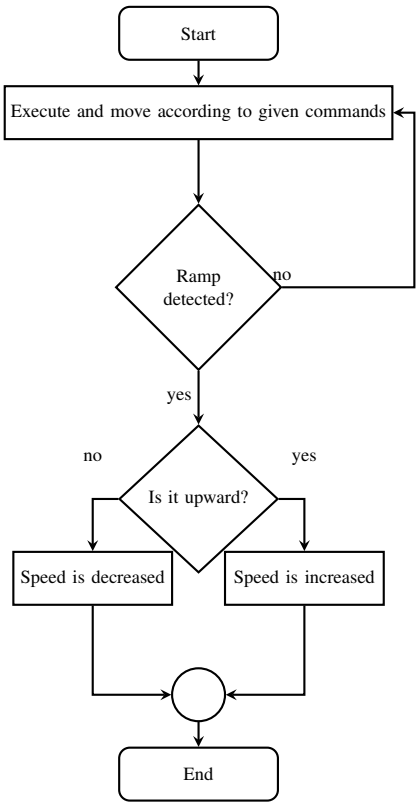
A. Features of the smart wheelchair

The smart wheelchair will include the following features to attain the objective of this research:

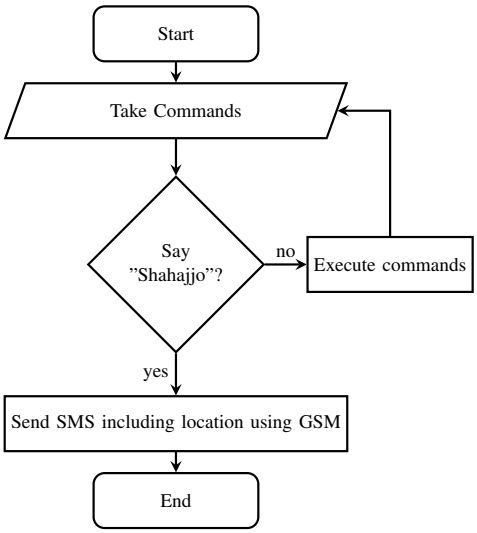
1) *Manual control*: Manual control is for those disabled people who are able to use their hands for controlling the movement of the wheelchair. A joystick is used for enabling the manual control which will move the wheelchair in the forward, backward, right and left direction.The joy stick is not fixed with the body rather it is wirelessly connected and movable.
2) *Controlling through voice commands*: Disabilities like quadriplegia make the users unable to use manual control for wheelchair for which voice commands has been incorporated with the system. The commands can be customized to any language. For this project Bengali language has been considered. The commands are 'Dane', 'Bame', 'Samne', 'Pichone', 'Thamo' which interpret in English as 'Right', 'Left', 'Forward', 'Backward' and 'Stop' respectively.
3) *Sending SMS for help*: Despite the various precautions taken, the user might still fall down from the wheelchair or need external help. For this, the voice command, ′Shahajjo′ (help) will send an SMS including the chair's current location to the user's caregiver asking for assistance. Figure 2(c) represents the working procedure of the system for Sending SMS.
4) *Longitudinal obstacle detection* : Obstacle detection is required so that the wheelchair does not bump into pillars, walls or any other object even if the user gives gives commands that would accidentally cause collision. Therefore, upon detecting, the wheelchair will automatically stop when a certain obstacle



(a) Light up



(b) Speed control



(c) Sending SMS

Fig. 2. Working procedure for light up (a), speed control (b), and sending SMS (c)

is present at the front or rear and will not execute any given command in that direction.

5) *Speed Control*: While going up and down the ramp, the speed needs to automatically counteract the increasing or decreasing speed due to gravity. So, for slopes or ramps, depending on the upward or downward direction the speed will automatically increase and decrease respectively.It works for both forward and backward direction. Figure 2(b) shows the working procedure of the process of speed control.
6) *Continuous Monitoring*: Many disabled people often have caregivers. Since it's not always possible for them to be present at sight, the continuous monitoring is kept so that the user's movement can be tracked with current location. Its done by receiving the satellite data through GPS. Then the WIFI module gets that data, decodes it into latitude and longitude and sends it to the cloud server. Thus the wheelchair's location can be accessed from the mobile phone. Figure 3 describes the architecture of Continuous monitoring using GPS and Wi-Fi Module.

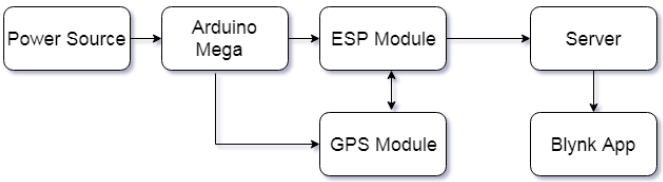


Fig. 3. Block Diagram of Continuous Monitoring

7) *Reclining the wheelchair*: Sitting down continuously in the same position is not good for the health of the user [27]. Moreover user's comfort have to be considered as well. So, option of reclining is kept to break the monotonicity and making it more comfortable. Therefore the back of the wheelchair can be reclined to any position as per the user's comfort.
8) *Light Up*: The user might have to move in indoor or outdoor environment with insufficient light or there might be unavailability of light due to power outage. In these circumstances the wheelchair will sense the shortage of light and automatically light up. Figure 2(a) demonstrates the working procedure of this feature.
9) *Controlling the Wheelchair from a distance*: The user will be benefited by this feature when he is not on the wheelchair. The user can give command from a certain distance and the wheelchair will move according to the command. This will prove to be helpful when the user is lying on the bed and the wheelchair has been moved away from him. We have preferred to use Bengali language for this.

IV. Implementation of the System

A prototype has been developed for demonstrating the above mentioned features which has been shown in Figure 4. The whole implementation was carried out in an academic environment. The wheelchair is powered with the help of two batteries. One low voltage battery is used to power up the Arduino which is programmed to control the wheelchair and another high voltage battery is used to power up the motor driver. The motor driver is able to drive only two motors. Therefore two pairs of motors on each side are connected in parallel and is being driven by the motor driver. There are four wheels for each of the gear motors.

The main hardware components are the different modules that are connected with the Arduino. An Arduino Uno is used for controlling the wheelchair from remote distance. Radio Frequency module is used for this purpose. Two Arduino Mega are used for integrating all other modules implanted on the wheelchair body. It could be done using one Arduino Mega. But to distribute the load and for proper wiring two Arduino Mega have been used.

A Joystick is used for controlling the wheelchair manually. Five commands can be given through this joystick. The commands are for forward movement, backward movement, right movement, left movement and stopping the movement.

Controlling the wheelchair by voice command has been done using voice recognition module. Firstly, the voice recognition module has to be trained. Based on the training, it recognizes the commands and moves accordingly. This wheelchair can be trained in any language. But the Bengali language has been chosen for our academic work. The commands are same as manual commands but in addition to that one more voice command has been added.

In case of danger or emergency,the user can speak out the voice command "Shahajjo" which means "help" in English. Then, it will execute the command and the current location of the wheelchair will be sent by SMS to the previously saved number with the help of GSM module. The current location of the wheelchair is computed from the data of the GPS module.

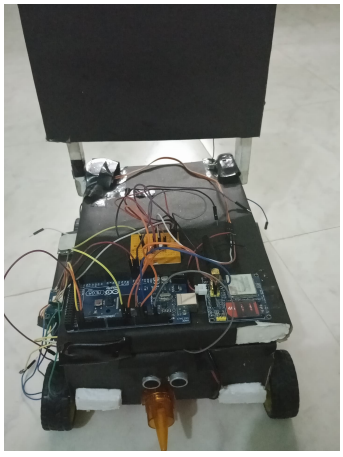


Fig. 4. Image of the Prototype

An additional pre-existing "Blynk" [28] application has been used for viewing the location for continuous monitoring. For this purpose, GPS module and WiFi module is paired up. GPS module collects the latitude and longitude of the current location of the wheelchair and sends to the Blynk app server using WiFi module.
In order to go up and down the ramps with constant speed, the previously mentioned setting of the wheelchair is not enough. So, a Gyrometer has been utilized for detecting the slopes and it has been programmed in such a way that the speed will increase while going up and decrease while coming down to combat the effect of gravity.

Two ultra-sonic sonar sensors have been used to detect obstacles. Inspite of getting command to move, if the sonar sensors sense obstacle in the forward or in the backward direction, the wheelchair will stop moving automatically and it will not further execute any commands in that particular direction till it is away from the obstacle.

If the illuminance of the wheelchair environment is below 50 lux, light sensor will sense this and a headlight which is placed in front of the wheelchair will automatically light up. When there is sufficient illuminance, the headlight will not turn on automatically.

Servo motor is used at the back of the wheel chair to recline the wheelchair according to user comfort. Two switches are placed in the wheelchair to recline the upper part of the wheelchair. Servo motor has been used for the prototype, but for actual implementation it might not be sufficient to balance the weight.

The total cost for making the prototype was near about $125 (10,500 Bangladeshi Tk). An additional $90 (7500 BDT) will be required for a real wheel chair. List of equipments and their costs are given in Table I.

TABLE I
LIST OF EQUIPMENTS WITH COST

Name of the Component	Quantity	Unit Price (BDT)	Price (BDT)	Price (USD)
Arduino Mega	2	750	1500	17.85
Voice Recognition module	1	2400	2400	28.56
Frame for Prototype	1	100	100	1.2
Joystick	1	95	95	1.13
GPS Module	1	870	870	10.35
Gyroscope	1	230	230	2.74
Sonar Sensor	2	95	190	2.26
Light Sensor	1	160	160	1.9
Wifi Module	1	325	325	3.87
GSM module	1	1850	1850	22.01
Battery 11.1V	1	1000	1000	11.90
Arduino Uno	1	410	410	4.88
Breadboard	2	40	80	0.95
RF Module	2	160	320	3.81
Light + 9V battery	1	50	50	0.59
Servo Motor	1	20	20	0.24
Gear Motor	4	71	284	3.84
Motor Driver	1	376	376	4.47

V. EVALUATION

The evaluation study was conducted at Software Engineering Laboratory of the authors' institute. The system was tested in a room of rectangular shape with 8 feet length and 6 feet width. The route and milestone points are showed in Figure 5. To test the effectiveness and efficiency of the system, a number of scenarios were trailed, that includes: (a) the system was run using joystick and then by Bengali voice command ('Dane, <Right>', 'Bame <Left>', 'Shamne <Forward>', 'Pichone <Backward>', 'Thamo <Stop>' from point A and ended at A via B, C, and D. (b) Help request were sent through a Bengali voice command 'Shahajjo <Help>'. (c) While the system was moving forward or backward the present location of the wheelchair was monitored via a mobile phone (that was supposed to be used by the caregiver). (d) the wheelchair was reclined in different position. (e) the wheelchair was run both in absence and presence of light (f) The wheelchair was called by Bengali voice command from facilitator positioned at point A, while the wheelchair was positioned at point C.

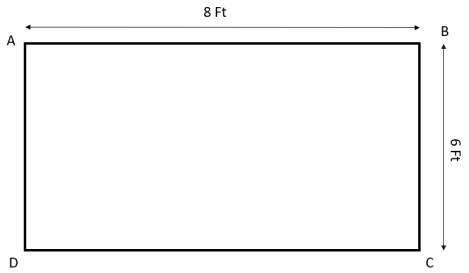


Fig. 5. Map of the place of experimentation

TABLE II
EXPERIMENTATION

Point	Time	Number of Attempts	Sequence of Commands	Collision	Command
A-B	12 sec	1	f	0	Forward
B-C	23 sec	1	r,s,l,s,l,s,f	0	Right, Stop, Left, Forward
C-D	11 sec	2	1st attempt: r; 2nd attempt: f	0	Right, Forward
D-A	4 sec	2	1st attempt, 2nd attempt: f	0	Forward

TABLE III
SUMMARY OF THE EXPERIMENT RESULTS

Feature	Experimental Scenario	Findings
Manual control	(a)	Left, Right, Forward and Backward command works properly and successfully complete the route. No difference was noticed between the manual and voice command.
Controlling through voice command		
Sending SMS for help	(b)	SMS were well successfully received with the location of wheelchair at mobile phone (used by the caregiver)
Longitudinal obstacle detection	(a)	The wheelchair was stopped in each obstacle/end of the route like at point A, B, C, and D (see figure 5)
Speed control	(a)	Accurately executed (observed the Increases and decreases the speed of the wheelchair)
Continuous monitoring	(c)	The present location of the wheelchair was consistently possible to monitor from the mobile phone (used by the caregiver)
Reclining the wheelchair	(d)	Accurately executed (observed the reclining of the wheelchair to different position)
Automatic light turns on	(e)	Light turns on automatically in absence of light in every trial-cases
Controlling/calling the wheelchair remotely	(f)	The wheelchair reached to the caller position successfully using Bangla voice command

To test the system, at first, the system was run along the length, from point A to B. Only forward command was given using voice and it took 12 seconds to reach to B. Since the entire

path is enclosed, the wheelchair stopped at B. The sonar worked successfully since the obstacle was detected. Next, to reach to C, a turn to the right had to be made in order for the wheelchair to be aligned across the path towards C.

One of the drawbacks of the prototype was that the speed was more while turning right and was less in comparison while turning left. So, initially when the right command was given, it had to be stopped first using the stop command and then the left command was used to align it properly. After the wheelchair was lined up properly, the forward command was given, to reach to the destination C. Since readjustments were required while turning, the overall time required to reach to C was comparatively more than before in spite of the distance being lesser than before. When moving from C to D, two attempts were required. In the first attempt, the right command wasn't working and the system had to be restarted. Then the forward command was given to move it from C to D. In the last traversal from D to A, two attempts had to be taken as well. The first attempt resulted in the failure of the voice commands since the transmitter's reception was hampered due to the surrounding noise. In the second attempt, the forward command worked successfully and it took only 4 seconds to reach A. In all of the mentioned steps of testing, the obstacle detection mechanism worked successfully.

VI. Conclusion

The wheelchair provides the means for people with disability to move around freely without requiring much help from the caregivers. It supports them by enabling voice command as well as manual command in order to cater to a large group of disabilities. Along with these, the features of continuous monitoring, sending SMS when in need, obstacle detection, speed control, ability to recline back and lighting up in dark environment work hand in hand to allow the disabled person to have maximum comfort. All of these combine to ensure relief for the disabled people. The prototype required an additional cost of $120 (Bangladeshi Tk 10,500) apart from the cost of the wheelchair itself which is cheaper than available wheelchairs based on the wide range of obtained functionalities.
Thus this can be affordable to the people of Bangladesh compared to the available ones in the market.

One of the limitations of this prototype was the inability to test out the wheelchair on disabled people since only a miniature version of the actual wheelchair was constructed.Future plan is to focus on constructing a normal sized wheelchair with the features mentioned above with the addition to autonomous blind navigation. This would provide the opportunity to create a wheelchair that can be used by disabled patients and required changes can be made according to the feedback obtained from them.

References

[1] "Different types of disabilities," www.services.anu.edu.au/human-resources/respect-inclusion/different-types-of-disabilities, accessed: 2018-06-2.

[2] "Motor impairment," http://www.neuromodulation.com/motor-impairment, accessed: 2018-07-27.

[3] "Physical disability," https://imvc.com.au/youthservices/broaden-your-horizons/disability-information/physical-disability/, accessed: 218-07-27.

[4] "World report on disability," www.who.int/disabilities/world_report /2011/report/en/, accessed: 2018-07-27.

[5] "Mobility is most common disability among older americans, census bureau reports," https://www.census.gov/newsroom/press-releases/2014/cb14-218.html, accessed: 2018-07-27.

[6] "Fff: Veteran's day 2015: Nov. 11, 2015," www.census.gov/newsroom/facts-for-features/2015/cb15-ff23.html, accessed: 2018-07-27.

[7] "How much is an electric wheel chair," https://kdsmartchair.com/pages/how-much-is-an-electric-wheelchair, accessed: 2019-04-12.

[8] M. Nishimori, T. Saitoh, and R. Konishi, "Voice controlled intelligent wheelchair," in *SICE, 2007 annual conference*. IEEE, 2007, pp. 336–340.

[9] N. S. Khan, S. Kundu, S. Al Ahsan, M. Sarker, and M. N. Islam, "An assistive system of walking for visually impaired," in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, 2018, pp. 1–4.

[10] R. C. Simpson, D. Poirot, and F. Baxter, "The hephaestus smart wheelchair system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 10, no. 2, pp. 118–122, 2002.

[11] K. Meena, S. Gupta, and V. Khare, "Voice controlled wheelchair," *International Journal of Electronics, Electrical and Computational System*, vol. 6, no. 4, pp. 23–27, 2017.

[12] M. Mazo, F. J. Rodriguez, J. L. Lázaro, J. Ureña, J. C. Garcia, E. Santiso, P. Revenga, and J. J. Garcia, "Wheelchair for physically disabled people with voice, ultrasonic and infrared sensor control," *Autonomous Robots*, vol. 2, no. 3, pp. 203–224, 1995.

[13] R. C. Simpson and S. P. Levine, "Voice control of a powered wheelchair," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 10, no. 2, pp. 122–125, 2002.

[14] "Smart wheel chair," www.transmitter.ieee.org/makerproject/view/c6e9e?fbclid=IwAR0Qdln9OKbfeFhs_4j-YieArtqvnu6uVrJttH58nFySTMkH1OdHt9zcSL0, accessed: 2019-04-12.

[15] D. Ding, B. Parmanto, H. A. Karimi, D. Roongpiboonsopit, G. Pramana, T. Conahan, and P. Kasemsuppakorn, "Design considerations for a personalized wheelchair navigation system," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE, 2007, pp. 4790–4793.

[16] H. A. Yanco, "Wheelesley: A robotic wheelchair system: Indoor navigation and user interface," in *Assistive technology and artificial intelligence*. Springer, 1998, pp. 256–268.

[17] S. P. Levine, D. A. Bell, L. A. Jaros, R. C. Simpson, Y. Koren, and J. Borenstein, "The navchair assistive wheelchair navigation system," *IEEE transactions on rehabilitation engineering*, vol. 7, no. 4, pp. 443–451, 1999.

[18] R. Simpson, E. LoPresti, S. Hayashi, S. Guo, D. Ding, W. Ammer, V. Sharma, and R. Cooper, "A prototype power assist wheelchair that provides for obstacle detection and avoidance for those with visual impairments," *Journal of NeuroEngineering and Rehabilitation*, vol. 2, no. 1, p. 30, Oct 2005.

[19] D. E. Perry, "Power stand-up and reclining wheelchair," Nov. 22 1994, us patent 5,366,036.

[20] D. Patterson, "Stabilized reclining wheelchair seat," Sep. 3 1991, us patent 5,044,647.

[21] R. A. Braga, M. Petry, A. P. Moreira, and L. P. Reis, "Intellwheels-a development platform for intelligent wheelchairs for disabled people," in *ICINCO 2008: Proceedings Of The Fifth International Conference On Informatics In Control, Automation And Robotics, Vol Ra-2: Robotics And Automation, Vol 2*, 2008.

[22] M. Prathyusha, K. Roy, and M. A. Shaik, "Voice and touch screen based direction and speed control of wheel chair for physically challenged using arduino," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 4, no. 4, pp. 1242–1244, 2013.

[23] "The mit intelligent wheelchair project developing a voice-commandable robotic wheelchair," https://bit.ly/2Ihugov, accessed: 2019-04-12.

[24] V. K. N. Amiel Hartman, "Human-machine interface for a smart wheelchair," *Journal of Robotics*, 2019.

[25] R. Simpson, E. Lopresti, S. Hayashi, I. Nourbakhsh, and D. Miller, "The smart wheelchair component system," *Journal of rehabilitation research and development*, vol. 41, pp. 429–42, 06 2004.

[26] L. Fehr, W. E. Langbein, and S. B. Skaar, "Adequacy of power wheelchair control interfaces for persons with severe disabilities: A clinical survey," *Journal of rehabilitation research and development*, vol. 37, no. 3, p. 353, 2000.

[27] "Here's what sitting too long does to your body," https://fitness.mercola.com/sites/fitness/archive/2015/05/08/sitting-too-long.aspx, accessed: 2018-07-27.

[28] "Blynk iot platform," https://blynk.io, accessed: 2019-03-27.

Wearable Visual-MIMO for Healthcare Applications

Tae-Ho Kwon
Dept. of Electronics Engineering
Kookmin University
Seoul, Korea
kmjkth@kookmin.ac.kr

Ki-Soo An
Dept. of Convergence Design & Technology
Kookmin University
Seoul, Korea
crc2017@kookmin.ac.kr

Partha Pratim Banik, and Ki-Doo Kim*
Dept. of Electronics Engineering
Kookmin University
Seoul, Korea
ppbanik006@kookmin.ac.kr
*Correspondance: kdk@kookmin.ac.kr

***Abstract*— For daily health checkup, wearable device plays an important role in healthcare applications. In this manuscript, we develop a visual multiple-input multiple-output (visual-MIMO) system for healthcare application. As a demo of healthcare device, we design a pulse oximeter device that can measure heart rate (HR) and blood oxygenation (BO) of a person. The HR and BO data are used to send through light-emitting diode (LED) to camera communication. We have used generalized color modulation (GCM) technique to modulate the HR and BO data into the LED array for transmission and smartphone camera is used as the receiver to demodulate the transmitted data. We demonstrate our proposed visual-MIMO system by developing an android app on smartphone.**

Keywords—heart rate, blood oxygenation, LED array, GCM, visual-MIMO

I. INTRODUCTION

For leading a healthy life, we have to take care of our health condition daily. From this sense, wearable medical device plays an important role in daily health checkup. Pulse oximeter is one kind of device that has remarkable application on healthcare. For primary checkup of a patient, doctors use this device to measure heart rate (HR) and blood oxygenation (BO). Based on the estimated HR and BO, they can take decision of further treatment processing. For monitoring HR and BO into the smartphone application, a wireless communication system is demandable. RF communication may one of the best option in this sense but it is not suitable in healthcare application because of the high sensitivity of electromagnetic medical devices to RF radiation. This sensitivity can occur electromagnetic interference and the performance of medical devices is largely affected by this interference [1, 2]. RF radiation is also bad for human health [2]. As a wireless communication technology, light-emitting diode (LED) to camera communication offers the solution against these RF-related effects on health.

In this manuscript, we have developed a visual multiple-input multiple-output (visual-MIMO) based system for healthcare application. Visual-MIMO refers to a wireless communication system between LED array and camera. Generalized color modulation (GCM) is one kind of color space based modulation technique by which information data can be modulated by selecting the colors from a color space. In receiver side, the receiver demodulates the transmitted data by applying image processing algorithm. J.-S. Kim described the application of visual-MIMO in fashion technology and also proposed an intensity variation on-off keying modulation technique [3]. T.-H. Kwon also showed a wearable visual-MIMO based implemented prototype system on fashion technology [4]. In this paper, we have shown an implementation prototype of pulse oximeter device based on visual-MIMO system for healthcare application. We use LED array as transmitter and smartphone camera as the receiver. But the demodulating operation is performed in server PC, MATLAB platform. We follow the method [5] to demodulate the information.

This research was supported by the National Research Foundation of Korea Grant funded by the Ministry of Science, ICT, Future Planning [2015R1A5A7037615].

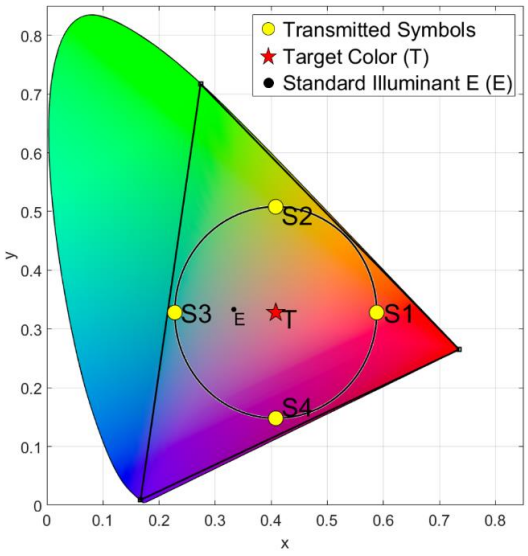


Fig. 1. Example of constellation diagram in CIERGB gamut area on CIE 1931 color space for GCM based visual-MIMO system.

We have organized the manuscript in six sections. Section II describes the GCM based visual-MIMO system. The system methodology is described in section III. We describe the system implementation in section IV. Section V discusses about the result and implementation outcome. We conclude the work through section VI.

II. GCM BASED VISUAL-MIMO SYSTEM

For developing a visual-MIMO based pulse oximeter device, we have to describe the modulation technique of the visual-MIMO system. For modulating information data, GCM demands a constellation diagram on a color space. In this case, we use CIE1931 color space. Fig. 1 shows the example of constellation diagram (a black circle) into the CIERGB gamut area (black triangle area) on CIE1931 color space. According to GCM technique, n-bit information data can be represented in 2^n points [6]. In this paper, we consider 2-bit information data and by taking a target color (T) as a center of the constellation diagram 4 color symbols (S1, S2, S3, S4) are constructed. In order to transmit a color symbol, the transmitter maps each color symbol into the RGB intensity of LED array. At the receiving end, the camera image sensor captures the LED array (transmitter) that emits light and processes the captured LED array to decide the transmitted symbol. For improving symbol decision accuracy, it is important to select a representative color which will be close to the transmitted color symbol.

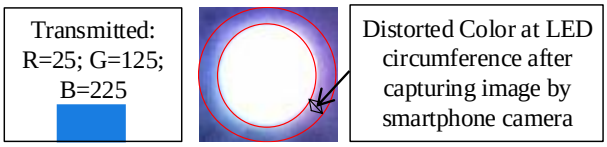


Fig. 2. Color distortion example of received LED image.

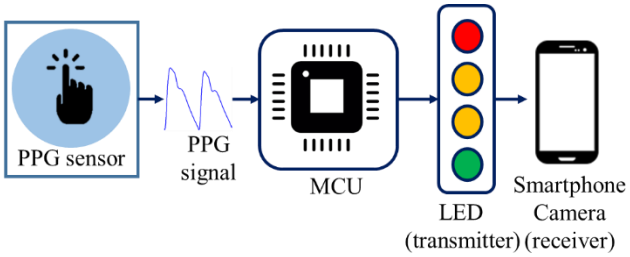


Fig. 3. Conceptual block diagram of the entire pulse-oximeter based visual-MIMO system.

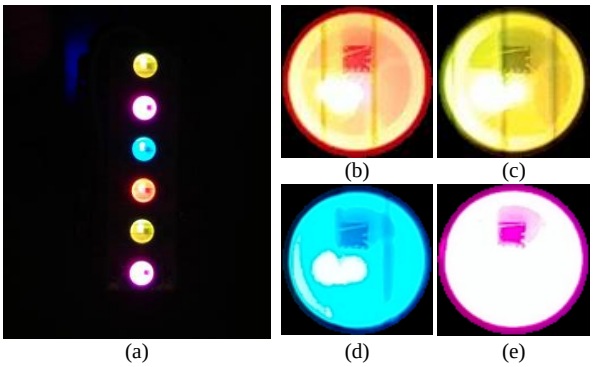


Fig. 4. (a) Captured LED array, (b)-(d) LEDs of four color symbol S1, S2, S3, S4, respectively.

Because of the limitation of smartphone camera dynamic range, the received LED image goes to the saturation. In a normal camera setting, this phenomenon is occurred, as shown in Fig. 2. As a result, the position of the fully saturated colors will be (x = 1/3, y = 1/3), which are coordinates of the standard illuminant E (R: G: B = 1: 1: 1). The edge of the light source circle (that means LED) bears the color information, as shown in Fig 2. This problem can be solved by capturing LED image with lower exposure by adjusting the camera settings [7]. To determine the representative color of the received color symbol, conventionally, representative color of the symbol is determined by taking average of all the pixels. This is an error-prone method. For this reason, we use the algorithm of paper [5] to determine the representative color of the transmitted symbol. In the symbol decision step, by using the Euclidean distance we decide the transmitted symbols according to the closest representative color position of the received symbol.

III. System Methodology

In this section, at first, we will describe conceptual block diagram of the entire system. Then, we describe the transmitter and receiver of our pulse oximeter based visual-MIMO system.

Fig. 3 shows the conceptual block diagram of the entire system. To develop the pulse oximeter device, we use MAX30100 Photoplethysmogram (PPG) sensor. We extract the PPG signal from this sensor. The PPG sensor is combined with two LEDs and a photodetector. Two LEDs

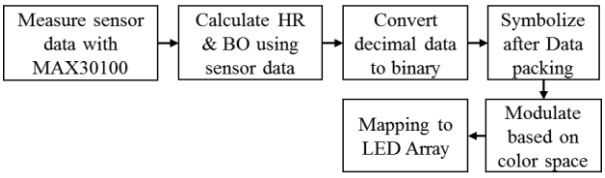


Fig. 5. Transmitter block diagram.

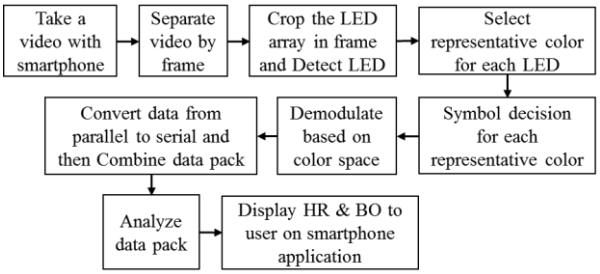
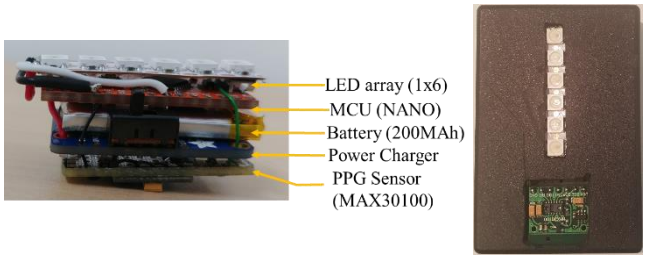


Fig. 6. Receiver block diagram.

are essential for measuring the BO rate. After extracting PPG signal, we give the two PPG signals as the input to the analog to digital converter (ADC) of microcontroller processing unit (MCU). The MCU measures the HR and BO from the PPG signal. Then, it converts the HR and BO to the binary data to send them into the LED array. In the receiver side, the captured LED array image by smartphone camera is processed by image processing algorithm. The LEDs are separated from the LED array by using circular Hough transform [8]. Fig. 4(a), and 4(b)-4(e) show the captured LED array, and separated LEDs for each color symbol (S1 to S4) from LED array.

Fig. 5 shows the block diagram of transmitter. In the transmitter side, at first, we measure the sensor data from MAX30100. Then we calculate the HR and BO by using this sensor data. We convert the HR and BO to the binary data. HR converts in 8-bit data and BO converts in 4-bit data. As every LED represents 2-bit data, we use 4 LEDs for HR and 2 LEDs for BO. So, in total, we use 6 LEDs for modulating HR and BO data by GCM technique. The reason of selecting 4-bit for BO is that in general, normal pulse oximeter is designed for the range of 90% to 100% [9]. But for our designed pulse oximeter, we set this range from 86% to 100% and 4-bit data can easily cover this range by subtracting 4-bit data from 100. Then, in total, 12-bit data is mapped to the 4 color symbols by using GCM technique. After that the symbols are mapped to 6 LEDs.

Fig. 6 shows the block diagram of receiver side. In the receiver side, at first, we take a video with smartphone and separate frames from video. Our processing is frame by frame. We crop the LED array that means region of interest (ROI). We do not need to do any image processing for selecting ROI region for further processing. Because we assume that user holds the camera in the given alignment of the LED array through the developed smartphone app. In [5], k-means clustering based algorithm is developed to select the representative color. Then we use Euclidian distance method to decide the corresponding transmitted symbol of the received symbol. After symbol decision, we can demodulate the binary data to decimal data by using binary to decimal conversion. Finally, the HR and BO are displayed in the smartphone application. Thus a visual-MIMO based pulse oximeter system is developed.



(a) (b)

Fig. 7. Prototype of pulse oximeter device; (a) side view, and (b) top view.

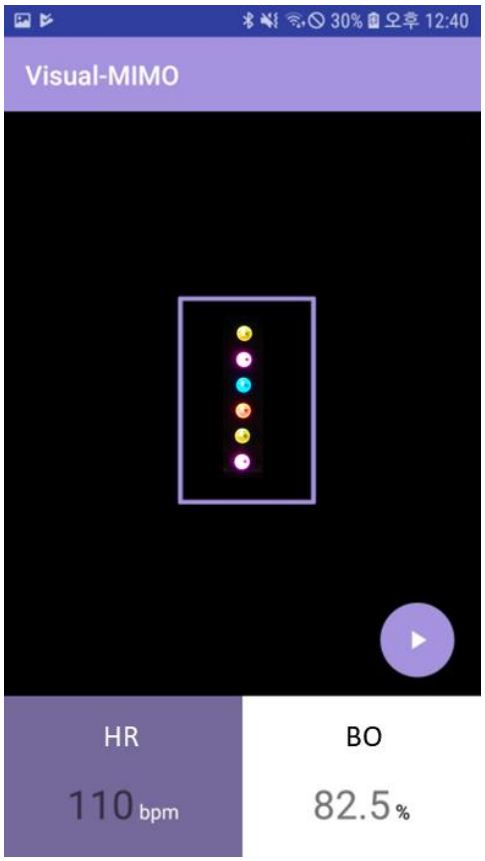


Fig. 8. The developed smartphone app for displaying HR and BO.

IV. System Implementation

We show the implemented hardware of the pulse oximeter device in Fig. 7. We use Arduino NANO MCU not only to configure the transmitter side (LED array) but also extract the HR and BO from the PPG sensor. A 200MAh battery is used to power up the MCU. Fig. 8 shows the developed android app to show the transmitted HR and BO by visual-MIMO system. An ROI region is drawn, violet rectangular box, shown in Fig. 8 for focusing the LED array on that box. The video on-off button is shown into right bottom corner, above the BO data display panel. This app is able to capture video in the constant frame rate of 30 with 50 ISO, 1/200 [s] shutter speed to achieve low exposure. Additionally, we can set the white balance by placing white paper in front of the camera lens with using auto white balance function of smartphone camera. As a result, we used automatic adjusted value (5400K) of white balance which is related to the temperature of the ambient light. We used smartphone LG G6, integrated IMX258 AF camera module with Android 8.0 operating system. We maintain 25cm distance between transmitter and receiver.

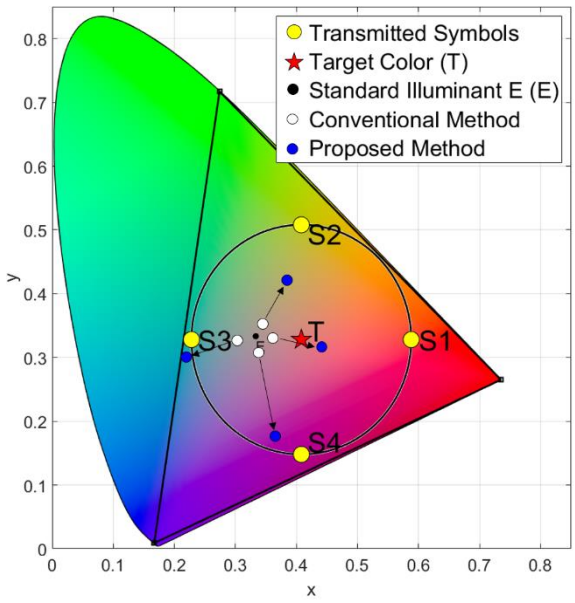


Fig. 9. Selected representative symbol color by using algorithm [5] for a single case of data transmission.

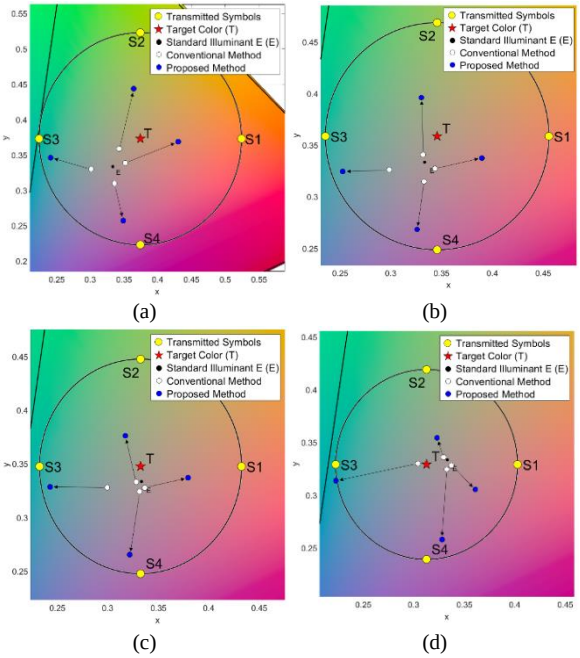


Fig. 10. Selected representative symbol color for different target color; for standard illuminant (a) F11, (b) D50, (c) D55, and (d) D65.

V. Results and Discussion

After capturing video, the demodulation is performed frame by frame. We set 3 clusters into the algorithm of [5] to select the representative color of the corresponding received symbol color for deciding the corresponding transmitted symbol color. Fig. 9 shows a sample case of the selected representative symbol color for the corresponding transmitted symbol color by using algorithm [5]. We also experiment it for different target color. We select F11, D50, D55, and D65 as target color because they are standard illuminants. The selected representative color symbol is also close to the corresponding transmitted symbol color for these target colors. Fig. 10 shows the selected representative symbol color for different target color.

VI. Conclusion

At last, we can claim that we successfully demonstrate the visual-MIMO based pulse oximeter system which is very useful for healthcare application. This type of LED to camera communication can reduce the extra installation and maintenance cost of RF device. Hospitals can easily install this type of system for monitoring the patients centrally and remotely. Moreover, the representative symbol color selection algorithm helps to reduce the color saturation noise in terms of transmitted symbol decision. The HR and BO are successfully transmitted through LED array by using GCM technique and receiver successfully demodulate the information data into the developed smartphone app. In future, we will measure the symbol error rate (SER) or bit error rate (BER) of this type of system by increasing the distance between transmitter and receiver.

Acknowledgment

This research was supported by the National Research Foundation of Korea Grant funded by the Ministry of Science, ICT, Future Planning [2015R1A5A7037615].

References

[1] H. Furuhata, "Electromagnetic interferences of electric medical equipment from hand-held radiocommunication equipment," 1999 International Symposium on Electromagnetic Compatibility (IEEE Cat. No.99EX147), May 1999, pp. 468–471.

[2] M. Z. Chowdhury, M. T. Hossan, A. Islam, and Y. M. Jang, "A comparative survey of optical wireless technologies: architectures and applications," in *IEEE Access*, vol. 6, pp. 9819-9840, January 2018.

[3] J.-S. Kim, T.-H. Kwon, and K.-D. Kim, "Variable intensity-on–off keying modulation and its application to wearable visual-MIMO," The 22nd Asia-Pacific Conference on Communications (APCC), Yogyakarta, Indonesia, August 2016, pp. 400-403.

[4] T.-H. Kwon, J.-E. Kim, J.-H. Oh, and K.-D. Kim, "Implementation of color-space-based wearable visual-MIMO System," Ninth International Conference on Ubiquitous and Future Networks (ICUFN), Milan, Italy, Juy 2017, pp. 776-779.

[5] K.-S. An and K.-D. Kim, "Performance improvement of symbol decision for Visual-MIMO system in RGB LED lighting," International Conference on Sensor Networks and Signal Processing (SNSP), Xi'an, China, China, October 2018, pp. 365-369.

[6] P. Das, B.-Y. Kim, Y. Park, and K.-D. Kim, "Color-independent VLC based on a color space without sending target color information," in *Optics Communications*, vol. 286, pp. 69–73, 2013.

[7] P. P. Banik, R. Saha, and K.-D. Kim, "Regression analysis for LED color detection of visual-MIMO system," in *Optics Communications*, vol. 413, pp. 121-130, April 2018, doi: 10.1016/j.optcom.2017.11.086.

[8] T. J. Atherton and D. J. Kerbyson, "Size invariant circle detection," in *Image and Vision Computing*, vol. 17, pp. 795-803, no. 11, September 1999, doi:10.1016/S0262-8856(98)00160-7.

[9] Mayo Clinic, "Hypoxemia," Mayo Foundation for Medical Education and Research (MFMER), 2019. [Online] Available: https://www.mayoclinic.org/symptoms/hypoxemia/basics/definition/sym-20050930. [Accessed: 05-Mar-2019].

RoboFI: Autonomous Path Follower Robot for Human Body Detection and Geolocalization for Search and Rescue Missions using Computer Vision and IoT

Ahmed Imteaj[1], M A Isfar Jubair Chowdhury[2], Mohammad Farshid[3], Abdur R Shahid[4]
[1,4]School of Computing and Information Sciences, Florida International University, Miami, FL 33199, USA
[2,3]Department of Computer Science and Engineering, [2,3]International Islamic University Chittagong, Bangladesh
Email: aimte001@fiu.edu, ishaar27@gmail.com, mdfar42@gmail.com, ashah044@fiu.edu

Abstract— **Numerous loss of lives is seen each year and the sum is consistently expanding because of destruction, fiasco, cataclysm and so forth these incorporate natural and man-made fiascoes. To counter this we have structured a rescue robot that is autonomous, it pursues track way by dint of 3 ultrasonic sensors in the front. For the processing work Raspberry Pi Model B is utilized which incorporates a grouping of single board processors. The smaller card size is handful as we are utilizing 4wd Car Chassis and a motor controller to drive it through. The robot features computer vision and raspberry pi camera module is used to help with the vision. We trained the camera utilizing algorithm to empower it identify movement, consequently after detecting movement the robot utilizes PIR to detect and affirm human presence. When human life is recognized in a destructed place the web application helps us to observe the live streaming of that found human and the closest environment. The coordination of GPS Module gives us a chance to see the area of where rescue activity is occurring. The web application that is designed by us likewise gives us a chance to do as such. After rescuing the detected individual, it goes on its merry way in repeating the flow of work that is assigned to it helping achieve maximum amount of success and lives saved. The reality to stun is, all these are incorporated in a solitary car chassis and the expense behind the work is insignificant. The reason is utilizing raspberry pi and our work to make a pi camera module ready to recognize movement by help of raspberry pi while incredible and expensive processors are required by and large.**

Keywords—Human Detection; Raspberry Pi; Autonomous Robot; Obstacle Avoidance; Raspberry pi Camera Module; Live Streaming; Location.

I. INTRODUCTION

At this moment in time People are fronting different sorts of disaster. The types and the damage these disasters can cause are various. Typically, in Earthquake and Wartime individuals do not possess any awareness about the places where injured people can be stuck. So for identifying the places and the humans who can be in danger, we built a tiny Robot which can detect human presence and also enrich us with the nature of the places where natural calamity or disaster is faced.

The core idea behind this work is to aid the rescue teams with an added component which can come more than handy in rescue operations. As a developing country we lack the kind of state-ofthe-art resources that the developed countries use to conduct their rescue operations. In countless number of times the cost to import or use these kinds of rescue robot is beyond our ease. Hence, we are working on a tiny rescue robot, that is autonomous and it is cost effective and efficient at the same time.

We researched in Embedded Engineering and Robotics, used Image and Real Time video recognition and Internet of Things (IoT). Rescue Robotics is something which is at the core of this work. A Rescue robot is a machine that is aimed for rescuing people. Rescue robots are used in situations such as mining mishaps, urban catastrophes, hostage circumstances, and explosions.

So, our robot is fortified with the functionalities of a rescue robot yet it is very little in size making it proficient of gathering information in difficult circumstances. We are using RPI3 as the main processor of the robot and the rest of the components will be interfaced with it. The main unit is this, with several sub-components such as GPS Module, GSM module and Pi Camera interfaced with the Raspberry Pi 3. The stand out feature is the inclusion of Pi camera which enables us to process images and videos in real time. Motors are used to make the robot move from one place to another. Power is supplied to the robot through power bank.

II. RELATED WORKS

Surveillance system requires vision to monitor, server and network to stream[2]. Raspberry Pi for processing, USB Camera for recording, simple cv for motion detection and MPJG streamer for streaming is used in this work. Our system uses Raspberry Pi Model B with Raspberry Pi camera module powered by OpenCV for motion detection. We have used PIR as double layer assurance of human detection. Added advantage of our system is, it is a robot capable of contributing in rescue missions compared to static surveillance system. [1] Raspberry pi is used for educational purposes where students were asked to build a system having raspberry pi and image processing programs. Energy saving, and low-cost setup are the unique features of this work that makes it stand out. Remotely controlled Autonomous systems are used to aid the parents monitor children from distance[3]. It requires human intervention and android application to operate and observe. RoboFI is Autonomous so it does not require human intervention. Web Application developed using Asp.net MVC is used to share location and streaming information. Human detection is done in a number of ways. Neural network with doppler radar is used to identify human. Which is a complex procedure[4]. We have opted for OpenCV and background subtraction algorithm for motion detection. Passive Infrared sensor makes sure that identification is done in a proper manner.

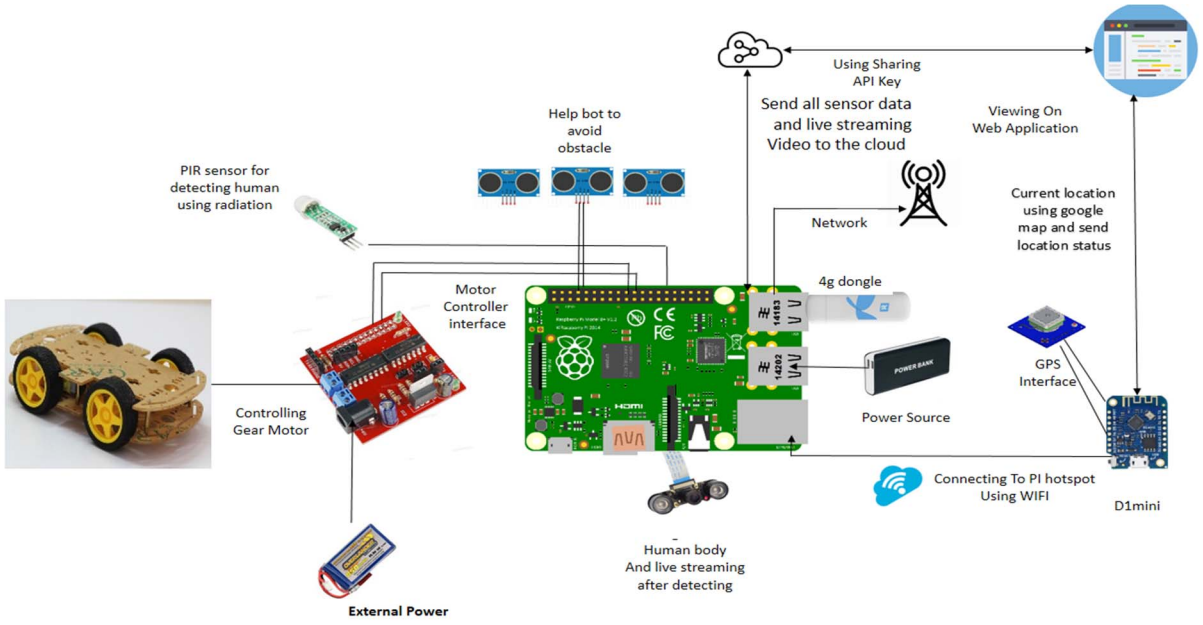


Fig. 1.System Architecture.

Obstacle avoidance is a major aspect of robotics. Various sensors such as visual, ultrasonic, infrared etc. are used to avoid obstacles. Ultrasonic sensor is used by them [5] for obstacle avoidance which is similar to ours. The system is capable of recognizing objects. RoboFI ensures human detection along with location and streaming of the human that we are looking to rescue. Surveillance systems require vast memory generally as it record and transmits video continuously. [6] OpenCV is utilized in order to detect motion using USB HD camera and it starts and also stops recording video based on motion. That ensures that minimum amount of space is consumed yet security is maintained. [7] It illustrates learning based approach for human detection using RGB-D images extracted from Microsoft Kinect sensors and it is indeed a challenge to detect human accurately. Depth sensing feature of these sensors is used in this work. But, they are quite costly to implement a system. Our sytem uses a more feasible approach of human detection. [8] Provides an in depth analysis of how to construct an autonomous robot. It [9] uses two cameras for vision and MATLAB for tracking and mapping. Rescue operations require adaptive rescuing. [10] uses Dijkstras shortest path algorithm after locating victims, to rescue them properly.

III. SYSTEM DESCRIPTION

A. System Architecture

Raspberry Pi 3 Model B is the core unit along with Raspberry Pi camera module powered by OpenCV to conduct the process of human detection. A couple of sensors namely Ultrasonic sensors and Passive Infrared sensor are used. These are placed on a 4wd car chassis which is driven by a motor controller. D1 mini GPS Module provides the location information of RoboFI. To view the streaming live Asp. net web application is developed by us that also plots the location in google maps by using the latitude and longitude data that is transmitted over the cloud (Thingspeak server) by D1 mini. Human is detected through vision and passive infrared sensor reassures the presence. Obstacle avoidance is achieved by using three ultrasonic sensors in the front of the robot. External power for raspberry pi is managed by using a power bank while a lithium polymer battery powers motor controller.

B. Logical Description

The system at first needs to detect any obstacles nearby and avoid them upon starting to move. For this, there are three ultrasonic sensors installed in the front side of the robot as stated in the system architecture. Three conditions for three sensors are used in order to avoid the obstacles. The obstacle can either be in front or on the right or left. So each of the cases are covered by us as illustrated in the logic diagram. The basics steps of distance measurement are:

1. Sonar[Trig, Echo]:
    a. Trig = 0
    b. time.sleep(.1)
    c. Trig = 1
    d. time.sleep(.0001)
    e. while:
        i. echo = 0
        ii. pulse start()
    f. while:
        i. echo = 1
        ii. pulse end()
    g. distance = pulse end() - pulse start()
    h. distance = round(distance,2)
    i. return distance
2. Sonar 1= Sonar[Trig1,Echo1]
3. Sonar 2= Sonar[Trig2,Echo2]
4. Sonar 3= Sonar[Trig3,Echo3]

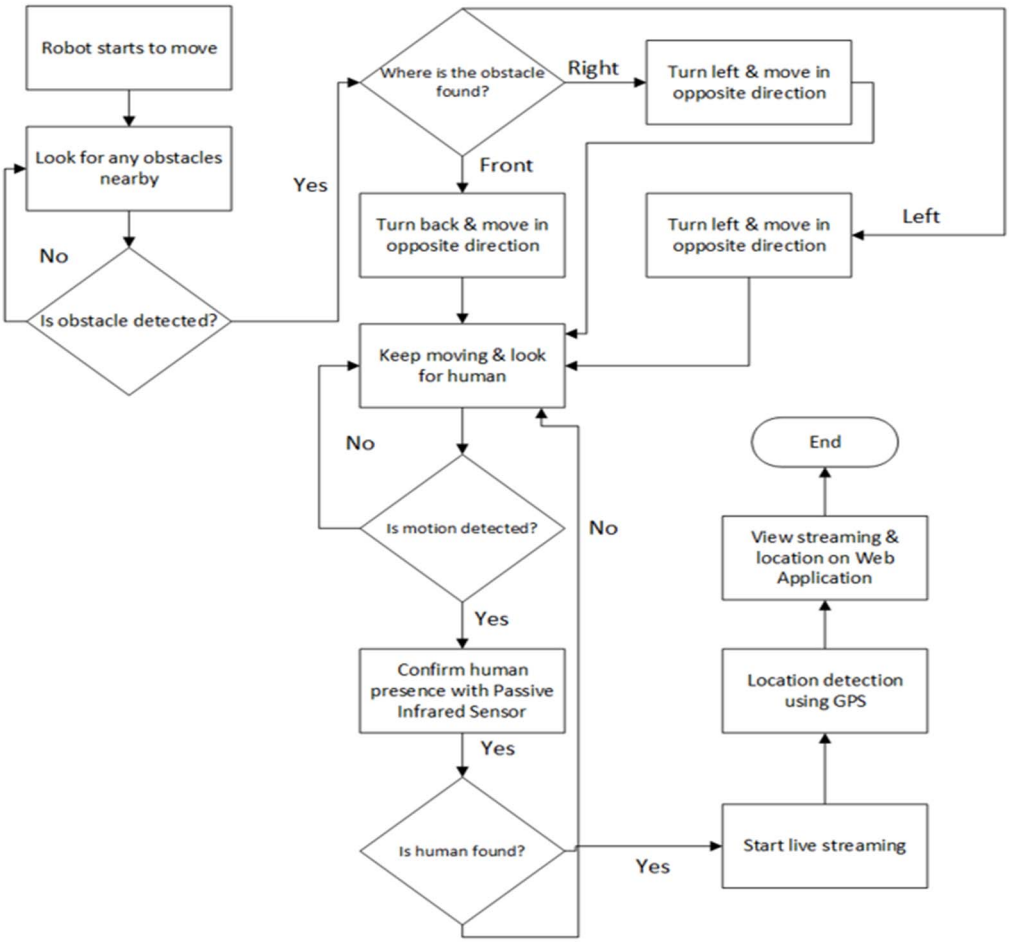


Fig. 2. Logic diagram of the system

The first task of this robot is to find a suitable path. Three ultrasonic sensors are used to do so. All measurements for the work will be considered based on the average of the values found from the three ultrasonic sensors.

The formula to find the average distance from all of the ultrasonic sensors is:

- average = (sonar1 + sonar2 + sonar3)/3

If the robot is switched on and placed somewhere, it starts looking for a path.

1. If the robot finds a path – (average>50 and all sonar value>45)

a) It traverses along the distance measured for that path.

b) After traversing, it starts checking for further path to move ahead and detect probable motion.

c) On detecting motion, it uses PIR sensor to assure human presence.

2. If it doesn't find a path or the path is blocked – (average <40)

a) It takes a turn of 180 degrees to the right. Then it looks up again at this position for navigation track.

b) If it finds no track, RoboFI starts rotating by 10 degrees to the left and keeps repeating the same task until it gets back to the starting position.

c) Coming back to the starting spot, it again rotates 180 degrees to the left to check for availability of path.

3. If the path is blocked and the distance to the obstacle is very low (average<10) then

It moves backward and repeats step 2.

All of these procedures are carried out based on the average of the values from the three ultrasonic sensors.

After the obstacle is successfully avoided RoboFI moves ahead in quest of human. To identify human we have used raspberry pi camera module and OpenCV. It allows us to use several built in libraries related to image processing and we used the method:

- Initialize the raspberry pi camera module.
- At the very beginning, capture a frame as background model that is the template image.
- Convert that frame into gray scale.

- Compute the difference between the frame that comes next and the background model.
- From the difference we find a threshold value.
- The average background image is then updated with the weighted running average.
- Noise reduction is done by thresholding the difference frame.
- Fit holes by dilating the threshold frame.
- Detection of the Contours.
- The motion is located by finding out the bounding boxes around the contours.

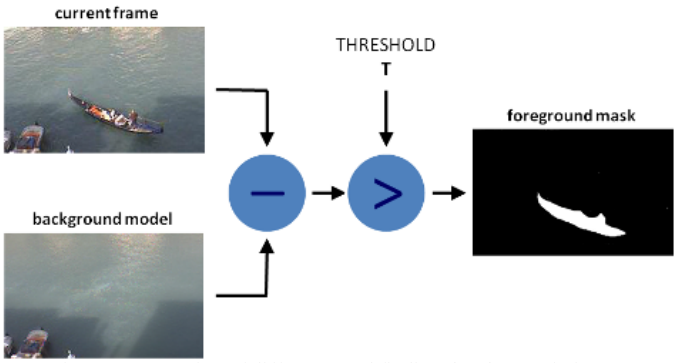


Fig. 3. Computation of difference and finding the changes in image

Once camera makes sure of human presence, Passive infrared sensor is used to reassure the presence. It is made to work as follows:

This sensor working principal depend on the motion. If any movement happen, it will start it's work, and the work is detecting Human.

- It will track 10 time for human searching. If it gets high, then human is there, and this recommends bot to start streaming.
- And if low then no human found, and this recommend bot to leave place. And the bot will move for others area.

After successful identification of human streaming of visulas begin on the asp.net web application. Location of the human found to rescue is also shown on the web application with the help of D1 mini and GPS Module which transmits location information to Thingspeak server and we fetch the location from cloud and plot on google maps in our web application.

IV. SYSTEM IMPLEMENTATION

To implement the system we are using two micro-controllers Raspberry Pi 3 model B and D1 mini arduino. Raspberry pi is the core processing unit. Raspberry pi camera module is interfaced with it using a CSI flex cable. By dint of the vision from camera we use OpenCV powered by python to conduct the process of human identification. Passive infrared sensor is connected to the raspberry pi to enable reassurance of human detection. These are placed on a 4wd car chassis which is controlled by a motor controller. To control the car and make sure that any obstacle is avoided in the track path three ultrasonic sensors are conncted to raspberry pi and placed in the front.

The motherboard named Raspberry Pi 3 acting as a Client is both a micro-controller and a CPU which has a processor of 1.2Ghz 64-bit quad-core ARMv8 Cortex A53 CPU alone with 1GB ram of 900Mhz, 4 USB ports, 1 HDMI port, 1 audio I/O port and 1 Ethernet port in it, including 40 GPIO pins which can be configured as digital input or output. The board Raspberry Pi 3 has built in wireless module in it that has both 802.11n wireless LAN and Bluetooth 4.1 including BLE feature by which we can both receive and transmit data wirelessly without any other partial device. We use these features to implement our system. After the initial phahses upon detection of human we are using D1 mini and GPS module in order to transmit location data to the cloud. The sreaming is done through the Ip address of raspberry pi using StreamingHandler and StreamingOutput library functions in OpenCV.

A. Steps of Implementation:

Step 1 (Configure the Programmable Devices): Seperately write code for each programmable device, raspberry pi and Arduino. Code for raspberry pi is written in python using OpenCV which is the library used for image processing and computer vision. Similarly, we wrote code in Arduino ide to extract location from gps module and then for transmitting it into thingspeak server.

Step 2 (Integrating the sensor modules): The ultrasonic sensors, passive infrared sensor and raspberry pi camera module all are integrated with raspberry pi. GPS module is integrated with D1 mini Arduino.

Step 3 (Connecting the motor controller): Motor Controller is connected with the motor of the chassis. Then In1, In2, In3 and In4 pin are connected to the GPIO pin of raspberry pi.

Step 4 (D1mini and gps Setup): The neo-6 gps module is setup with D1 mini. Neo-6 gps has 4 pins and an antenna. Antena is connected with neo-6 gps and vcc and gnd pin is connected with d1 mini 5v and gnd pin. Rx is connected with D3 pin and tx is connected with D5 pin. We have written a code for sending latitude and longitude data to the thingspeak through d1 mini and d1 mini is connected to the raspberry pi 3 using hotspot.

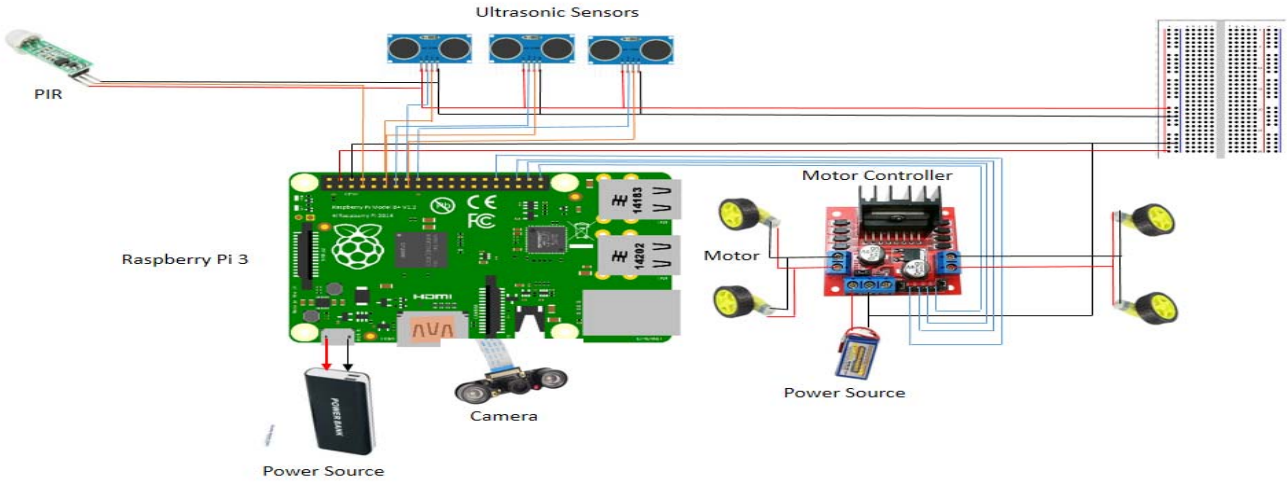


Fig. 4.Circuit diagram of the system.

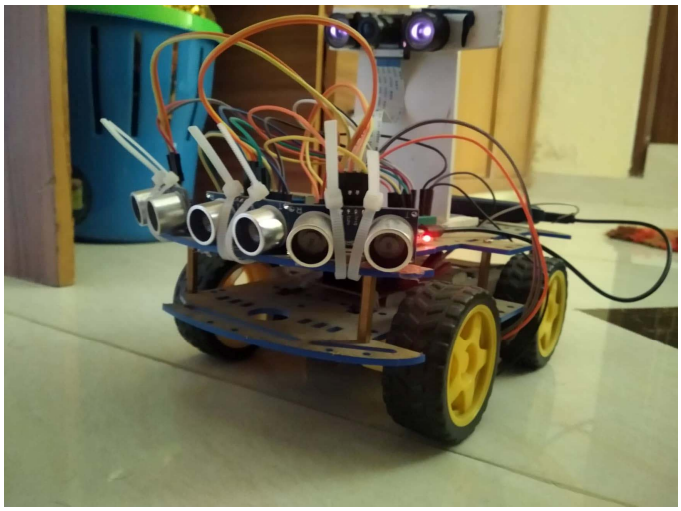


Fig. 5.RoboFI after implementation

Step 5 (Setting the whole system up in chassis): Next, we set the whole system on wooden chassis board. In this chassis we setup the raspberry pi along with camera module, three ultrasonic on front, pir also in front and motor controller. We set the four motors in the wooden chassis. We also setup the gps system on this board.

Step 6 (Powering up all the devices): In our system, the main devices that will be directly connected to the power sources are the Raspberry Pi 3 and Motor controller. Raspberry Pi needs to be connected with a 5v 2.5 amp. power supply powerbank. Motor controller needs external power. This is the reason that we set a lithium polymer battery to the motor controller. Camera Module is directly connected to camera space of raspberry Pi. The Three Ultrasonic and Pir needs exactly 5 volt and this power is coming from Raspberry Pi. D1 mini is connected in pi's usb and from this it gets the power. GPS is connected with the d1 mini's 5 v and GND pin as stated earlier.

V. EXPERIMENTED RESULT

After the system is implemented we run the robot and this provides the data of human detection, confirmation with passive infrared sensor, obstacle avoidance data from ultrasonic sensors, live streaming visulals and mapping of location in web application. All of these and efficiency and precision of the system is illustrated in this section. The results are provided here respectively. Firstly, different situations are provided while detecting human, confirmation with PIR and live streaming follows. The distance that is measured by ultrasonic sensors is also shown in this section along with the plotting in map using google maps api key and using our web application.

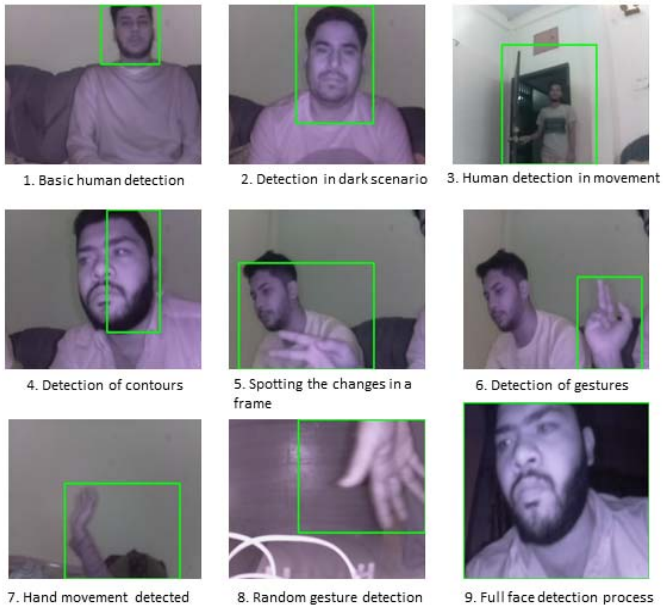


Fig. 6.Human detection in various situations

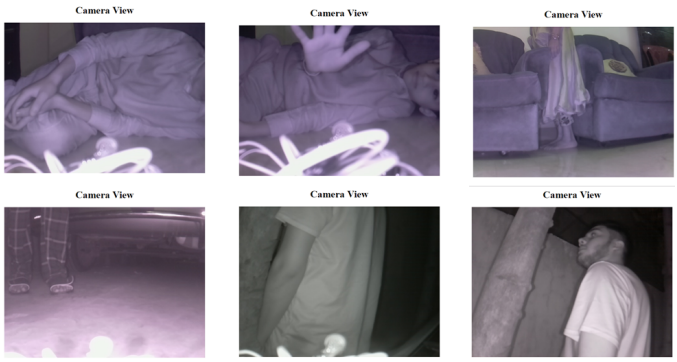


Fig. 7.Live streaming of detected humans

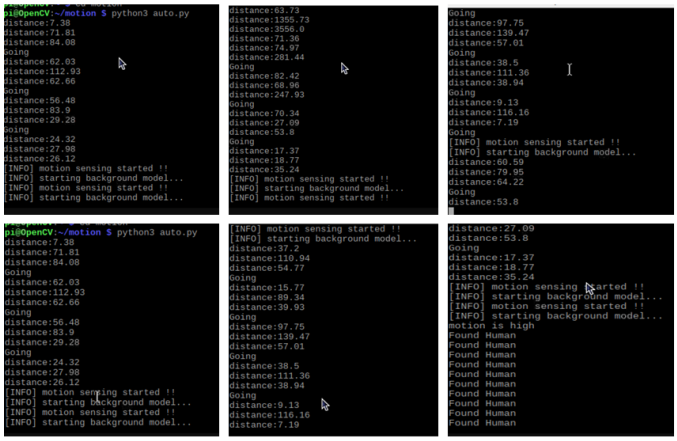


Fig. 8.Distance measurement for obstacle avoidance

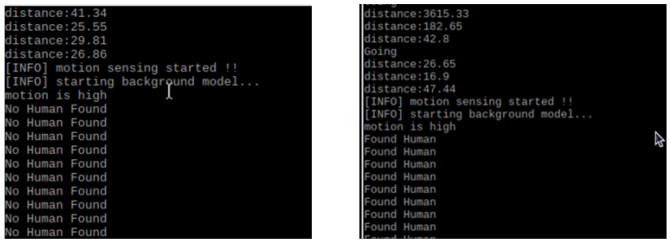


Fig. 9.PIR data when human is found and also when not found

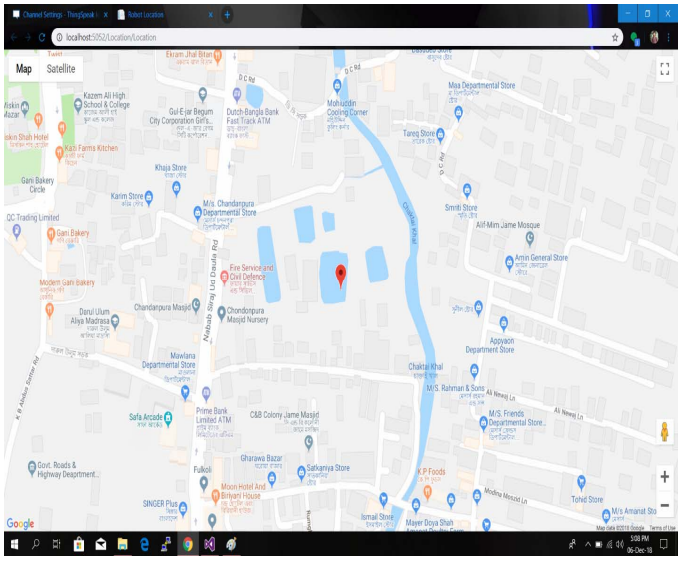


Fig. 10.Location information in web application

Fig. 11 depicts the time elapsed in traversing to human avoiding obstacles and accuracy of the detection. In the figure high or 1 signifies that the human was detected properly compared to low or 0 signifying inaccurate detection of human through random iterations of RoboFI.

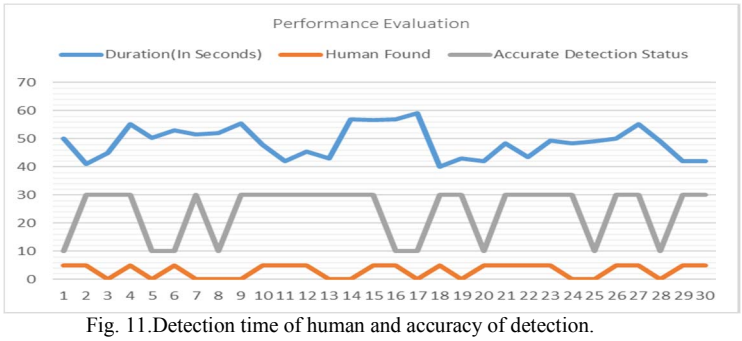


Fig. 11.Detection time of human and accuracy of detection.

CONCLUSION

In this paper, we illustrated our work on autonomous rescue robotics capable of human detection and Geolocalization. Couple of sensors and their working on our work is also discussed in this paper. Succesful use of this, can lead to a significant amount of success in search and rescue missions with respect to reduction in detection and locating. Saving people in time during crisis remains to be the motto of RoboFI, the autonomous path follower robot.

REFERENCES

[1] Marot, Julien, and Salah Bourennane. "Raspberry Pi for image processing education." Signal Processing Conference (EUSIPCO), 2017 25th European. IEEE, 2017.

[2] Menezes, Virginia, Vamsikrishna Patchava, and M. Surya Deekshith Gupta. "Surveillance and monitoring system using Raspberry Pi and SimpleCV." Green Computing and Internet of Things (ICGCIoT), 2015 International Conference on. IEEE, 2015.

[3] Pujari, Sachin S., Mahadev S. Patil, and Subodh S. Ingleshwar. "Remotely controlled autonomous robot using Android application." I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2017 International Conference on. IEEE, 2017.

[4] Kwon, Jihoon, and Nojun Kwak. "Human detection by neural networks using a low-cost short-range Doppler radar sensor." Radar Conference (RadarConf), 2017 IEEE. IEEE, 2017.

[5] Balasubramanian, Karthi, et al. "Object recognition and obstacle avoidance robot." Control and Decision Conference, 2009. CCDC'09. Chinese. IEEE, 2009.

[6] Murugan, KH Shakthi, V. Jacintha, and S. Agnes Shifani. "Security system using raspberry Pi." Science Technology Engineering and Management (ICONSTEM), 2017 Third International Conference on. IEEE, 2017.

[7] Santoso, Patrisia Sherryl, and Hsueh-Ming Hang. "Learning-based human detection applied to RGB-D images." Image Processing (ICIP), 2017 IEEE International Conference on. IEEE, 2017.

[8] Obiozor, C. N., T. A. Giuma, G. Rowlands, T. Dinkins, J. Brian, and A. Wilson. "Construction of an autonomous robot." In Southeastcon 2000. Proceedings of the IEEE, pp. 213-216. IEEE, 2000.

[9] Kho, Yau Hee, Aminath Elaf Abdulla, and Jessica Chan Zhai Yan. "A vision-based autonomous vehicle tracking robot platform." Industrial Electronics & Applications (ISIEA), 2014 IEEE Symposium on. IEEE, 2014.

[10] Kiyani, Midhat Noor, and Muhammad Uzair Masood Khan. "A prototype of search and rescue robot." Robotics and Artificial Intelligence (ICRAI), 2016 2nd International Conference on. IEEE, 2016.

Development of Microcontroller based Light Following Emergency Evacuation Chair

Delowar Hossain, Hasan Alam, Ajmain Ar Rafi, Kamruzzaman Rakib and Umma Hany

Department of Electrical and Electronic Engineering, Bangladesh Army International University of Science and Technology, Cumilla-3501, Bangladesh, E-mail:delowarhossain.baiust@gmail.com

***Abstract*— Emergency evacuation is required for fast evacuation of people during disaster. In this paper, we develop a light following emergency evacuation chair to rescue and evacuate physically weak or disable people. The light following evacuation chair includes proximity detection and obstacle avoidance features in order to avoid collision. The chair follows the light sources fixed at different turn points in the Emergency exit track to determine the route and direction to the Emergency exit. We develop the Light Follower using light dependent resistors (LDRs). A motor driver circuit is used to control the direction of the chair depending on response of the LDRs. We use ultrasonic sensor and the motor driver to develop the proximity detector and obstacle avoidance system. A micro-controller is used to control the operation of LDRs, untrasonic sensor and the motor driver. We simulate the performance of the Evacuation chair in forward, left and right direction using an Emergency exit route model and verify its accuracy for different angle of arrival (AoA) of the incident light. We observe an average movement accuracy of 93% in all directions for all AoA of light.**

Keywords*— *Emergency Exit, Emergency Evacuation Robot, Light Following Robot.

I. INTRODUCTION

An Emergency exit is a special exit for dangerous scenarios during emergencies when people are required to be evacuated from a location or building quickly and it also provides an alternative if the route to the regular exit is blocked [1]. People does not have enough time and idea on navigation of the emergency exit path. Thus, the existing system such as emergency exit signs and evacuation maps, as well as alarm and audio announcements are installed to guide people to find the way of the exit. However, people with disabilities or physical weakness in a large venue as hospitals, business centers, super stores or shopping malls are unable to escape from the venue or walk down steps in order to reach a safe area.

The Disability Discrimination Act (DDA) 1995 and Equality Act 2010 works to protect people with disabilities including blind and partially sighted people from discrimination [2]. In order to comply with the DDA, it is the responsibility of different institutions to install evacuation systems which should be autonomous and efficient to rescue and evacuate such people with disability from the site immediately during the fire or other dangerous scenarios. Robots can play an important role in such emergency evacuation. Robots can be used for navigation task and for evacuation of people with disability from the site.

There are several works in the literature to develop evacuation systems using different sensors, wireless sensor network and robots to evacuate people. In [3], the authors propose an approach for detecting alive humans in destructed environments using an autonomous robot. The proposed robot in [3] uses an ultrasonic sensor in order to detect the existence of living humans and a low-cost camera in order to acquire a video of the scene as needed. Additionally, the robot includes sensors as temperature, fire and metal detector to detect the presence of bomb in warfield and in rescue operations. In [4], the authors develop a multiple source multiple destination Robot, which could differentiate various colors by using sensors and choose a desired line to reach the target location. It could also detect presence of obstacle using close loop control and correction system. In [5], the authors design a PC controlled robot which can detect live humans and transmit the location details wirelessly. In [5], PIR sensors are used to detect the motion of any kind, GSM is used for wireless communication and GPS is used for location estimation. In [6] the authors discuss the disastrous effect of overtrust on robot in emergency evacuation scenarios. In [7], the authors discuss current robotic technologies of crowd evacuation and also demonstrate how autonomous robots are effectively deployed for evacuation assistance task in disaster, as well as search and rescue missions. However, the robotic evacuation technologies presented in [3-7] do not address the evacuation issue of the patients or people with disability during emergency.

In this paper, we develop a micro-controller based Emergency evacuation chair to evacuate people with disability. The developed system uses a light following navigation system to guide the chair to the Emergency exit route. To implement the light following navigation system, the light sources are installed in different turn points in the Emergency exit route and light dependent resistors (LDRs) are used to sense the lights emitted from the light sources. A micro-controller unit (MCU) is used to control the operation of LDRs, ultrasonic sensor and the motor driver according to the instructions loaded. Light sources are installed at different points along the pathway to direct the chair to its destination. The light is sensed by three LDRs and the motor driver circuit is activated based on the response of LDRs by the MCU to move the wheels of the chair according to the light sensed by the front, left or right LDRs. We also use ultrasonic sensor to detect proximity of obstacle. When the ultrasonic sensors sense obstacle, the center LDR signal is turned off and the wheels stop moving to stop forward movement. We simulate the hardware and conduct experiment using an emergency exit route model to verify the performance of the light following emergency evacuation chair in forward, left and right direction. We also verify the performance of the robot by varying the angle of arrival of light and observe 93% average accuracy in all directions.

II. SYSTEM LAYOUT

Figure 1 shows the system layout of the emergency evacuation robot. As we can see from Fig. 1, the robot includes

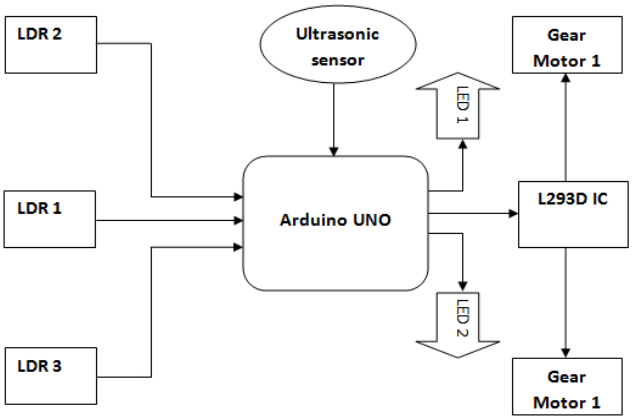


Fig. 1 System layout

three (03) LDR sensors which receive light and generate signals as it can sense light emitted from the light sources. LDR1 is used to sense light received from the front, LDR2 is used to sense light coming from right side and LDR3 is used to sense light received from left side. The system also includes an ultrasonic sensor to sense the proximity of any obstacle and avoid the collision. LDRs and the ultrasonic sensor sends their response signals to the Arduino UNO microcontroller unit (MCU). Based on the received signals from LDRs and ultrasonic sensors, the Arduino UNO operates the LEDs and the Motor driver IC L293D according to the instructions loaded in it. The right LED is lit when LDR2 receives light and Gear motor 1 is activated. The left LED is lit when LDR3 receives light and Gear Motor 2 is activated. The LEDs show in which direction the robot will go by blinking and the motor driver will drive the two Gear motor according to the signal received from Arduino UNO.

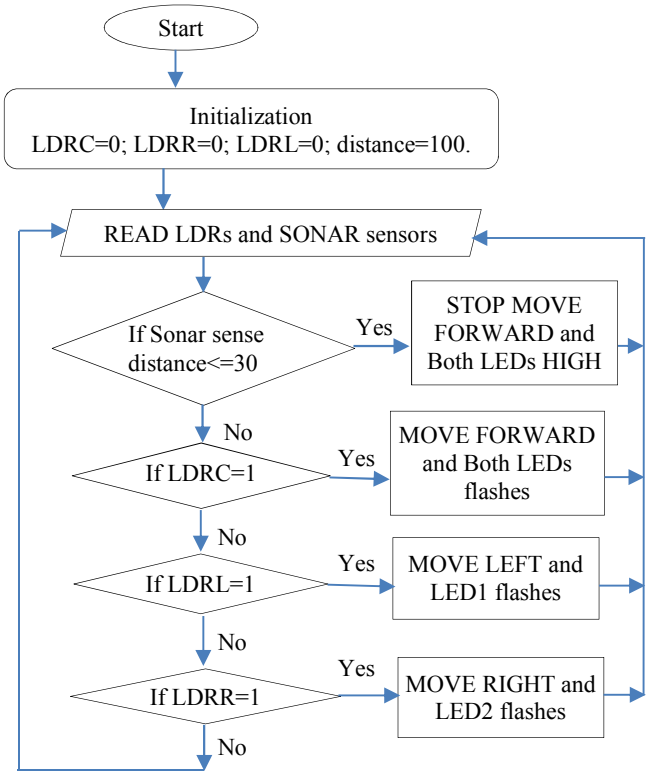


Fig. 2 Flow chart of emergency evacuation chair.

III. System Development of the Evacuation Chair

The Emergency Evacuation chair works using the working principle as shown in the flow chart in Fig. 2. We can see from the flow chart that the MCU always checks for response of the sensors. Whenever it receives any signal from LDRs or sonar sensor, it sends the corresponding signal to the motor driver for necessary movement action. The MCU is programmed using the language of Arduino 1.8.3 software.

TABLE I
LIST OF COMPONENTS AND THE COST OF DEVELOPMENT

Sl	Components	Quan-tity	Price ($)
1.	Arduino Uno R3	1	9.0
2.	L293D Dual DC Motor Controller	1	3.25
3.	Ultra-sonic Sensor (HC-SR04)	1	3.00
4.	LDR	3	4.50
5.	Robotic Chassis (2 Wheel with Gear Motors)	1	15
6.	Mini Breadboard	1	.50
7.	Battery 9V	1	3
8.	Jumper	30	1.5
9.	9V Battery Connector	1	0.10
Total			39.85

Using the above software development step, we develop the Emergency Evacuation Robot prototype as shown in Fig. 3. The components used to develop the core hardware of the system is listed in Table I which costs approximately $39.85. As we can see from Fig. 3 that the Evacuation system is developed in a shape of a chair. Thus, we call it Emergency Evacuation chair which can easily carry people with disability. The chair is powered on when people sit on this. The movement and direction of movement of the chair is guided by the light sources fixed on the Evacuation route. In this way, the chair can be used to rescue people and reach them to the exit point by following lights. The system consists of several sub-systems. The step by step development of each sub-system is explained as in the following

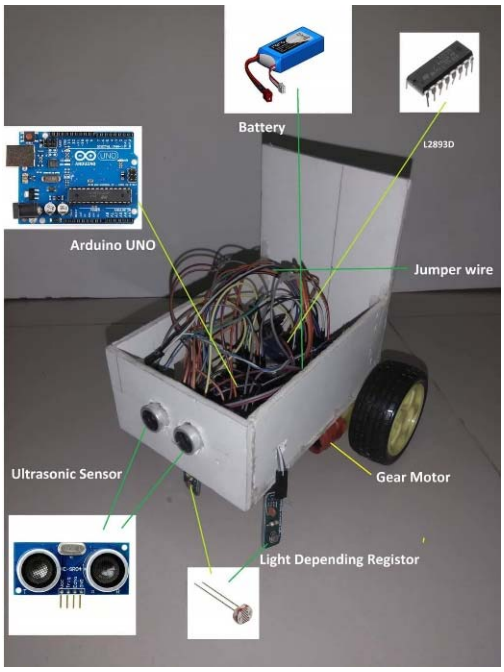


Fig. 3 Emergency evacuation chair.

A. Light Follower Circuit

The Light follower is developed using light sources and light dependent resistors (LDRs). The light sources are installed in different turn points in the track. Photo resistors, also known as LDR, are light sensitive devices most often used to indicate the presence or absence of light, or to measure the light intensity. In the dark, their resistance is very high, sometimes up to 1MΩ, but when the LDR sensor is exposed to light, the resistance drops dramatically, even down to a few ohms, depending on the light intensity [8]. Due to resistance variations, the voltage levels are also varied between 0 and 5V using voltage divider rule. The robot follows the light emitted from the sources by means of the central LDR (LDR1), the right LDR (LDR2) and the left LDR (LDR3) sensors. Figure 4 shows the connection diagram of LDRs and Arduino UNO.

Due to light variations if the resistance of LDR3 is dropped to R, then the voltage across R1 which is connected to Pin A0 of MCU can be expressed using voltage divider rule as,

$$V_{R1} = V_{total} \times \frac{R1}{R+R1} \quad (1)$$

Similarly the voltage across R2 and R3 can be expressed as follows

$$V_{R2} = V_{total} \times \frac{R2}{R+R2} \quad (2)$$

$$V_{R3} = V_{total} \times \frac{R3}{R+R3} \quad (3)$$

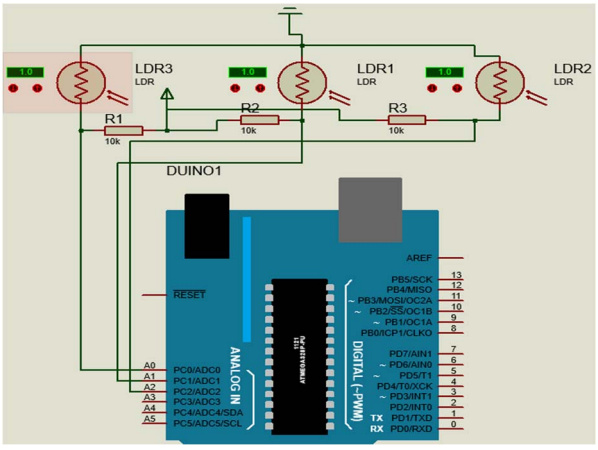


Fig. 4 Schematic diagram of Light Follower Circuit.

B. Obstacle Avoider

The ultrasonic sensor (HC-SR04) uses sonar to determine the distance of an object, just like bats. It offers excellent contactless range detection with high accuracy from 2 cm to 80 cm. The operation of the sensor is not affected by sunlight or black materials as sharp telemeters. It has four pins for VCC, GND, Trigger and Echo. The sensor works with the following formula

$$\text{Distance} = \text{Speed} \times \text{Time} \quad (4)$$

The untrasonic transmitter (trigger pin) sends a high frequency sound signal. When the signal finds an object, it is reflected and the untrasonic receiver (echo pin) receives it as shown in Fig. 5.

To calculate the distance using the above formula, we need speed and time of traveling signal. The speed of untrasonic (US) wave at room conditions is 330m/s. The time can be calculated by measuring the time difference between transmitting and receiving US wave by the trigger pin and the echo pin. Then, we can simply calculate the distance using equation (4) using the microcontroller (MCU). The MCU of the evacuation chair can avoid collision by measuring the distance of obstacles (within 2-80 cm) and by operating the motor driver circuit according to stop movement of the chair. Figure 6 shows the connection diagram of the Arduino UNO and the ultrasonic sensor.



Fig. 5 Untrasonic sensor transmitter and receiver.

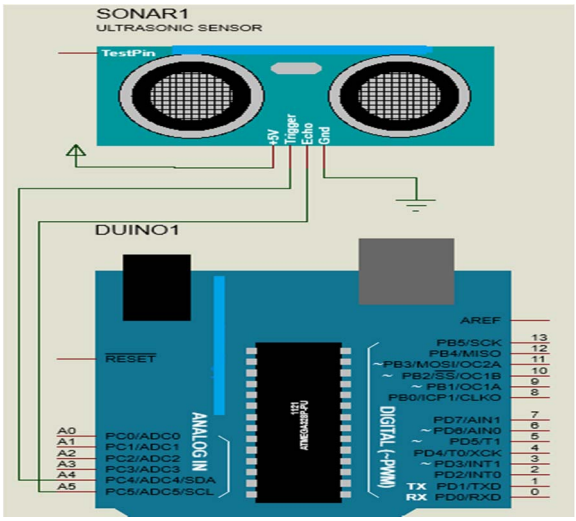


Fig. 6 Schematic diagram of Obstacle Avoider circuit.

C. Motor Driver Circuit

When the MCU gets signals due to light variations in LDR1, LDR2, LDR3 or the untrasonic sensor, the MCU sends signal to the motor driver circuit to move left, right, stop or forward actions accordingly. We use Motor driver IC L293D to operate motors of the two wheels of Evaculation chair.

It works on the concept of H-bridge. H-bridge is a circuit which allows the voltage to be flown in either direction. Voltage need to change its direction for being able to rotate the motor in clockwise or anticlockwise direction, In a single L293D chip there are two h-Bridge circuit inside the IC which can rotate two dc motor independently. Figure 7 shows the motor driver connections with Arduino UNO.

D. *LED Indicator circuit*

The LED indicator helps others to know the direction and movement of the chair to avoid accident or collision. As the chair stops, the MCU sends signal to the two back LEDs to flash continuously. As the chair turns left, the MCU sends signal to the left LED which flashes to indicate left movement. In the same way, the right LED flashes during right movement of chair. The LED indicators D1 and D2 are connected to the MCU Arduino UNO as shown at the bottom of Fig. 8 to indicate stop, left or right movements of the chair.

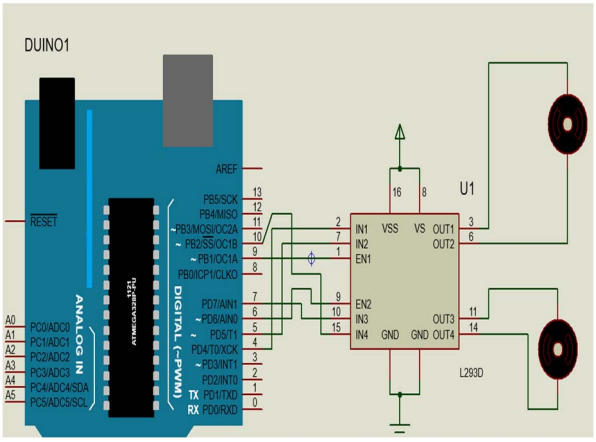


Fig. 7 Schematic diagram of motor driver circuit.

E. *Microcontroller Unit (MCU)*

An Arduino UNO R3 micro-controller (MCU) is used as the central processing unit of the system. The MCU receives the input signals from the sensors (LDRs and ultrasonic sensor) and sends the output to the motor driver and LEDs according to the instruction sets loaded in it according to the flow chart as shown in Fig. 2. The connection diagram of all components to the MCU is shown in Fig. 8.

The Evacuation chair is developed by combining all the above sub-systems. Figure 8 shows the combined circuit diagram of the whole system developed for Emergency Evacuation Chair.

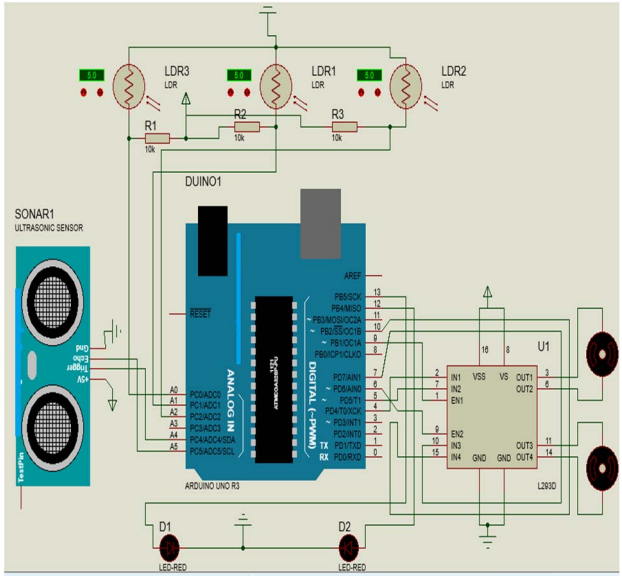


Fig. 8 Schematic diagram of the Emergency Evacuation chair.

IV. SIMULATION AND RESULTS

We simulate the performance of the Emergency Evacuation chair using the Exit Route model as shown in Fig. 9. When the Evacuation Chair travels from "Start" point, the LDRs continuously sense the presence of light. Light is emitted from the light sources indicated using two torches in the track as shown in Fig. 9. It detects the light emitted from the first light source in the track using LDR1 and move forward to follow the light until it detects proximity of any obstacle. It detects proximity by measuring distance of obstacle using untrasonic sensor. When it gets closer to the wall (<=30 cm), it stops moving and the back LEDs flashes. Then it detects the light emitted from the second source of light using LDR2 and turns "Right" to follow the light. During turning "Right", the right LED continues to blink. It continues to move forward in the right track until it detects any obstacle or wall. Then it stops again when it gets closer to the second wall at the end of right track and the back LEDs starts flashing. When it detects the light from second source of light fixed at the termination point using LDR3, it turns "Left" and the left LED continues blinking. Then it continues moving at the left track to follow the light set at the termination point. It moves forward until it reaches to the "Stop" or exit point of the Evacuation route.

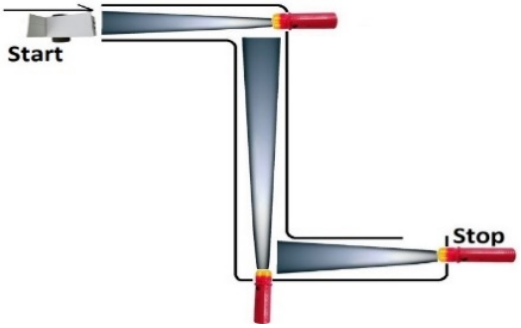


Fig. 9 Evacuation route model.

TABLE II
MOVEMENT ACCURACY IN FORWARD, LEFT AND RIGHT DIRECTION.

No. of exp.	Response of the LDRs and Motor Driver			Movement Accuracy		
	Forward	Left	Right	Forward	Left	Right
1	Yes	Yes	Yes	94%	87%	87%
2	Yes	Yes	Yes			
3	Yes	Yes	Yes			
4	Yes	Yes	No			
5	Yes	Yes	Yes			
6	No	No	Yes			
7	Yes	Yes	Yes			
8	Yes	Yes	No			
9	Yes	Yes	Yes			
10	Yes	Yes	Yes			
11	Yes	No	Yes			
12	Yes	Yes	Yes			
13	Yes	Yes	Yes			
14	Yes	Yes	Yes			
15	Yes	Yes	Yes			

To evaluate the performance of the Evacuation Chair, we simulate the system for different angle of arrival (AoA) of the light and evaluate the performance of the chair in forward, left

and right direction while traveling through the evacuation route model. Few of the simulation results for 220° AoA are summarized in Table II. In Table II, "Yes" means correct response and "No" means wrong response. We can see that the movement accuracy of the system in forward, left and right directions are 94%, 87% and 87%, respectively. Table III summarizes the movement accuracy considering different angle of arrival (AoA) of the light. As we can observe from the results in Table III that the average movement accuracy of the system in right, forward and left direction are 91%, 95% and 92%, respectively.

TABLE III
MOVEMENT ACCURACY FOR DIFFERENT ANGLE OF ARRIVAL.

Movement Direction	Angle of arrival	Movement Accuracy	Average Accuracy
Right Direction	220°	87%	
	315°	88%	91%
	330°	89%	
	345°	92%	
	360°	93%	
	30°	95%	
	10°	95%	
Forward Direction	25°	94%	95%
	40°	92%	
	55°	93%	
	70°	96%	
	85°	97%	
	220°	94%	
	115°	95%	
Left Direction	130°	92%	92%
	145°	92%	
	160°	94%	
	175°	96%	
	190°	94%	
	205°	90%	
	220°	87%	

From the practical simulation results, we observe that the Emergency evacuation chair can move in right, left and forward direction by following the light emitted from the source. The movement of the robot can be controlled in forward, left or right direction by using the light following LDRs. It follows light to reach to the exit point and can detect obstacle in the path of the robot to avoid collision. Therefore, it can move by avoiding damage caused by collision to an obstacle that makes it more robust and reliable. LED indicators helps the robot to avoid collision from backside and helps others to know the direction of the robot.

We can also see that the performance of the Evacuation chair is fairly accurate in all directions and for all AoA of the light. For any angle of arrival of the light, the movement accuracy of the chair is 95% in forward direction, 91% in right direction and 92% in left direction, i.e., in average the movement accuracy is 93% in all directions.

Thus, the developed Emergency Evacuation chair can be used to rescue and evacuate disable people from any place or hospital during emergency. It can carry disable people and reduce the risk of life of the patient and the rescue team.

V. CONCLUSION

In our paper, we have developed a Micro-controller based light following emergency evacuation chair. We have used three LDRs as light sensors to follow the source of light. We have used untrasonic sensor to detect obstacle in order to avoid collision. We have also used LED indicators to indicate the direction of movement of the chair. We have simulated the performance of the Evacuation chair for different angle of arrival of the light using an Emergency exit route model. We have observed that the evacuation chair can move in all directions following the light source fixed at different locations in the route. The performance of the chair in terms of movement accuracy is 93% in all directions for all AoA of light. Thus, the developed light following Emergency evacuation chair can be used to rescue people with disability and reduce the risk of their life by evacuating them from the location during emergency.

REFERENCES

[1] *"Emergency Exit"*, https://en.wikipedia.org/wiki/Emergency_exit, 24 July 2018.

[2] *"Disability Discrimination Act (DDA)",* Royal National Institute of Blind People (RNIB), UK, 1995.

[3] Bhatia, S., Dhillon, H.S. and Kumar, N., "*Alive human body detection system using an autonomous mobile rescue robot*," IEEE India Conference, December, 2011.

[4] K. M. Hasan, A. AI-Nahid, K. J. Reza, S. Khatun, and M. Basar, "*Sensor based Autonomous Color Line Follower Robot with Obstacle Avoidance*," IEEE Business Engineering and Industrial Applications Colloquium, pp. 598–603, 2013.

[5] Purnima G., Aravind S., Renju M. V.,Neethu A. M., Gayathri C. S., "*Alive Human Body Detection and Tracking System Using an Autonomous Pc Controlled Rescue Robot",* International Journal of Emerging Technology and Advanced Engineering, vol. 4, Issue 12, December 2014.

[6] Robinette, P., Li, W., Allen, R., Howard, A.M. and Wagner, A.R., *"Overtrust of robots in emergency evacuation scenarios*," ACM/IEEE International Conference on Human Robot Interaction, pp. 101-108, IEEE, March, 2016.

[7] Sakour, I. and Hu, H., "*Robot-assisted crowd evacuation under emergency situations: A survey*," Robotics, 6(2), p.8, 2017.

[8] *"Photo resistor"*, http://www.resistorguide.com/photoresistor/, 2018.

A Prototype of Heavyweight Carrying Robot for Garments Industry in Bangladesh

Md. Mizanur Rahman[1], Sanjoy Kumar Sana[2], Md. Rokib Hasan[3], Sadia Kamal[4], Thajid Ibna Rouf Uday[5], Junaed Jahin[6]
{(United International University, Madani Avenue, Dhaka 1212, Bangladesh) [1,2,3,6], (Ai Robotics Asia Ltd., Mohammadpur, Dhaka 1207)[4, 5]
(mm.rahmanbd24@gmail.com[1], sanjoysks000@gmail.com[2], mhasan151321@bscse.uiu.ac.bd[3],
sadia.kamal94@gmail.com[4], ibnauday@gmail.com[5], junaed066@gmail.com[6])}

***Abstract*— In this modern world, we are developing the most innovative scientific solutions, where the robotics and embedded field also experienced an immense growth. Automation and robotics are widely used for solving many major problems that are part of our life. Robotics systems already make life easier. In critical situations such as natural disasters, robots can be utilized instead of humans to avoid more casualties. In this research, we focus on our garments industry to automate the warehouse management using robotics system. The problem is handling heavyweight objects which are to be kept in a storeroom and preparing them to be shipped. As a solution, we introduce the multi-robot system to pick up and drop those heavy objects to the designated point(s). The whole system is designed to locate objects at some point and pick them up based upon their need. The proposed system aims to contribute to the above mentioned applications using a proposed model to control a warehouse using the robot.**

Keywords—line follower bot, warehouse management robot, garments automation, heavyweight carrying.

I. INTRODUCTION

In recent years robotics and the embedded system sector has been revolutionizing the world industry with its solutions. The global competition of production efficiency is growing day by day. Garments Industry is the key export channel and giving huge support to the GDP in Bangladesh. At present, the country generates about $5 billion worth of products each year by exporting garments [1]. The global industry automation is already catching attention. Bangladeshi garments industry also needs to move with this growth rate. Many strategic leaders now think that automation in the warehouse sector has become a basic need in order to sustain in today's economy.



Fig.1.Scenario of a garment factory, source: BGMEA

The carts of products which were being picked up manually in the previous years are now getting more flexibility in terms of mobility because of automation. The performance of collaborative mobile warehouse robots has attracted many leading organizations such as DHL, XPO, and many others to accelerate their working process. Over the years DHL introduced conveyors and shuttle systems to see if it can fulfill the needs of customers and to maintain commitment towards them. But that did not work out eventually. Barcode scanning system was also introduced but scanning consumes time so this method is also not that effective [2]. So now they are also investing in flexible autonomous systems/robots which can scale with the volume of orders [3]. So robots can work as a human assistance [4]. At present, the design optimization of traditional distribution center picking system has caught attention and high automation is being integrated to improve the system [5]. In this paper we are proposing an autonomous solution which is a heavyweight carrying robot. Its impact will be all over the garments industry of Bangladesh as it will bring enormous changes by minimizing the cost and risk factors.

II. Related Works

In recent work, the system uses QR code-based technology to proceed. By scanning pre-coded QR stickers which are placed in a few locations the robot reaches to its destination [6]. In this way, the robot takes less processing time so the whole process is but if the system fails to recognize a QR sticker it will lose its way as there is no other way to receive information. Whereas our proposed solution receives information from various sources so the chance of getting lost is low. Another fully automated warehouse robot uses computationally efficient local path planning algorithm, named the RERAPF [7]. The computational cost of the algorithm is quite high and the whole process is lengthy, so despite providing the assurance of a robot not getting trapped in a fixed region, it is not maybe the best choice. On the other hand, our robot goes to the location based on the priority list that is stored in the stack which makes the whole process simple and less time-consuming.

III. MOTIVATION

Around Bangladesh, hundreds of textile industries need customized solutions and perform research on their activities to increase their productivity and minimize costs. According to the predictions made by PWC prediction, Bangladesh is going to become the 28th largest economy by the year2030. Industrial development will play a significant role in building this potential as it will be one of the most important factors which will help to flourish our economy. In our country, the ready-made garments (RMG) industry has immense strengths and lot of potentials to establish this growth [8]. A lot of

support is needed for the industry according to customization solutions.

IV. Optimization Process

The proposed solution is a prototype of a heavyweight carrying system which will collect objects from some points and deliver to another point inside a closed perimeter ideal warehouse or storeroom, which is likely similar with the fig.2. The robot can work in an orderly manner avoiding obstacles, which happens mainly due to the manipulation of the software system [9]. We consider garment factories that are producing wearable products and keeping them some points for the next stage of processing.

Definition, There are 04 locations with labeled, each location is a point of collecting the product upon on user calling system. The call is defining by the different switch for each location which has two states. When the product is ready for pickup, Call=1, else Call=0.

The environment and operations flow belongs to a garments production house, the product produce and holds in four (04) different locations labeled by A, B, C, and D. The produced product have to pick and delivered to a storeroom which location is labeled by E.

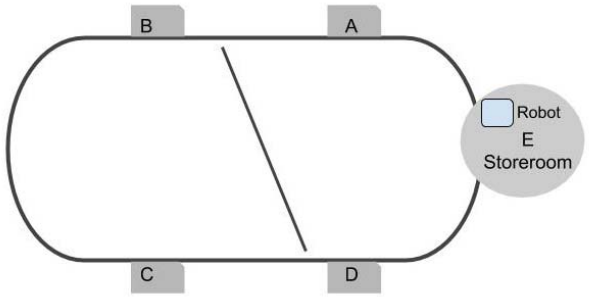


Fig.2. Operations Flow.

These four (04) locations are related with a specific time, t and there will be a priority rank system. From table 1, Location A has the top priority 1, as it contains a product that must be collected and delivered immediately. The robot must move by A and afterload, route the shortest part to location E. It's mean that whatever the calling state of four (04) locations, the robots will pass by the location A.

Table1: Priority ranking

Priority rank	locations
1	A
2	B
3	C
4	D
0	E

Table2: State table

Locations	State	Call (t=1s)
A	0	0
B	1	1
C	2	0
D	3	1
E	4	0

a. Define the state:

In order to understand the programming flow, the system is mapped using some index number according to the locations. After encode the locations name into index numbers with respect to the following the state table 2.

Where, the first line of each matrix, which has index 0, corresponds to location A. The second line/column, which has index 1, corresponds to location B. The robot are connected with control station (calling station) using Bluetooth communication and start the operations depending on their availability.

Form the table2, Say at t=1s, there are two incoming calls for pick up the product when robots position is at E. After call Robot is serving for location B, which means state 1. Robot will operate from A to B and come from B to C, D. After unload from E. Robot will again serve the D, which means state 3 and the call stored using array programming. The Robot will operate from A to D using the shortest part and come from D to E.

The Robot are collecting those products autonomously and dropping them at the designated points, as shown by the ideal workflow in figure 2. The right-hand rule has used for line flowing methodology [10]. This process will be continued until the completion of all tasks.

V. Electronics System

The Electronics system is based on the ATmega328P microprocessor. In this research project, we used Ultrasonic sensors, Load cell, IR sensor array, DC motor and controller, HC05 - Bluetooth module and Li-Po as the battery.

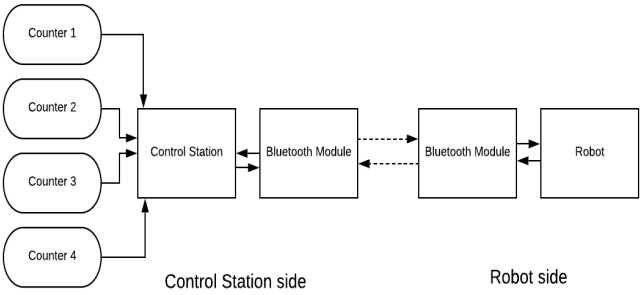


Fig.3. Over all system block diagram.

A. Embedded System (Robot)

We have used popular and very time effective controller, the Arduino Uno board in which the processor is the ATmega328P. The board consists of 14 digital input/output

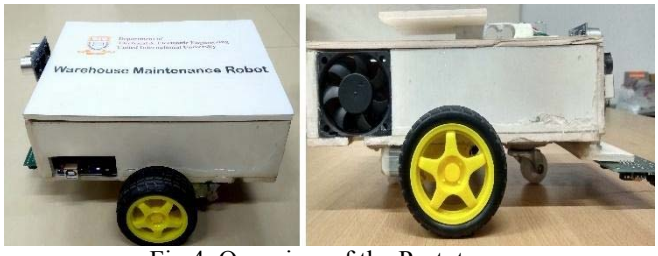


Fig.4. Overview of the Prototype.

Pins along with 6 PWM pins. We have used it to get a faster simulation for this research output. As a motor driver, we have used the L298N which is a monolithic circuit in a 15-lead multi-watt and PowerSO20 packages. It is a dual full-bridge driver designed in a way by which it will accept standard TTL logic levels, it operates at high voltage and current and it also works in low saturation voltage. The device can be enabled or disabled independently through two given enable input signals. An extra supply input is available through which the logic works at a lower voltage. The

maximum operating voltage is 46 V, and the total working DC current, is up to 4A.

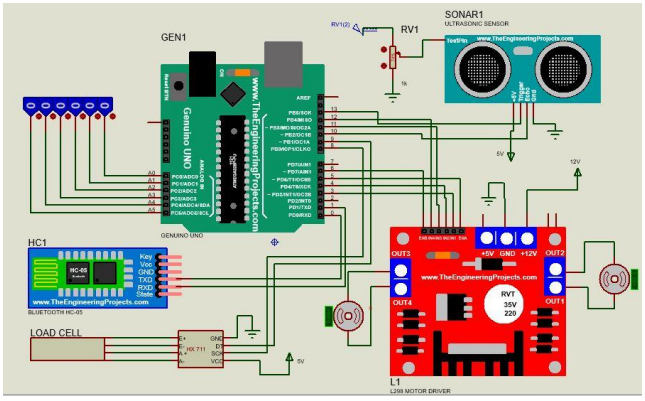


Fig.5. Simulation Circuit Diagram (Robot)

We used few sensors to make it perform autonomous movement. An ultrasonic sonar sensor was needed to determine the distance to an object like some creatures do such as bats or dolphins. To serve that purpose we have used the HCSR04 ultrasonic sensor. Excellent non-contract range detection with high accuracy and stable reading is offered by it. This device is user-friendly. Its detection range is 2cm to 400 cm or 1 inch to 13 feet. Sunlight or any black material like sharp rangefinders do not affect this senor's operation. Ultrasonic transmitter and receiver modules complete it.

For line follower robot and obstacle avoiding robot multipurpose infrared sensor is a must which gives the robot the ability to determine if there is any obstacle in front of it. The IR sensor we have used here detects the radiation which is reflected from its own infrared LED. It is able to detect light or dark lines or even obstacles in front of it by measuring the amount of reflected infrared radiation. There is a RED LED integrated on to the board in order to detect the presence of an object or line. Sensing range can be customized by integrating a built-in variable resistor.

As a force sensing module load cell has been used here, it is designed carefully with a structure of metal and tiny elements called strain gauges placed in some precise locations on the structure. Specialized amplification is required for this cell and the electrical signal output is very small. The 1046 PhidgetBridge will perform all the necessary amplification and measurement needed for the electrical output. This type of cell can measure force in only solo direction. In another direction, force can be measured also but then the sensitivity would be different.

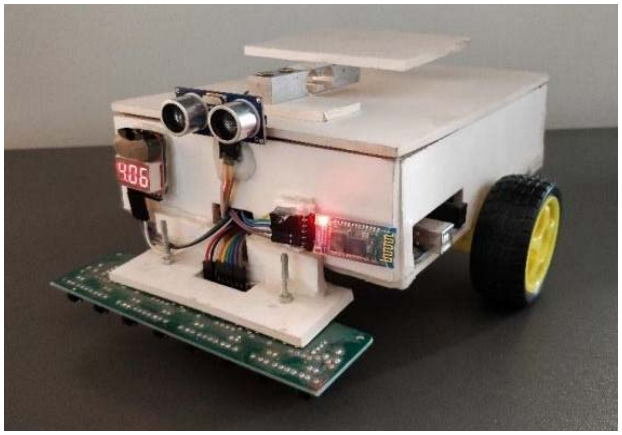


Fig.6. Prototype with Sensors.

B. Embedded System (Control Station)

The robot can take a maximum of four (04) "task calls" from the counter, and the bot will serve one call at a time. At the control side, we used the ATmega328P microprocessor based Arduino Uno board with the Bluetooth module, HC-05. As our ideal model has four counters we add four (04) "task call" buttons to call the robot.

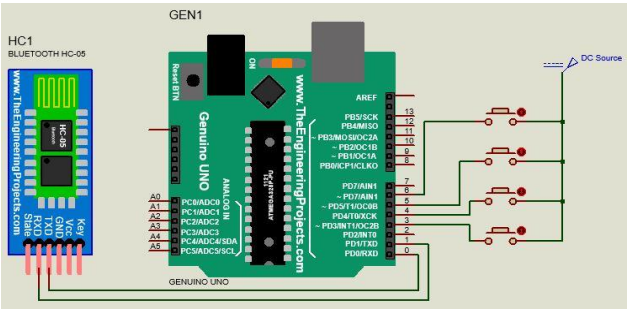


Fig.7. Simulation Circuit Diagram (Control Station)

C. Communication System

The communication system is built using Bluetooth module HC05, which is designed for wireless data transmission between short distance operations and it can also be used as personal area network (WPAN) as it works at ultra-high frequencies (UHF). Two Bluetooth modules have been used here for building the communication system. When pairing them, one of the Bluetooth modules is configured for the master role and another module is configured for the slave role. The modules are arranged in order response work mode so that commands are passed to them. The communication mode is enabled in both modules so that data can be shared amongst them.

D. Load Calculation:

According to the datasheet, for measuring the force a formula has been used to convert the load in to mv/V.

Force = X * Measured mV/V + C (offset)

The road cell output rated is 1.0±0.15mv/v and total capacity of the sensors is 5kg.

So, X is,

Capacity = X * Rated Output

X = Capacity / Rated Output

X = 5 / 1.0

X = 5

As the offset (C) is a variable, so for each sensor it has to be calculated.

C = 0 - 5 * Output

E. Cooling system:

12V DC Axial cooling fan has been used to cool the system. For motor driver, we using heat sink. The total cooling system consumption is 0.08 Amp.



Fig.8. Cooling system.

F. Safety System:

Robotics System are getting populated day by day and the safety system of robotics system has been deployed more. Robots are working with human and for that safety system is need to be properly design [13]. With the advancement of technology, the industry is seeing more robots that can perform tasks autonomously. But with all this development the industry is seeing a new dimension of risk. The robot

contains an emergency terminating button so that in a case of mechanical failure the robot can be turned off manually.
During the process of data collection, communication failure might occur. In such cases, Robot is working with the last command and after that buzzer will buzz for malfunctions. In order to prevent with obstacles sonar sensor data have been used.

VI. Embedded Software System

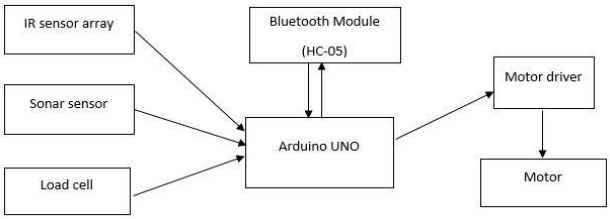


Fig.9. System Overview of the Embedded System (Robot)

The embedded system has been developed using unique method using Advanced C programming. The software system is divided into two subdivision, one is the robot operations system and another is control station system.

a. Robot Operations System:

The programming language that is used to develop communication is C language. Once the communication is established through the Bluetooth module, the robot will check the next command. It can receive four (04) "task calls" at a time currently, using array programming. It will execute the command by going to the counters one by one. The visiting order will be stored in the array. As per the instructions given by the user, the robot will start moving towards its destination by following a black line path. If there is any obstacle or object in front of the robot then it will stop and wait until the object is removed.

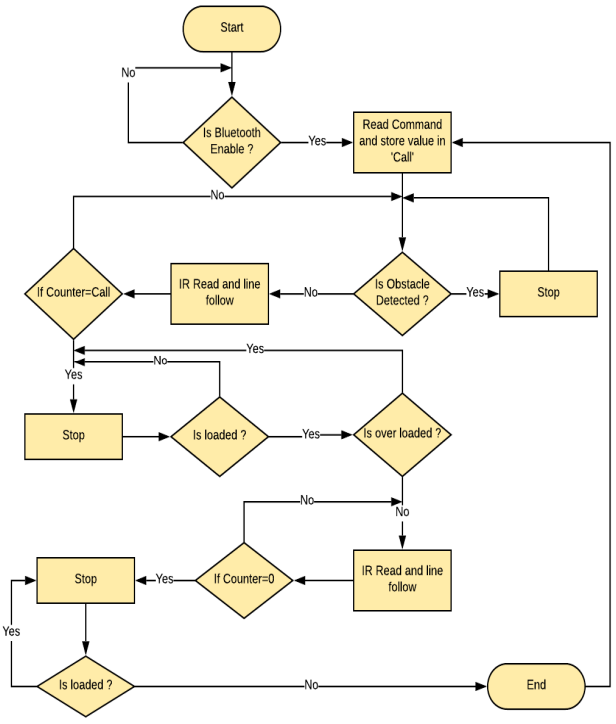


Fig.10. Software System Flowchart

When the robot arrives at the destination, it picks up the box. At this point, the bot will check if the box is overweight. If the weight crosses its limit then a red LED will be turned on to give a signal. If the weight turns out within limit then it will return to the storeroom by following the same black path line and unload the boxes there. If the bot receives another call from the control station while it is completing one task then it will receive the command and store it in the array like mentioned previously. The robot will continue to accept new tasks after it has placed the boxes back to the storage area [9].When the first task is finished it will go to the next destination to fulfill that task.

b. Control Station:

At the control station side we used some modern simple logic for generating faster "task calls" and increase the success of the real-time operation. When a "task call" button is pushed the data will receive the microcontroller and send the data to the bot through the Bluetooth communication system.

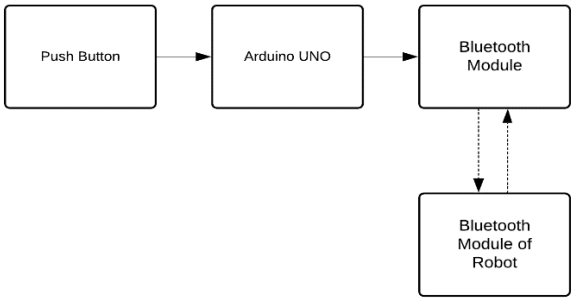


Fig.11. Control System block diagram.

VII. Power System:

The whole power system of this robotic system is divided into two individual parts. One is for the onboard power system and another is for the control station. The onboard power system is a 12V DC system. The total capacity of this system is 07 Watts. We used an integrated monolithic circuit as a motor driver, L298N. This package operates the full-bridge method with a high current dual driver. For this project we used a DC gear motor which is operated by 12V. The consumption of this motor is 500 mA.

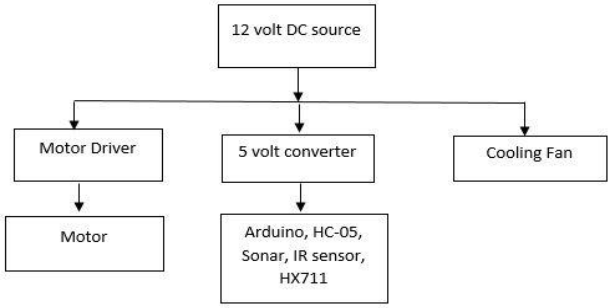


Fig.12. Block Diagram of Power System

All other devices, controller and sensors are operated by 5 Volts. A buck converter LM2596 is used to convert 12 Volts to 5 Volt. The total Watt consumption of this device is 1 Watt. The Control System is operated by a 5 Volt system. The total Watt consumption of this side is 0.50 Watt. We used a Lithium Polymer Battery, widely known as Li-Po Battery and as our system is in the prototype stage, we used a 2.2 Amp 12V Li-Po Battery.

a. Charge Indicator System:



Fig.13. Charge Indicator

There is a charge indicator system for tracking the voltage range. When the below voltage sets value, a buzzer loudly sets off, with a pre-set value of 3.3V. One button setup of alarm voltage

b. Buck Converter: (12V to 5V converge)

We used the voltage regulator LM2596, step down (buck) switching with 3A Load. Using this regulator we buck 12V to 5V for the sensors and motor driver.

VIII. Result

a. Trials & success:

According to proposed solutions the robot is tested in different conditions of calling states. Table 3 is showing the summery of trials results.

From the trail result the system is perfectly working 97.142%, which means robot had missed the task 2.857 times.

Table: 3 Trials & success

Condition Matrix [1=true, 0= False]	High priority state		Trial Results (Delivered to E)	
Initial state, [A,B,C,D]=[0,0,0,0]	Location	State [table2]	Success (out of 10)	Error (out of 10)
[1,0,0,0], call from A	A	0	10	0
[0,1,0,1], call from B, D	B,D	1,3	10	0
[1,1,0,0], call from A, B	A, B	0,1	10	0
[1,1,1,0], call from A, B,C	A,B,C	0,1,2	09	01
[1,1,1,1], call from A, B,C,D	A,B,C,D	0,1,2,3	10	0
[0,0,0,1], call from D	D	3	10	0
[0,1,1,1], call from B,C,D	B,C,D	1,2,3	09	01

b. Robustness:

Most of the robots are not perfectly operates and delivered their task like the human. [10] The system has been designed to help the human to carry the heavyweight product. The robots are enough capable to take the challenge of the current garments need. Currently, the total capacity of the maximum call is 04. But the robustness of the program and processor can take the more call and serve the purpose.

c. Limitation:

There is some limitation in the system. When the call state of C is high. The robot will pass through A and B which is not optimizing the shortest path. But at the call state of D is high, the robot will use the shortest path from A to D.

The robot is taking more than one call, but serve one position till comeback to E position (storeroom). It can be minimized by increasing capacity of the load carrying.

There is no automatic charging system. The user has to charge when the robot buzz for Low battery. This manual process will be overcome with the future adaptation of the charging station.

IX. Conclusion

This research project is just an inspiration of future real-time Artificial Intelligence-Based Systems. For future development, we have planned to research and implement on a robot navigation system, a control updated application improvement and analyze the data of some sensors to control loading/unloading time effectively. In upcoming days we plan to manage the whole inventory through robots. The robots will sort and rearrange the vast inventory 24 hours a day through incorporating deep learning methods. A complete automated system is designed to bring out the utmost efficiency of the garments industry. A fully automated and constant modernization of the processes has become the demand of all the industrial sectors. In this research we have presented our proposed solution to make our industry dynamic. The goal is to make people's life easier and reducing the percentage of error. As there will be less manual tasks chances of error will also be less. The whole world is going through a technological evolution and in order to keep pace with it, bringing improvement in present technologies which are currently used here is our target. Our proposed solution can bring immense change in the industry by reducing manual labor and making the process autonomous.

Acknowledgment

We would like to express our earnest gratitude to the Dept. of EEE, United International University for providing financial aid. We would also like to thank our honorable supervisor Dr. Kaled Masukur Rahman for providing necessary technical supervision. Finally, we are thankful to our Coordinator Dr. Hasan uz zaman for the coordination and supervision. We would also like to thank Ai Robotics Asia Ltd. for giving support through technological planning.

Reference

[1] Masud, A., Hoque, A.A.M., Hossain, M.S. and Hoque, M.R., 2013. Corporate social responsibility practices in garments sector of bangladesh, A study of multinational garments, CSR view in dhaka EPZ. Developing Country Studies, 3(5), pp.27-37.

[2]Mather, T.W and Hsieh, M.A “Ensemble Modeling and control for congestion management in automated warehouses”, Automation Science and Engineering, vol. 10, no. 5, pp. 390-395, Aug 2012.

[3] Jain, D. and Sharma, M.Y., 2017. Adoption of next generation robotics: A case study on Amazon.

[4] Ronsse, R., Vitiello, N., Lenzi, T., Van Den Kieboom, J., Carrozza, M.C. and Ijspeert, A.J., 2011. Human–robot synchrony: flexible assistance using adaptive oscillators. IEEE Transactions on Biomedical Engineering, 58(4), pp.1001-1012.

[5] D’Andrea R, Wuman P. Future challenges of coordinating hundreds of autonomous vehicles in distributionfacilities//Techonologies for Practical Robot Applications,2008. TePRA2008. IEEE International Conference on. IEEE,pp 80-83, 2008.

[6]P. Teja and A. Kumaar, "QR Code based Path Planning for Warehouse Management Robot", 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018.

[7] K. Tsang, Y. Ni, C. Wong and L. Shi, "A Novel Warehouse Multi-Robot Automation System with Semi-Complete and Computationally Efficient Path Planning and Adaptive Genetic Task Allocation Algorithms", 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), 2018.

[8] Intelligent Manufacturing in the Context of Industry 4.0: A Review, Ray Y. Zhong, Xun Xu, Eberhard Klotz, Stephen T. Newman

[9] Industry dynamics. “Amazon use robots handle orders”.Journal of Robot Technique and Application, vol. 1, pp58-60,2015.

[10] Cai, J., Wan, X., Huo, M. and Wu, J., 2010, June. An algorithm of micromouse maze solving. In 2010 10th IEEE International Conference on Computer and Information Technology (pp. 1995-2000). IEEE.

[11] Semwal, V.B., Singha, J., Sharma, P.K., Chauhan, A. and Behera, B., 2017. An optimized feature selection technique based on incremental feature analysis for bio-metric gait data classification. Multimedia tools and applications, 76(22), pp.24457-24475.

[12] J. Enright, P. R. Wurman. “Optimization and Coordinated Autonomy in Mobile Fulfillment Systems. Automated Action Planning for Autonomous Mobile Robots,” 2011.

[13] Semwal, V.B., Mondal, K. and Nandi, G.C., 2017. Robust and accurate feature selection for humanoid push recovery and classification: deep learning approach. Neural Computing and Applications, 28(3), pp.565-574.

Q-Learning Based Particle Swarm Optimization Algorithm for Optimal Path Planning of Swarm of Mobile Robots

Syed Irfan Ali Meerza
Electrical and Electronics Engineering Department
American International University Bangladesh
Dhaka, Bangladesh
irfanmeerza00185@gmail.com

Moinul Islam
Electrical and Electronics Engineering Department
American International University Bangladesh
Dhaka, Bangladesh
rumon.moinul@gmail.com

Dr. Md. Mohiuddin Uzzal
Electrical and Electronics Engineering Department
American International University Bangladesh
Dhaka, Bangladesh
drmohiuddin@aiub.edu

***Abstract*—For a swarm of mobile robots in an unknown environment, the most challenging task is to plan the optimal path and also to learn the environmental parameters. Machine learning is one of the solutions to this problem. This paper proposes a combination of particle swarm optimization (PSO) and Q-value based reinforcement learning (Q-Learning) for a swarm of mobile robots to find the optimal path in an unknown environment and to learn the environment. Q-learning combined with PSO enable the robots to learn the unknown environment with reward and action selection policy while reducing the time to find the optimal path in the environment using the iterative improvement method of PSO. The proposed algorithm is simulated and found performing faster than Q-learning and PSO performing alone. We also compared with some other established algorithm where the proposed algorithm outperforms each of them in accuracy and speed.**

Keywords—particle swarm optimization, reinforcement learning, q-learning, swarm intelligence, mobile robot

I. INTRODUCTION

Our mother nature is an inspiration for a lot of the machine learning algorithms, for example, Genetic Algorithm is inspired by the natural evaluation of the living beings, reinforcement learning is inspired by the behavioral psychology, swarm intelligence is inspired by the natural swarms present in the nature like birds flock, fish school.

There is two types of machine learning, i) supervised learning and ii) unsupervised learning. In supervised learning, the learning parameters are updated based on the training data. The training data is gathered upon the desired output for a known input. However, for the unknown environment, these data are not present. This is where unsupervised learning comes handy. In unsupervised learning, agents learn using the reward and action concept. There is a set of actions from which the agent selects an action and gets a reward for that action. The main aim of the agent is to find the action which gives the greatest reward. For reinforcement learning problems Temporal Difference (TD) is one of the most used concepts. TD is step based concept where the results are found after single time steps. Q-Learning is based on TD, but it is computationally simple and easy to implement [1]. Q-learning works by updating the state-action table where the values are the rewards associated with that action and state.

There a lot of evolutionary algorithm (EA) that are used to solve the reinforcement problems. Some of the popular ones are the Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO). They are population-based algorithms. Many agents are exploring through the solution space and collaborating to find the global optima. As this algorithm uses exploration and collaboration, there is a minimal possibility of them to trap into the local optima. However, in EA there is no limit on the actions that can be taken by the agents, which leads to longer time to reach the optima. On the other hand, reinforcement learning can select the desired actions, but there is no fitness function in these methods which can indicate how soon the agent can control the system's state into a set of goal states [2].

In our proposed method we combined the features of reinforcement learning and the evolutionary algorithms. This enables the action selection by the Q-Table added with the fitness of the PSO, so that the agents find the optimal path and learn the environment faster and with less associated complexity.

Rest of the paper is organized as follow: In section II related works of Q-Learning and PSO are summarized. Section III presents some background on PSO and Q-Learning. Section IV describes our proposed methodology for optimal path planning and environmental learning for a swarm of robots. In Section V our training and testing methods are described. Section VI explains the simulation results and performance of the implemented algorithm. Finally, In Section VII, we conclude the paper.

II. RELATED WORKS

Use of machine learning algorithms is getting popular in robotics. As the robots are used in different environment researchers, tend to use more unsupervised learning algorithms. On the other hand, due to the simplicity and the reliability of evolutionary algorithms, researchers are using hybrid algorithms with reinforcement learning.

In [3] researcher introduced Q-Learning as a TD based approach. While in [1] researchers proved that Q-learning convergences to optima and there is a relationship with dynamic programming and TD. Yi-Chang and Sheng proposed Q-value based PSO for a neuro-fuzzy system [2]. In [4] James Kennedy and Russell Eberhart introduced the PSO while they were studying the behavior of natural swarms.

Many researchers had worked on many hybrid systems as well. Like in [5] researches used Ant Colony Optimization (ACO) to select Artificial Neural Network (ANN) topology and PSO to adjust ANN weights. In [6] they presented reinforcement hybrid learning algorithm (R-HELA) combining the Compact Genetic Algorithm (CGA). Orawan

and Peerapun used modified Q-learning and PSO for target finding application [7]. In [8] researchers used Q-learning to help particle swarm optimization to avoid being trapped into the local optima. In [9] researchers proposed A Q-learning based swarm optimization algorithm for economic dispatch problem.

III. Background Study

A. Q-Learning

Q-Learning is model-free reinforcement learning algorithm [2]. It uses the action selection policy for the agents that are exploring through the solution space. The policy mainly defines the action that is related to a specific state. These actions are selected based on the reward it gets for that particular action. The algorithm maintains a Q table whose values are updated by the reward matrix using (1) after each action is taken.

$$Q(s, a) = r(s, a) + \gamma max_a (Q(s', a)) \quad (1)$$

Where

s is the current state

r is a reward

a is an action

s' is the next state

γ is the discount factor

Discount factor plays a vital role in learning. γ value is confined between [0,1]. Value closest to 0 (zero) makes the agent prone to short term rewards, while a value close to 1 (One) makes the agent more attracted to the long term rewards. In Fig. 1 the structure of Q-Learning is shown.

B. Particle Swarm Optimization (PSO)

PSO is one of the most commonly used evolutionary algorithms. It is a stochastic population-based algorithm inspired by the natural swarm behaviors. Each agent in the algorithm is considered as a particle in the multidimensional space. Individual particle tries to find the global optima by iteratively adjusting their trajectories. Each particle updates their position and velocity based on the best performance by itself and the best performance found in the swarm. The velocity and the position are updated by using (2) and (3).

$$v_i = w*v_i + c_1 * r_1 * (P_i - x_i) + c_2 * r_2 * (P_g - x_i) \quad (2)$$

$$x_i = x_i + v_i \quad (3)$$

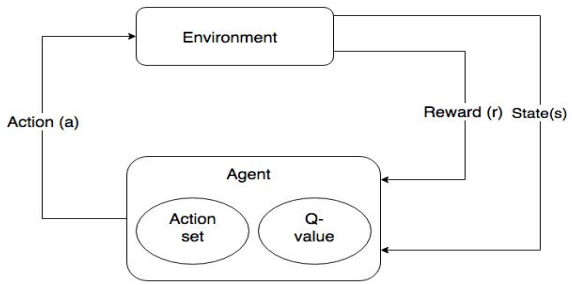


Fig. 1. Structure of Q-Learning.

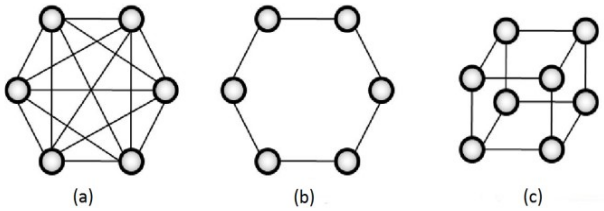


Fig. 2. Three topologies of PSO a) Fully connected,b) Ring and c) Square.

Here v_i and x_i is the velocity and position of a single particle. While P_i is the best position found by that particle and P_i is the best position found the swarm. Here r_1 and r_2 are a uniform random number, and c_1 and c_2 are the constants called the learning factor. Using c_1 and c_2 the algorithm's biases on the local or global best position can be adjusted. To control the effect of the previous velocity in the new velocity a constant w is used.

PSO work with different topologies depending on the population type [10]. The topologies are shown in Fig.2. Fully connected PSO trajectory of each particle is affected the global best performance while for square and ring each particle is affected by their neighbors.

IV. Proposed Algorithm

Our proposed algorithm combines Q-Learning and PSO, where the action is selected by the Q-value and trajectory is defined by the swarm optimization. The primary goal of this algorithm is to guide a swarm of mobile robots in an unknown environment toward the target. The architecture of the algorithm is shown in Fig.3. Here a swarm of particles is initialized with random trajectory, and Q-table is initialized with zeros. Based on the state an action is selected from the action set. The action set contains only four actions which are forward, backward, left and right. Depending on the action and state a reward is given to the particle. These rewards are the Q-value for the table. Using the q-value table of all particles best position in the swarm is determined; also particles learn their own best position by using the q-value. Then PSO operation is done on each particle, and a new trajectory is calculated. It creates a new swarm with updated trajectory and updated Q-table. In Fig.4 flowchart of the total algorithm is shown. We added a limited number of trial for

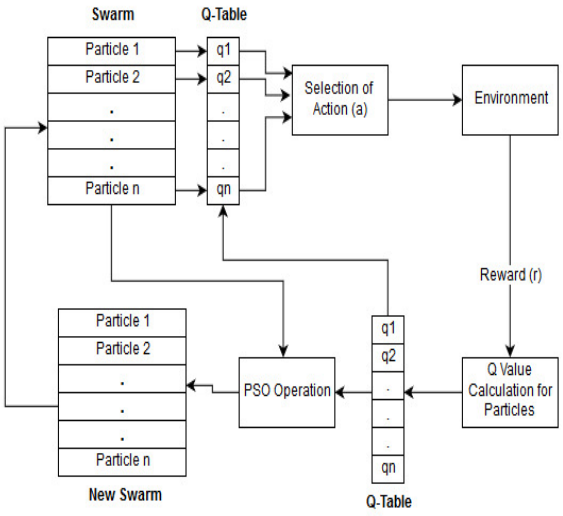


Fig. 3. The architecture of Q-Learning based PSO.

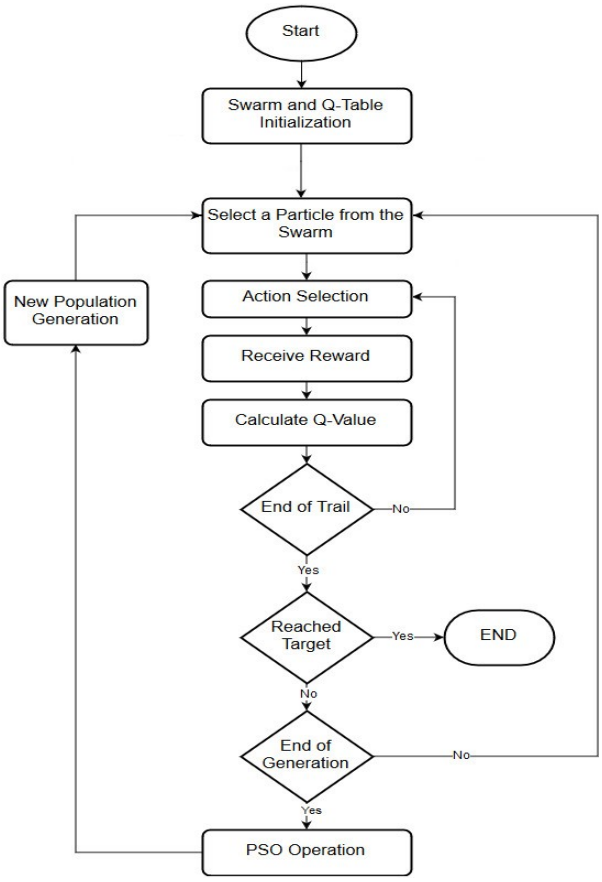


Fig. 4. Flowchart of the proposed algorithm.

each generation so that the algorithm does not stick in local optima.

V. Training And Testing

We simulated the algorithm using Processing which is an open source graphical library and integrated development environment. We created three environments that consist of different type of geometrical obstacles. These environments are characterized by Easy, Normal and Hard. This characterization is done by random walk algorithm where an agent walks randomly into the environment and counting the steps is required to reach the target. This process is done 100 times per environment, and the average number of steps are taken in the count. Environments are shown in Fig.5,6,7 and in Table I the average number of steps required for each environment is shown.

A. Learning

First, we initiate the learning phase to learn each environment. In the learning phase, 100 of agents are initiated at the start point and wanders around the environment till either the trial ends for all the agents in one generation or 95% of the agents each at the target destination which is a red dot in the Fig.5,6,7. After each learning phase, the Q-table is stored. We used different learning rates and found out that the system worked better at a learning rate of 0.9.

TABLE I. Step Required For Different Environment

Quantity	*Easy*	*Normal*	*Hard*
Average number of steps required for each environment	336	594	685

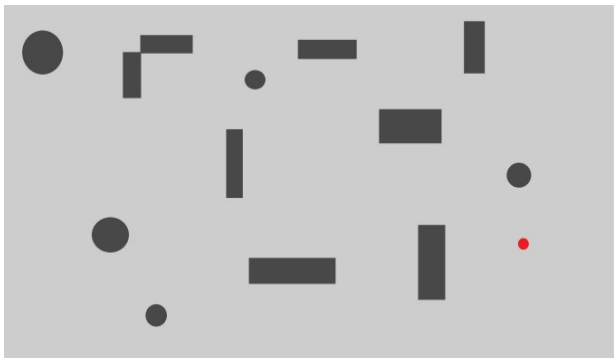


Fig. 5. Easy Environment.

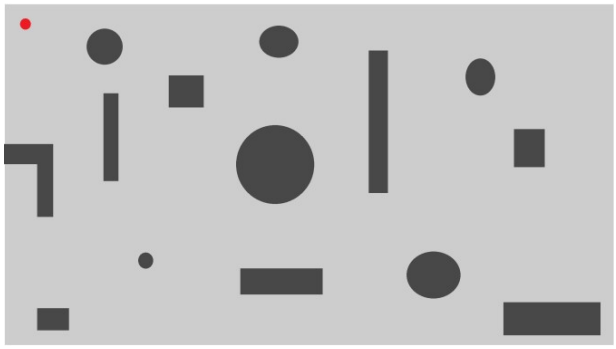


Fig. 6. Normal Environment.

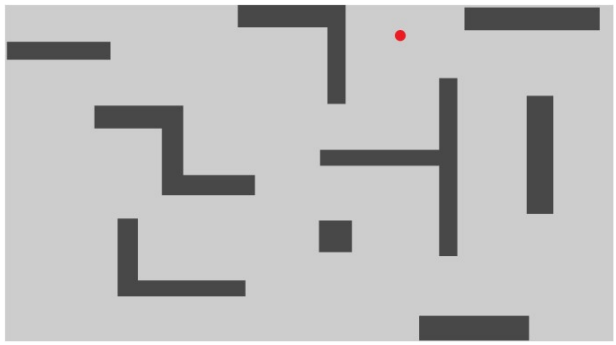
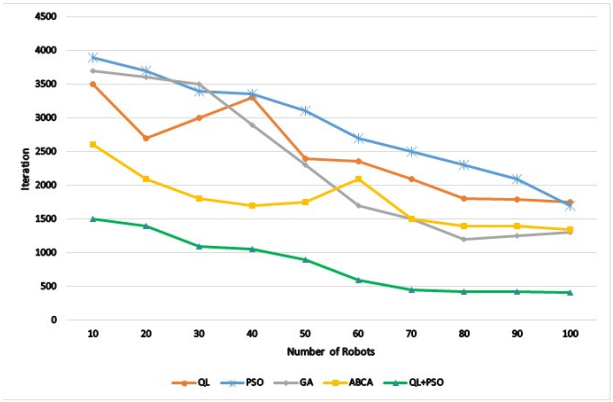


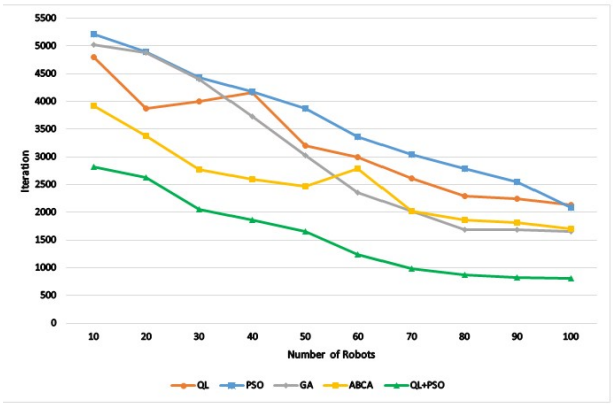
Fig. 7. Hard Environment.

B. Testing

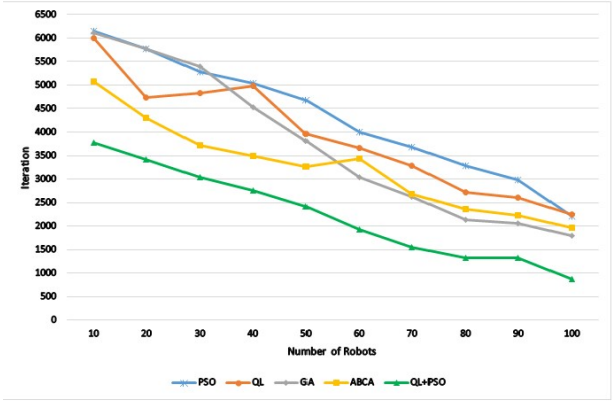
After the learning phase, the system is tested with a different number of agents in the three different environments and measured the performance by evaluating the iteration required to reach the target and also the time to reach the target. We also compared different algorithms like Genetic Algorithm (GA), Q-Learning and PSO and Artificial Bee Colony Algorithm (ABCA). Random walk algorithm is also included as the baseline for the comparison and to judge the difficulty level of the environment.



a) Easy Environment



b) Medium Environment



c) Hard Environment

Fig. 8. Iteration required in training phase

VI. Result And Analysis

We tested the performance of the algorithm by measuring the iteration it required to have 95% of the robots to reach the target in the training phase. Other criteria of performance measurement are the percentage of trials finished with 95% of the robots reaching the target. After the training phase, we measured the performance in the testing phase by measuring the time required to reach 90% of the robots reaching the target. We compared the results with other algorithms. We found that the performance of the algorithms is higher with the higher number of agents in the solution space. In Fig. 8 the required iteration for each algorithm is shown for three

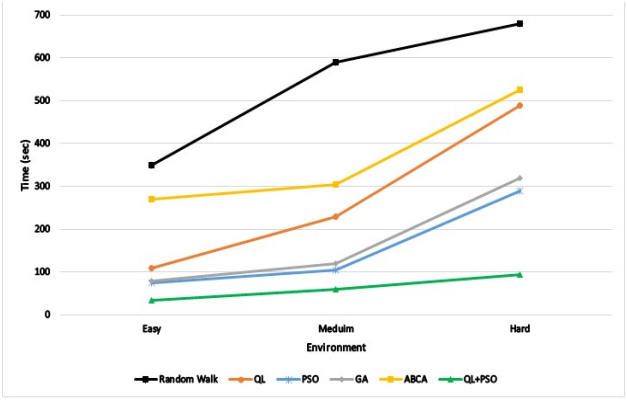


Fig. 9. Test Results

environments. It is visible that the proposed QL based PSO algorithm outperforms all other algorithms. Though the number of iteration required in the hard environment is much higher in other environments. It is due to the nature of the obstacles as the hard environment contains all rectangular obstacles which make it difficult for the robots to pass the obstacles. In Table II the percentage of the trials having 95% of the robots reach the target is shown. Here also the proposed algorithm having QL and PSO working together shown the higher performance than the other algorithms. In Fig.9 the test results are shown. We measured the time required for 90% of the robots to reach the

target in the testing phase. We used the same number of robots as the training phase and used the training data for the same batch of robots in the testing phase. We also compared with the time required for the random walk algorithm to reach 90% of the robots to the target.

TABLE II. Comparison Of Different Algorithms

Quantity	*QL*	*PSO*	*GA*	*ABCA*	*QL+PSO*
Percentage of trials on which 95% of the robots reaches to the target	62	89	87	74	96

VII. Conclusion

In this paper, an algorithm which has the convergence ability of Particle Swarm Optimization and the ability to reduce the number required action feature of Q-Learning is proposed. This algorithm is also compared with other popular machine learning algorithm for performance evaluation. It seemed that the number of iteration required in an unknown environment and time required to reach the target for the swarm is much lower than the other algorithms. Using two algorithms simultaneously increases the complexity and required computation power, but for an increase in accuracy of about 7-8%, it can be considerable.

References

[1] C. J.Wakins and P. Dayan, “Technical note: Q-learning,” Machine Learning, vol. 8, pp. 279–292, 1992.

[2] Y. C. Cheng, S. F. Lin and C. Y. Hsu, “Q-Value Based Particle Swarm Optimization for Reinforcement Neuro-Fuzzy System

Design," in International journal on Computer Science and Engineering (IJCSE), vol. 3, No. 10, pp. 3477–3489, 2011.

[3] C. J. Wakins, "Learning from delayed rewards," Ph.D. dissertation, King's College, Cambridge, U.K., 1989.

[4] Kennedy, J. and Eberhart, R., "Particle swarm optimization", in IEEE International Conference on Neural Networks, vol.4, pp.1942-1948, 1995.

[5] C. J. Lin and Y. C. Hsu, "Reinforcement Hybrid Evolutionary Learning for Recurrent Wavelet-Based Neuro-Fuzzy Systems," IEEE Trans. Fuzzy Systs., vol. 15, no. 4, 2007, pp. 729-745.

[6] G. R. Harik, F. G. Lobo, and D. E. Goldberg, "The compact genetic algorithm," IEEE Trans. on Evolutionary Computation, vol. 3, pp. 287–297, Nov. 1999.

[7] O. Watchanupaporn and P. Pudtuan, "Multi-rovot Target Reaching Using Modified Q-Learning and PSO," International Conference on Control, Automation and Robotics, 2016, pp. 66-69.

[8] M. Khajenejad, F. Afshinmanesh, A. Marandi and B. N. Araabi, "Intelligent Particle Swarm Optimization Using Q-Learning", in IEEE swarm intelligence symposium, 2006, pp 7–12.

[9] Y. Z. Hsieh and M. C. Su, "A Q-learning-Based Swarm Optimization Algorithm for Economic Dispatch Problem", Neural Computing and Applications vol. 7, 2016, pp. 2333-2350.

[10] S. I. A. Meerza, M. Islam, M. M. Uzzal, "Optimal Path Planning Algorithm for Swarm of Robots Using Particle Swarm Optimization Technique", in Proceedings 3rd International Conference on Information Technology, Information System and Electrical Engineering, 2018, pp. 331-335.

Offloading SLAM for Indoor Mobile Robots with Edge-Fog-Cloud Computing

V. K. Sarker[1], J. Peña Queralta[1], T. N. Gia[1], H. Tenhunen[2] and T. Westerlund[1]

[1]Department of Future Technologies, University of Turku, Turku, Finland
[2] Department of Electronics, KTH Royal Institute of Technology, Stockholm, Sweden
Email: [1]{vikasar, jopequ, tunggi, tovewe}@utu.fi, [2]hannu@kth.se

***Abstract*—Indoor mobile robots are widely used in industrial environments such as large logistic warehouses. They are often in charge of collecting or sorting products. For such robots, computation-intensive operations account for a significant percentage of the total energy consumption and consequently affect battery life. Besides, in order to keep both the power consumption and hardware complexity low, simple micro-controllers or single-board computers are used as onboard local control units. This limits the computational capabilities of robots and consequently their performance. Offloading heavy computation to Cloud servers has been a widely used approach to solve this problem for cases where large amounts of sensor data such as real-time video feeds need to be analyzed. More recently, Fog and Edge computing are being leveraged for offloading tasks such as image processing and complex navigation algorithms involving non-linear mathematical operations. In this paper, we present a system architecture for offloading computationally expensive localization and mapping tasks to smart Edge gateways which use Fog services. We show how Edge computing brings computational capabilities of the Cloud to the robot environment without compromising operational reliability due to connection issues. Furthermore, we analyze the power consumption of a prototype robot vehicle in different modes and show how battery life can be significantly improved by moving the processing of data to the Edge layer.**

***Index Terms*—Edge, Fog, Cloud, SLAM, Efficiency, Computation, Offloading, Energy, Performance, Mobile, Robots**

I. Introduction

Small indoor robots which work alongside humans have become ubiquitous in different fields such as home-care, restaurant business and manufacturing plants. For instance, lightweight robots are used in exhibitions for guiding visitors to a desired stall, and are the operational backbone in logistic warehouses of the world's top e-commerce companies. In smart homes, small robots can be used as a pet which can communicate, watch over or play with children and elderly people and can notify or trigger alarm when abnormal situation occurs [1]. Companion robots can remind elderly people to take medicine on time. However, small robots have limited resources such as low battery capacity and light processing capability. In most cases, these robots cannot operate for a long period of time or cannot make useful decisions in complex situations due to these limitations.

Internet of Things (IoT) can be defined as a virtual platform which allows both physical and virtual objects to be interconnected and communicate with each other. It consists of advanced technologies such as sensor fusion, wireless sensor networks and Cloud computing to help improve the quality of services [2]–[6]. Besides, global storage allows users to access data anywhere at any time [7]. Proper utilization of IoT can help overcome some of the drawbacks of small robots. Particularly, real-time position of a small robot can be tracked and a system administrator can send control commands in cases of complex situations [8]. However, IoT-assisted small robots still have some disadvantages. For example, traditional gateways in IoT platform are mainly responsible for collecting data from robots and forwarding the data to Cloud servers. When the data volume is tremendous e.g., due to a large number of multi-sensor equipped robots operating simultaneously, the network bandwidth may be not adequate or the communication latency become too high [9]. Hence, an appropriate architecture is required which can overcome these limitations.

Edge and Fog computing can be described as extra layers between sensor devices, gateways and Cloud servers for enhancing the quality of services. These bring Cloud computing paradigms to the Edge of the network and help reduce the burden of Cloud providing faster services unsupported by Cloud computing. Edge and Fog computing help to reduce energy consumption of sensor nodes and diminish overall latency [10], [11]. The combination of Edge and Fog computing with IoT can provide a suitable approach for enhancing overall performance.

In this paper, we exploit the concept of IoT and Fog along with Edge gateways to enhance energy efficiency and operational performance of small indoor robots. We present a proof of concept for a hybrid Edge-Fog-Cloud architecture which facilitates use of Fog/Edge and Cloud computing for offloading computationally expensive tasks from robots with limited processing power. We discuss related works on how intensive tasks can be offloaded to Edge, Fog or Cloud services in Section II. Section III presents the proposed system architecture applying Edge and Fog computing. Section IV illustrates the results of different tests and discusses the advantages in terms of power saving and performance boost. Section V concludes the work and depicts on future improvements.

II. Related Work

Several works focus on increasing the performance of different algorithms which cannot be fully exploited in resource-

constrained devices. Dey *et al.* analyzed the advantages of leveraging Fog and Edge computing to offload computationally heavy tasks within SLAM algorithms for mobile robots [12]. They provided extensive simulations for different Robot-Edge-Fog-Cloud computational tasks distribution schemes. The authors concluded that Edge and Fog computing can bring significant enhancement to computationally intensive tasks such as localization and positioning tasks. Edge and Fog computing are shown to provide the same computational performance than Cloud offloading, with the benefit of reduced bandwidth and thus faster response time. However, the authors did not perform real-life tests where the communication layer can play an important role and become a bottleneck due to the high data rates between mobile robots and gateways.

Offloading Simultaneous Localization and Mapping (SLAM) tasks to Edge and Fog layers has only been considered recently while Cloud-based SLAM has been studied for the last decade. Benavidez *et al.* deployed an instance of the Robot Operating System in the Cloud to increase the computational capabilities of constraint robots for Visual SLAM (VSLAM) [13]. They pointed how this helps to overcome traditional bottlenecks in robots with limited computational resources with VSLAM performing feature identification and matching which usually need large databases. They demonstrated how parallel computing can be exploited with a multi-node Cloud deployment to facilitate multi-robot SLAM and with limited impact on the on-board computers that control the robots. A similar, more comprehensive work was carried out by Riazuelo *et al.* who presented C2TAM, a Cloud framework for cooperative tracking and mapping [14]. The authors used the unlimited computational resources available in the Cloud to enable the implementation of algorithms which fail to run on single board computers or even more powerful CPUs when the number of agents increases.

Similarly, Turnbull *et al.* utilized Cloud servers to enable advanced neural command for formation control in large groups of robots [15]. However, the network connection between robots and Cloud servers is unreliable and cannot be guaranteed. Therefore, a trade-off exists between the robustness of the solution in terms of available computational resources, operational reliability and performance. This is further explored by Salmeron-Garcia *et al.* in their work on Cloud offloading of a vision-based navigation system [16]. The authors stress in their paper that processing time can be matched to communication period with parallel computation in the Cloud and inferred of a bottleneck due to communication bandwidth and latency. They concluded that Cloud is necessary if the system needs to be easily scalable or precise operation is required. This occurs where the localization or navigation requirements are such that high quality images need to be processed. Extensive works exist in this direction by other authors in the field [17] [18], taking the trade-offs into account both in terms of communication and computational performance.

Kumar *et al.* gathered computational offloading techniques and algorithms into a survey where they accentuate how shifting of intensive applications will be essential in the future for battery powered devices [19] and explored the energy-saving potential of offloading computation to the Cloud.

Extensive work has been carried out on leveraging Cloud computing for offloading heavy computational tasks. However, most of this work have been focused on leveraging virtually unlimited computational resources of the Cloud. Fewer research have targeted the Edge and Fog computing paradigms as a solution to the high latency and unreliability of Cloud services. Keeping these in mind, we focus in both energy efficiency and performance improvement by placing the most expensive computational tasks in Edge and Fog layers leaving the Cloud only for monitoring and management and thus not entirely relying on network link to the Cloud services for normal operation.

III. System Architecture

The proposed system architecture consists of 4 layers: *Robot Layer*, *Edge Layer* for data processing and analysis, *Fog layer* for distributed storage and *Cloud Layer* for monitoring and general mission control. As shown in Fig. 1, we have combined the Edge and Fog layer because the Fog layer consists of the smart Edge Gateways and they use the same physical resources. The architecture we present is easily scalable, modular and distributed by definition. This section thoroughly describes the role of each layer and the distribution of the computational load through the network. Energy consumption, latency and computational power at different layers are the key factors taken into account when deciding the role of a layer.

In general, data is acquired by robots where minimal analysis is performed. Real-time decisions are taken at the Edge gateways to minimize the latency and share the computational load of different connected robots. The Fog layer provides additional services such as distributed storage and processing to enable an efficient handover mechanism with minimal data loss when a robot switches from one gateway to another. The handover includes sharing previous map data acquired by the robot switching gateways so that the result of the SLAM algorithm remains constant through the connection change.

Edge gateways also play an important role in reducing the energy consumption at the end nodes or robots by running the most computational expensive operations. Both Fog and Edge computing are essential in safety critical situations where they provide a more robust situational awareness and overall system control, and robot state knowledge. Finally, the Cloud layer enables end-users or administrators to control the system by giving general instructions and monitor the performance at different layers.

A. Robot Layer

In the robot layer, sensor data is gathered and streamed in real time to the smart Edge gateways. The control of the robot is run online and movement instructions are given in real-time. In the proposed system, the robot can be either aware or unaware of its current state depending on how extensively information is analyzed on-board and whether an

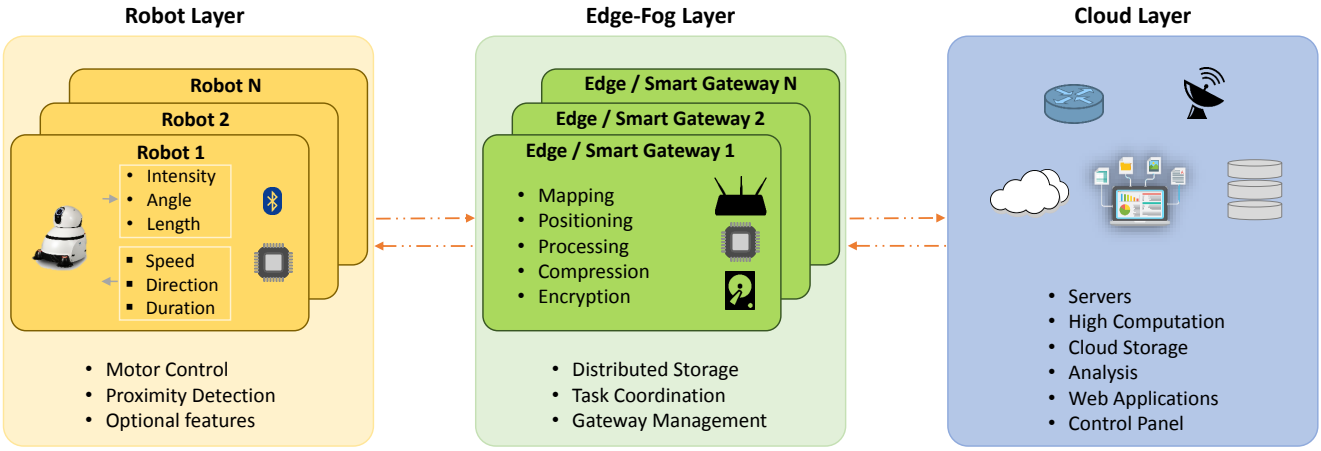


Fig. 1. Proposed System Architecture

Edge gateway provides enhanced state information to the robot or only movement instructions.

The robot relies on low power wireless communication technologies such as Bluetooth or nRF if the bandwidth requirements are met, or Wi-Fi for more bandwidth-intensive applications such as high-quality video transmission. By streaming sensor data directly to the gateway without being processed by the MCU, the robot gains in energy efficiency. Furthermore, low-power techniques can be applied to the MCU in order to further reduce the overall power requirements.

Compared to a traditional approach where sensing, obstacle detection, path planning and localization are made onboard the robot, we can significantly reduce the computing requirements of the robot's control system and downgrade them from processor-based to MCU-based. Only the most basic information is analyzed at the robot itself. This includes, for instance, wheel odometry or inertial data so that the robot is aware of its current acceleration, velocity and orientation. This can be combined with state estimation performed online at the Edge, Fog or Cloud layers to allow more accurate movement and simplify the instructions given to the robot.

B. Smart Edge Gateway Layer

Edge gateways receive sensor data from one or multiple robots in real-time. This includes, for example, images from on-board cameras or data from range sensors such as Lidars and radars, which is analyzed at the gateways to perform obstacle detection and avoidance, path planning, localization or mapping. Odometry data (e.g. wheel odometry, visual odometry) and inertial data (e.g. magnetometer, accelerometer and gyroscope data) is also transmitted from and processed at the gateway.

Acting as a central element, Edge gateways ensure a low-latency and robust solution and provide fast decision making while relocating computationally expensive tasks from the robots. Moreover, if multiple robots are connected to a single gateway, data from different sensors and sources is aggregated and analyzed in order to obtain a more comprehensive understanding of the environment. For instance, two robots operating in the same environment and connected to the same gateway are able to obtain information of larger areas through sensors of other robots and see through other agents if these are nearby and blocking their field of view.

The role of the Edge layer is of superior importance in safety-critical situations and scenarios such as industrial environments where human share operational space with autonomous or semi-autonomous robots. Compared to the more traditional practice of moving complex tasks to the Cloud, the Edge layer reduces the latency to the point where safe operation is achievable even in the case of network connectivity failure. From an operational point of view, smart Edge gateways receive information from sensors aboard multiple robots and are in charge of decision-making in terms of robot movement and task allocation. These instructions are transmitted continuously and wirelessly to the robots.

C. Fog Layer

The Fog layer includes the smart Edge gateways as part of it. Nonetheless, it is defined as a separate layer because its role in the proposed system architecture differs from the data analysis and instantaneous control role of the Edge layer. The Fog layer represents the interconnection of different Edge gateways, together with other services such as distributed databases or location services. In particular, the Fog layer takes care of the handover mechanisms when the robots disconnect from one gateway and connect to another one together with the procurement of additional services including distributed storage, external tracking and monitoring or location services.

Detailed research has been performed to minimize latency and data-loss during the handover [20] [21]. An alternative for managing the handover is to deploy Edge gateways just behind conventional gateways such as Wi-Fi routers so that the gateways can directly store the received information in distributed storage. Thus, the handover problem would be reduced to the distribution of the computational load of individual robots into

the set of Edge nodes, since all nodes are able to access robot data in real-time from the shared storage.

Localization algorithms are mostly run in the Edge layer in which real-time sensor data is matched with an area of an existing map. This is crucial in situations where the robot operates in a partially or totally unknown environment and simultaneous localization and mapping algorithms are run in the smart Edge gateways. In these cases, it is required that the local map stored in one gateway is shared to other gateways before or during the handover. Accordingly, our proposed system architecture includes a distributed storage deployment in the Fog layer. This can be implemented, for instance, in the form of a decentralized database where the maps and other essential robot data are stored and available to all nodes in the Edge/Fog layers at any time.

Other services included in the Fog layer are external sensor management, collaborative processing and monitoring. For example, while Edge gateways process the on-board sensor information from the robots, external monitoring cameras and other sensors can be used to increase the accuracy of localization algorithms or assure collision avoidance in cases of sensor failure or other local problems that might appear in the robot and go undetected by the Edge gateway. This results in an enhanced situational awareness [22] where both externally and internally collected information about a robot is available in the Fog layer. By introducing additional services in the Fog layer, we can enhance overall system robustness and its fault tolerance. For instance, if a gateway abruptly disconnects or any kind of failure occurs, all the information that it was handling will be available to other gateways due to the shared storage resources. This allows the robot to reconnect to the next available gateway and continue its operation with minimal data loss, operational interruption and latency.

D. Cloud Services

Time-series data of the robot state, including basic information such as position, velocity, acceleration, steering drift factor, drive torque variation, instantaneous current consumption and battery level is uploaded to the Cloud for monitoring, control and visualization by users or administrators [23]. Instead of uploading raw sensor data, it is pre-processed and compressed in the Edge/Fog layers and only sporadic local map updates or critical information is sent. This can be implemented in a dynamic way in which, for instance, the map around a robot or of a given area in the operation environment is updated and uploaded to the Cloud at different frequencies depending on whether an end-user is actively monitoring the operation or not.

The use of Cloud services facilitates generic control instructions to the robot and stable operation [24]. For example, factory managers can easily have an overview of the position of different robots around the factory floor, or override autonomous roaming to involve more robots assigned within a certain group or to prioritize specific tasks. Besides, the Cloud allows global access for the administrators and mission planners to have a collaborative overview in a cost-effective way.



Fig. 2. Prototype car with Lidar running on top, used for the experiment

However, due to comparatively higher end-to-end latency between the robot layer and the Cloud, it should not be used for primary control purpose as it can be subject to variation in network performance. The Cloud provides access to pertinently processed data from the robot layer which is periodically backed up at a reasonable interval. This yields a conclusive operation and data history which can be analyzed using high performance computing in the Cloud for machine learning, better prediction and big data analytics.

IV. Implementation and Experimental Results

In order to test the feasibility and validate the proposed architecture, we have constructed a car controlled via Bluetooth communication to produce a map of the perimeter it is bounded by. The car shown in Figure 2 incorporates a 360° Lidar on top of the chassis, two independent motors- one for steering control and another for driving forward or backward with appropriate motor control circuitry. The car is connected to a gateway through a Bluetooth Classic module. An AVR 8-bit MCU is used to control acquisition of raw data from the Lidar and send to gateway via the serial port profile of the Bluetooth standard. A prototype is built using the chassis from an *Reely 1:10 Elektro-Monstertruck NEW1* RC car and replacing the control circuit and motors with lower power variants with an RPLIDAR A1 M8 on the top, a low cost 360° 2D LASER scanner from *Slamtec* [25].

We have run experimental tests with a minimal setup where only Lidar is used for both localization and mapping purposes. Adding inertial sensors would provide more accurate positioning and control over the vehicle; however, that is not the main objective of this work and we focus on the implementation of the hybrid Edge-Fog-Cloud architecture and its benefits.

The MCU controlling the movement of the car does not save or analyze the data from the Lidar sensor in any form and is forwarded directly to the gateway as serial stream over Bluetooth. The gateway then runs an adapted version of the BreezySLAM algorithm by S. D. Levy [26]. The algorithm has been modified to take into account the intermittent nature of

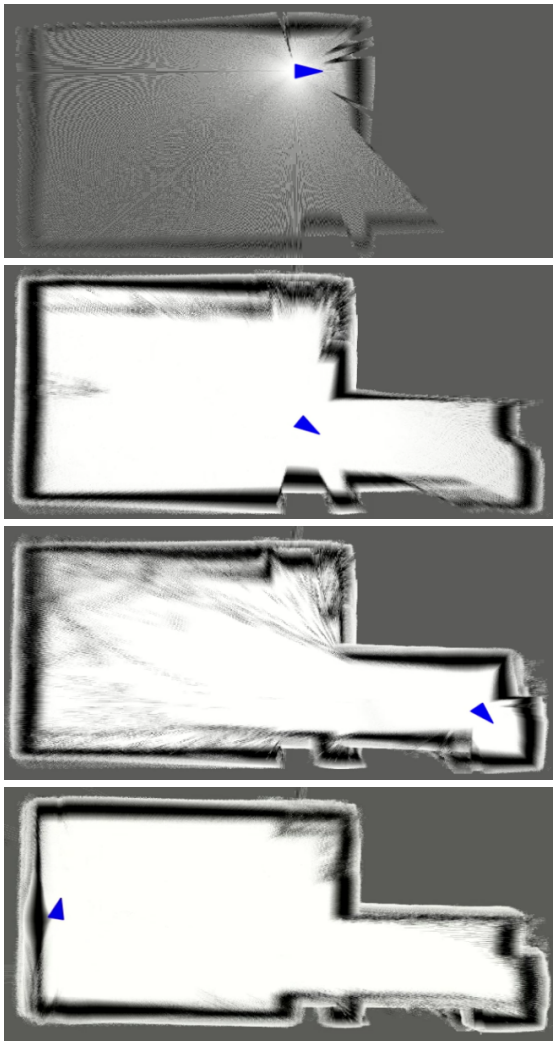


Fig. 3. Mapping and Localization at the Edge. (The position of the car is marked with a blue pointer)

the connection between the robot and the gateway, and execute accordingly when processing successive batches of data.

Figure 3 shows the result of applying the mapping and localization algorithm in the Edge gateway. As we are using an inexpensive Lidar and the SLAM algorithms run without any inertial measurement unit such as accelerometer and magnetometer, accurate positioning is hard to achieve in complex environments. However, with more input data from multiple sensors would suffice for enabling operation in non-trivial local environments.

The first map in Figure 3 shows the map obtained after the first few Lidar frames have been analyzed. The *RPLIDAR A1 M8* used in the tests is a low-cost Lidar and therefore several scans are processed before obtaining a good quality map. The granularity in the first map is derived from the relatively low scan speed of the device compared to more advanced sensors. Also, while the position within the map is accurate and consistent through the test, noise appears near the distant walls from the Lidar due to the low accuracy. However, the walls and corners in the local area around the robot are with little to no noise.

TABLE I
POWER CONSUMPTION OF THE PROTOTYPE VEHICLE AT 5V SUPPLY

Operation / Approach	**Average Current (mA)**	**Average Power (W)**
MCU Idle	4	19
MCU + Lidar	339	1695
MCU + Lidar + Streaming	378	1890
R.Pi Idle	301	1503
R.Pi + Lidar	622	3108
R.Pi + Lidar + Edge Processing	829	4152
Car (Drive / Steer)	122	607
Car (Drive + Steer)	184	918

In terms of performance, the recorded path elapsed a total of 44.3 seconds in which 314 Lidar frames were processed. This translates into 7 frames/s, which matches with the Lidar specification for a maximum speed scan of 10 Hz with an average car speed of 0.3 m/s. This frame rate can be compared to 1 frame/s processed on a Raspberry Pi 3 Model B running Ubuntu Mate Desktop with the graphical map display in real-time. While this can be optimized running a more trimmed, customized version of Linux on the Raspberry Pi, its computational capability limits the processing to about 5 frames/s. Since a MCU is able to handle the vehicle control and transmission of Lidar data, not only it significantly increases the system performance by offloading the SLAM tasks to the Edge gateway, but a considerable amount of energy is saved by replacing a power-hungry single-board computer such as the Raspberry Pi with a low-power MCU.

In particular, the specific power consumption and average current drawn by the prototype in different states are documented in Table I. In idle state, the Lidar consumes 300 to 350 mA, representing over 90% of the total power consumption. This can be replaced with low-power cameras or more energy efficient sensors for a particular SLAM implementation. However, when the prototype is moving, the Lidar's share of the power consumption goes down to around 60%. When choosing the Lidar, general purpose motors have been used to show that even with out-of-the-shelf devices, the overall energy efficiency can be improved by around 40% when a single-board computer such as a Raspberry Pi in this test is replaced with a MCU. This yields a significantly longer battery life and an increase in operational performance due to larger ratio of active-time with respect to idle or charging time.

V. DISCUSSION

Although Edge computing can bring advantages such as reduction in latency, bandwidth conservation, improvement in application robustness and security [11], [27], [28], there are inherent challenges. As the processing of the data stays near the Edge, specific data-oriented applications which require comparatively higher user interaction will have insignificant

performance gain. In the worst case scenario, performance can deteriorate due to the long path data has to travel.

As multiple robots can connect to the same smart gateway, communication latency and inter-unit interference can be an issue. Depending on the application, an appropriate communication medium must be chosen to ensure that specific parameters such as bandwidth, channel spacing, protocol, sampling frequency and data rate are properly considered for an interruption-free data transfer. Also, for wireless units, it is useful to examine the wake-up and sleep time of the radio communication module so that transmitter and receiver circuitry is not toggled unnecessarily and sleep modes are used appropriately. This yields better energy efficiency and communication at an optimal duty cycle.

Keeping the Edge gateways near the robot layer and a Fog layer provides distributed and scalable computing. However, as the number of robots and frequency of data transmission per node significantly increases, the gateways may not handle all requests within the highest permitted latency. In such a case, having more powerful and redundant gateways along with advanced load-sharing algorithm in the Fog and distributed data storage can effectively improve performance.

VI. Conclusion and Future Work

In this paper, the comparative advantage of using Edge computing for robots and nodes are discussed keeping the focus on energy efficiency and operational latency. Instead of directly transferring a large amount of data from robots to the Cloud, we proposed an Edge-Fog-Cloud based system for performance advantages. Run from the mains supply, the powerful Edge layer can pre-process and analyze data easily, implement advanced features and ensure data security by applying complex encryption algorithms. The Fog layer can intelligently manage the smart gateways of Edge layer.

This approach effectively takes the heavy computational tasks near the Edge layer and thus reduces the required energy for operation of the node or robot itself. The shorter path of data for a complete cycle improves latency and frees up valuable network bandwidth, especially when a lot of robots are connected to the same gateway. In addition, the reduced robot-to-Cloud transferable data eliminates potential bottleneck in the network and lowers the chance of system downtime due to a single point computing failure as in centralized computing architecture. We demonstrated an example of SLAM showing how a car in the robot layer can take advantage of the power of Edge computing for mapping a confined perimeter.

In future, we plan to implement a comprehensive system consisting of several vehicles, multiple smart Edge computing gateways constituting a Fog layer and an application for the Cloud layer. In addition, we plan to develop an autonomous roaming algorithm for the robot car so as to perform the mapping operation in shortest possible time. Besides, some stress tests to estimate the performance of smart gateways will be evaluated to emulate the scenario when the number of connected cars is sufficiently high.

References

[1] N. Mitabe and N. Shinomiya. An intelligent care support system for the elderly with an autonomous robot in ambient sensors. In *2017 IEEE GCCE*, pages 1–4, 2017.

[2] M. Jiang *et al.* Iot-based remote facial expression monitoring system with semg signal. In *2016 IEEE SAS*, pages 1–6. IEEE, 2016.

[3] S. R. Moosavi *et al.* Session resumption-based end-to-end security for healthcare internet-of-things. In *IEEE CIT*, pages 581–588. IEEE, 2015.

[4] M. Ali *et al.* Autonomous patient/home health monitoring powered by energy harvesting. In *IEEE GLOBECOM 2017*, pages 1–7. IEEE, 2017.

[5] T. N. Gia *et al*. Iot-based continuous glucose monitoring system: A feasibility study. *Procedia Computer Science*, 109:327–334, 2017.

[6] I. Tcarenko *et al.* Energy-efficient iot-enabled fall detection system with messenger-based notification. In *International Conference on Wireless Mobile Communication and Healthcare*, pages 19–26. Springer, 2016.

[7] R. Petrolo *et al.* Integrating wireless sensor networks within a city cloud. In *IEEE SECON Workshops*, pages 24–27, 2014.

[8] T. Kitanouma *et al.* Cloud-based self-organizing localization with virtual network topology for wireless sensor networks and its implementation. In *2016 IEEE PIMRC*, pages 1–7, 2016.

[9] W. Yu *et al.* A survey on the edge computing for the internet of things. *IEEE access*, 6:6900–6919, 2018.

[10] B. Negash *et al.* Leveraging fog computing for healthcare iot. In *Fog Computing in the Internet of Things*, pages 145–169. Springer, 2018.

[11] T. N. Gia *et al.* Exploiting fog computing in health monitoring. *Fog and Edge Computing: Principles and Paradigms*, pages 291–318, 2019.

[12] S. Dey and A. Mukherjee. Robotic slam: a review from fog computing and mobile edge computing perspective. In *MobiQuitous*, pages 153–158. ACM, 2016.

[13] P. Benavidez *et al.* Cloud-based realtime robotic visual slam. In *2015 Annual IEEE Systems Conference (SysCon) Proceedings*, pages 773–777, 2015.

[14] L. Riazuelo *et al.* C2tam: A cloud framework for cooperative tracking and mapping. *Robotics and Autonomous Systems*, 62(4):401 – 413, 2014.

[15] L. Turnbull and B. Samanta. Cloud robotics: Formation control of a multi robot system utilizing cloud infrastructure. In *2013 Proceedings of IEEE Southeastcon*, pages 1–4, 2013.

[16] J. Salmerón-Garcı´a *et al.* A tradeoff analysis of a cloud-based robot navigation assistant using stereo image processing. *IEEE Transactions on Automation Science and Engineering*, 12(2):444–454, 2015.

[17] Y. Hong *et al.* Energy efficient content-based image retrieval for mobile systems. In *2009 IEEE International Symposium on Circuits and Systems*, pages 1673–1676, 2009.

[18] J. Hauswald *et al.* A hybrid approach to offloading mobile image classification. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8375–8379, 2014.

[19] K. Kumar *et al.* A survey of computation offloading for mobile systems. *Mobile Networks and Applications*, 18(1):129–140, 2013.

[20] X. Wang *et al.* Offloading in internet of vehicles: A fog-enabled real-time traffic management system. *IEEE Transactions on Industrial Informatics*, 14(10):4568–4578, 2018.

[21] T. N. Gia *et al.* Fog computing approach for mobility support in internet-of-things systems. *IEEE Access*, 6:36064–36082, 2018.

[22] C. V. Smith *et al.* Enhanced situational awareness in autonomous mobile robots using context-based mapping. In *2013 IEEE CogSIMA*, pages 134–138, 2013.

[23] B. Zhou *et al.* Remote control system of mobile robot based on cloud platform. In *2017 2nd International Conference on Robotics and Automation Engineering (ICRAE)*, pages 94–98, 2017.

[24] C. Briese *et al.* Cloud-based active disturbance rejection control for industrial robots. In *2018 IEEE ETFA*, pages 559–565, 2018.

[25] Slamtec. RPLIDAR A1. [Online] Available: http://www.slamtec.com/en/lidar/a1. Accessed: 20 Feb. 2019.

[26] S. D. Levy. Breezyslam. [Online] Available: https://github.com/simondlevy/BreezySLAM. Accessed: 20 Feb. 2019.

[27] T. N. Gia *et al.* Fog computing in healthcare internet of things: A case study on ecg feature extraction. In *2015 IEEE CIT*, pages 356–363. IEEE, 2015.

[28] N. Abbas *et al.* Mobile edge computing: A survey. *IEEE Internet of Things Journal*, 5(1):450–465, 2018.

Simulation of Ferrofluid Based Soft Actuator Using Magnetic Field Optimization Approach

Tasnova Tanzil Khan
Mechatronics, Department of Industrial Systems Engineering,
Asian Institute of Technology
Pathumthani, Thailand
tasnovatanzil@gmail.com

A. M. Harsha S. Abeykoon
Mechatronics, Department of Industrial Systems Engineering,
Asian Institute of Technology
Pathumthani, Thailand
harsha@ait.ac.th

Tanujjal Bora
Nanotechnology, Department of Industrial and Systems Engineering,
Asian Institute of Technology
Pathumthani, Thailand
tbora@ait.ac.th

Golam Fahad Mullick
Department of Electrical and Electronics Engineering,
American International University Bangladesh
Dhaka, Bangladesh
fahadmullick89@gmail.com

***Abstract*— In recent times, soft robot has become an unconventional field of research in robotics. In comparison to classic robotics, soft robot shows an excellent adaptability in nature, surveillance with more pliability and of course a simpler structure. In this paper, a new approach for ferrofluid based soft actuator which shows various vertical motions when varying magnetic field is applied, has been introduced for the first time. Finite element method is used here to model different motions of the ferrofluid tube together with field, force and stress analysis. Eventually this simulation approach gave us values of the force exerted on the tube, the magnetic stream lines, direction of the flux, force direction and also the displacement of the tube when introduced to varying magnetic field.**

Keywords— ***ferrofluid, soft actuator, magnetic field, vertical deflection, finite element modelling***

I. INTRODUCTION

Harsh and extreme environments, on different locations in the universe, are a combination of fluctuating pressure, temperature and harmful radiation. Therefore more complex yet intelligent robotic devices are in need to be established which withstand these unfavorable conditions [1]. Introducing soft and flexible robots can be advantageous in many applications as they can pass though small tunnels, bear extreme conditions and interact with human. Meeting these conditions with soft robots can make a crisis free future with easy problem solving ability. As nature has always been an inspiration in the field of research with soft robots, starting from designing, fabricating and controlling the structure have grasped recent attentions of researchers from different field of studies, such as chemistry [2], physics [3], biology [4], and robotics [5][6]. The flexibility and smooth movement of robots is possible now due to the invention of soft materials. This simple yet flexible movement of soft robots has reduced the need for hard robots and of course lowers complexity of the system itself with less budget [7].

In comparison to classic form of robotics, fluid based soft robots are more advance in form of adaptability, flexibility and sovereignty [8]. A pneumatic fiber-reinforced actuator where the air pressure and external fiber arrangement replicated complex motions of the index and thumb of a human being [9], a soft anthropomorphic finger made with a low cost soft hyperplastic material were some of the recent breakthroughs in the world of robotics [10].

Beside those pneumatic actuators, some researchers used smart material like ferrofluid to make soft robots. Ferrofluid is a colloidal fluids which is composed of nano-sized ferromagnetic particles, made of magnetite ($Fe3O_4$) or hematite (Fe_2O_3) that are called ferromagnetic particles. It is a kind of smart fluid that flows to the direction of magnetic flux density therefore can be easily manipulated by applying magnetic field. A pumping device made with ferrofluid which changes its directions according to the change in magnetic fields has been developed in the year 2012, using permanent magnet [11], altering its own shape and position according to varying magnetic strength [12]. Also in 2013, performance of a solenoid type electromagnetic actuator with nano-ferrofluid was investigated and analyzed [13].The air gap of the actuator was filled with ferrofluid and was considered as a magnetic gap. A more improved version of it was designed later in the year 2014, for low energy consumption which caused the electromagnetic force increased up to 62% [14].

In addition, imitation of biology systems with soft robot has increased the softness of the robots and added more degree of freedom than rigid robots. In 2015, a soft robot with ferrofluid that resembles *Amoeba* like movements controlled by electromagnets was introduced. [15]. Ferrofluid actuation using both permanent and electromagnet is being used in energy harvesting system where an electromagnetic force is induced due to the movement of ferrofluid inside the device [16] and also in efficient levitation method [17],[18], [19]. Moreover, ferrofluid has also been used in smart drug delivery systems [20].

This paper presents the simulation technique used to evaluate the magnetic force and the displacement of the ferrofluid tube when varying magnetic field is applied on it by different coil placements. A thin rubber tube is filled with adequate amount of ferrofluid and introduced to tiny electromagnets with varying input voltage. The input voltage and the coil position is changed in such a way to create

various vertical movement pattern of the tube showing the degree of flexibility of the tube. Finite element modelling method is used for the magnetic field and solid mechanics simulation purpose. Simulation showed the values of the force exerted by the coils on the tube, the magnetic stream lines, direction of the flux, force direction and also the displacement of the tube when it was introduced by magnetic field produced in the coils. Recently in the year 2018, similar approach have been reported with ability to calculate reasonable velocity fields, forces, and temperature profiles of a ferrofluid-based passive cooling system [22].

II. METHEDOLOGY

The simulation procedure involves three stages: processing stage, solving stage and lastly the post-processing or result stage. In the processing stage, the fundamentals of the model are defined, that includes structural modelling, selection of materials and the meshing criteria. The solving stage includes two solver modules, magnetic field and solid mechanics which solves the model with related mathematical equations and appropriate study type (in this case, stationary study has been chosen). Finally after solving the model, the simulation shows the direction of the magnetic field lines, direction and values of the electromagnetic forces and the structural bending of the tube along with the displacement values. The flowchart showing the procedure involved in each of the stages is shown in Fig. 1.

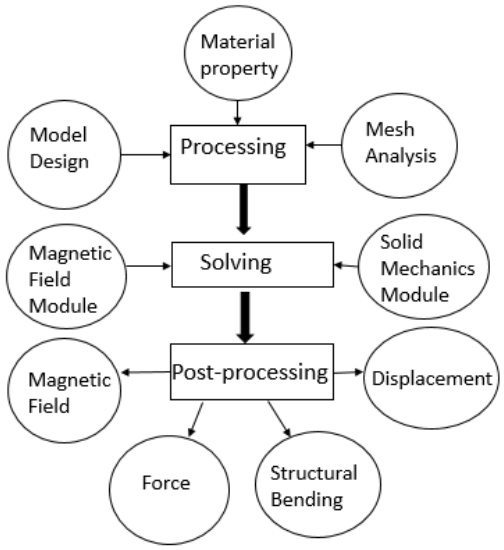


Fig. 1. FEM procedures involved in the simulation

III. MODEL DESIGN

A. *Flexible Tube*

A rubber tube of the height 13 cm and radius 0.7 cm has been considered for this study. The tube is filled with the ferrofluid which is a stable colloidal suspension of single-domain magnetic nanoparticles considering a relative permeability of 1.05, relative permittivity of 1, density of 1210 kg/m^3, Young's modulus of 0.05 GPa and 0.49 of Poisson's ratio.

B. *Electromagnet*

Standard Wire Gauge (SWG) 26 of 0.46mm diameter has been used as winding of the ferrite core (Ferrite F 3000 mu). The ferrite core is winded with 400 turns of wire to give the optimal force needed to create the required motion of the tube. The dimension of the electromagnet and the tube used are shown in Table I.

TABLE I. GEOMETRICAL SPECIFICATIONS OF THE ELECTROMAGNET AND THE RUBBER TUBE USED IN THE STUDY.

Geometrical parameter	**Length (cm)**
Diameter of Ferrite Core	0.4
Diameter of electromagnet	2.3
Height of electromagnet	2
Height of Ferrite Core	2
Height of tube	13
Inner diameter of tube	0.6
Outer diameter of tube	0.7

Three model designs built are shown in Fig. 2. Three different placement of coils, each with 2 cm gap between the opposite coils are used to show 7 different movement patterns of the ferrofluid tube. The coils are placed at different positions along the side of the tube and held in place by two plastic stands. For the first model shown in Fig. 2(a), three coils are placed 4 cm apart from one another, along each side of the stand. All the coils are given equal value of voltage. When one set was on the other set was turned off, so the tube deflected to the side which had the maximum magnetic field. For the second model shown in Fig. 2 (b), same arrangement of coil are used but coil 1 and 3, coil 2 and 4 had 2 cm distance between them. But coil 5 and 6 are 4 cm away from the other two coils in the same column. For the third model shown in Fig. 2 (c), the coils are placed at the three corners. The three coils are kept on, one at a time and the tube showed a three directional movement.

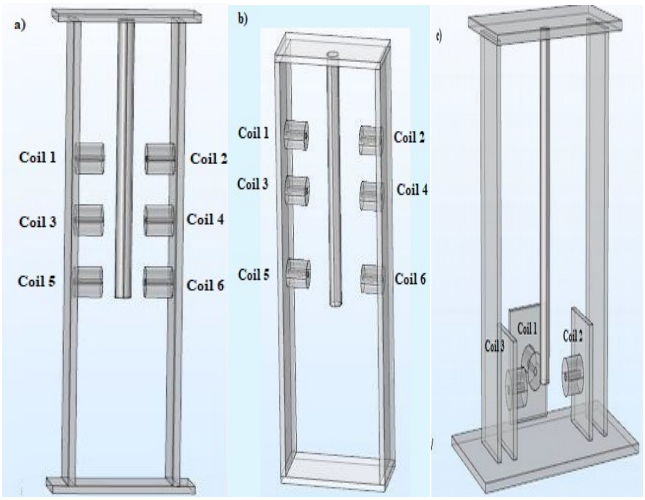


Fig. 2. Three different placement of electromagnet used in the study

IV. MATERIAL PROPERTY

Five important domains are considered in this simulation model. They are: (i) Ferrofluid inside the rubber tube, (ii) Copper for winding the electromagnet, (iii) Ferrite for electromagnet core, (iv)Rubber for the tube and (v) Air in the surrounding space. Individual properties of the materials used are described in the Table II.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

TABLE II. PROPERTIES OF THE MATERIALS USED IN THE STUDY

Materials	Relative Permeability, μ_r	Relative Permittivity, ε_r	Electrical Conductivity, σ (S/m)
Ferrofluid	1.05	1	0
Copper	0.99	1	5.99×10^{7}
Ferrite	3000	1	1×10^{-12}
Rubber	2.4	7	10×10^{-14}
Air	1	1	3×10^{-15}

V. SOLVER SETTINGS

The model design is solved using (i) Magnetic field and (ii) Solid Mechanics solvers.

A. Magnetic Field Interface

The magnetic field and force are the two factors needed to be considered first in the simulation, therefore the magnetic field module has been chosen. The current direction and the voltage applied to the electromagnet are also specified in this module.

The circular coil type with specific current direction is chosen in the first place for each coil winding. Ampere's law and force calculation interface were the next things added to the module. This module determines the magnetic flux density, magnetic flux lines, magnetic field vector direction and electromagnetic force for the electromagnets. It is concluded that, regardless of the current direction (clockwise or anti clockwise) the direction of magnetic force vectors applied to the tube show the same pattern. This is because the core material of the tube is filled with Ferrofluid, which is composed of "Paramagnetic" nano particles that shifts its polarity according to the magnetic field applied on it. Fig. 3 shows the current and magnetic field directions when clockwise and anti-clockwise current is added to the coils.

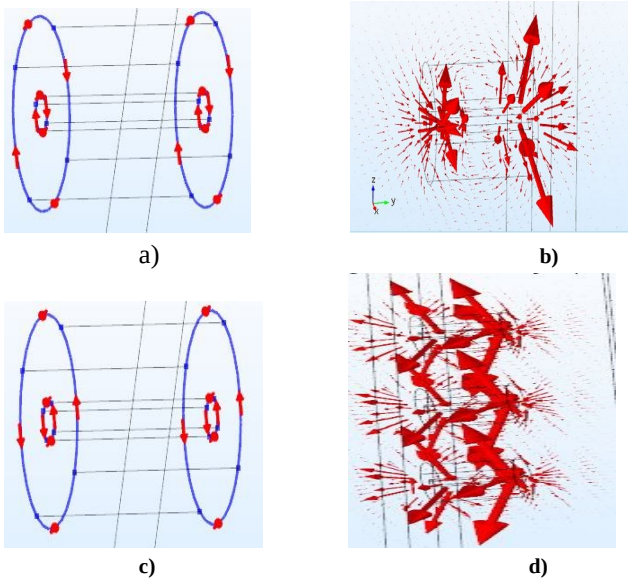


Fig. 3. a) Circular current path inside the electromagnet showing the direction of current flow when clockwise current is supplied across it, b) Directions of the magnetic field line vectors in the electromagnet for the current supply in clockwise direction, c) Circular current path inside the electromagnet showing the direction of current flow when anticlockwise current is supplied across it, d) Direction of the magnetic field line vectors around the electromagnet for the current supply in anticlockwise direction

B. Solid Mechanics Interface

To observe the different bending patterns of the ferrofluidtube this module is used. The fixed region along the tube together with the boundary load are the few parameters that are needed to be considered. The Boundary load used here is the electromagnetic force calculated from the previously discussed magnetic field module.

VI. MESH

One of the essential part of the pre-processing stage is selecting mesh for the model. Here triangular mesh is chosen i.e. the entire model is divided to non-overlapping triangular elements. Selecting mesh is the final stage of pre-processing after the model geometry has been drawn and the materials used is defined. In this design, an optimal physics controlled mesh with normal element size is selected to give a good accuracy of the results with less use of computational memory that in turn takes less processing time for the simulation without compromising the accuracy of the simulation results. The mesh for the model is shown in Fig. 4.

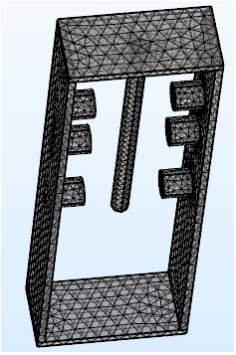


Fig. 4. Model design divided into non-overlapping triangular elements

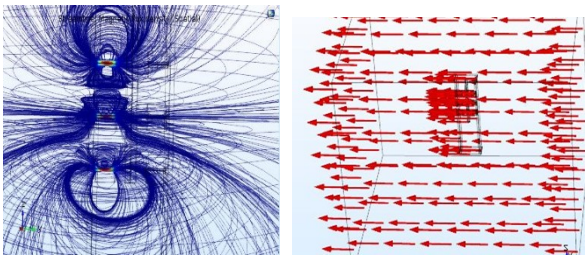
VII. RESULTS AND DISCUSSION

After all the parameters have been set, the simulation is evaluated and the result shows the various motion of the tube according to the magnetic field induced in the coils.

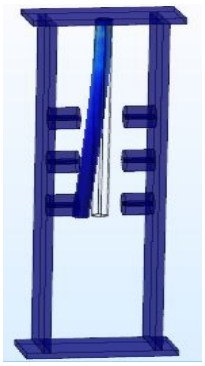
A. Flickering Motion

When the coils placed on the left hand side of the tube are given voltage supply of 9.2V, the tube shows the movement in the direction of the magnetic field introduced i.e. towards the left side. The results for the left flickering motion is shown in Fig. 5. Magnetic flux lines for the arrangement of coils are shown in Fig. 5(a), direction of force on the tube in Fig. 5(b) and finally the left hand flicker of the tube in Fig. 5(c).

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE



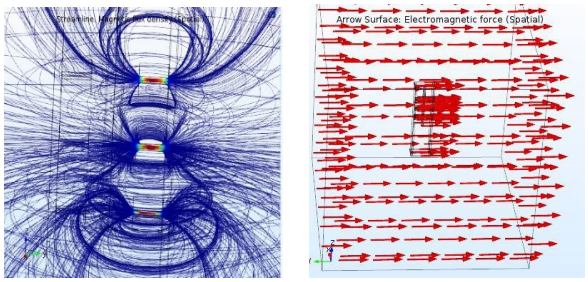
a) b)



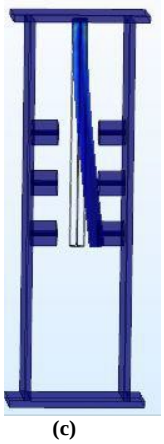
c)

Fig. 5. a) Magnetic flux lines surrounding the three electromagnets in the model design, b) Direction of magnetic force vectors surrounding the model, c) left flickering of ferrofluid tube when voltage supply is given to the left hand coils

The procedure is repeated with the coils on the right hand and the corresponding results are shown in Fig. 6. Magnetic flux lines for the arrangement of coils are shown in Fig. 6(a), direction of force on the tube in Fig. 6(b) and lastly the right hand flicker of the tube in Fig. 6(c).



(a) (b)



(c)

Fig. 6. a) Magnetic flux lines surrounding the three electromagnets in the model design, b) Direction of magnetic force vectors surrounding the model, c) Flickering of ferrofluid tube when voltage supply is given to the right hand coil

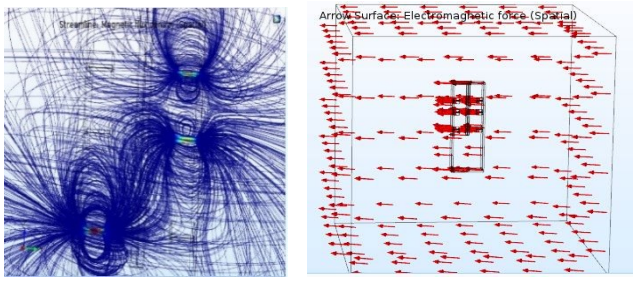
The electromagnetic force and the displacement of the tube for the flickering motion are shown in Table III.

TABLE III. ELECTROMAGNETIC FORCE AND DISPLACEMENT FOR FLICKERING MOTION

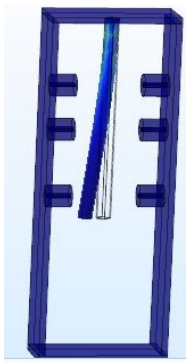
	Electromagnetic force components (N)			**Resultant Displacement (cm)**
	x	***y***	***z***	
Right flicker	0.29	0.68	-0.21	3.09
Left flicker	0.28	0.63	0.27	3.14

B. Bending Motion

When coil number 2 and 4 are given with 4.2V, whereas Coil number 5 is given 9.2V the tube shows a bending motion towards the left direction. The magnetic flux lines are shown in Fig. 7(a), direction of force on the tube in Fig. 7(b) and finally the left hand bending of the tube in Fig. 7(c).



(a) (b)



(c)

Fig. 7. a) Magnetic flux lines surrounding the three electromagnets in the model design, b) Direction of magnetic force vectors surrounding the model, c) Bending of ferrofluid tube in the left hand side

Alternatively, when coil number 1 and 3 are fed with 4.2V input signal, whereas Coil number 6 is given 9.2V input voltage, the tube shows a bending motion towards the right hand side. The magnetic flux lines surrounding the three electromagnets in the model design are shown in Fig. 8 (a), direction of force on the tube in Fig. 8(b) and finally the right hand bending of the tube in Fig. 8(c).

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

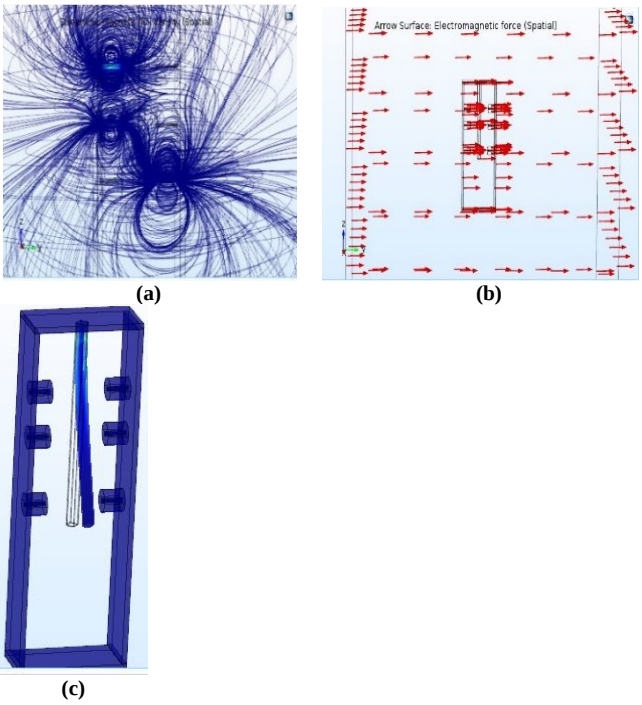


Fig. 8. a) Magnetic flux lines surrounding the three electromagnets in the model design, b) Direction of magnetic force vectors surrounding the model, c) Bending of ferrofluid tube in the right hand side

The numerical value of the electromagnetic force and the displacement of the tube for bending motion, both right and left are shown in Table IV.

TABLE IV. ELECTROMAGNETIC FORCE AND DISPLACEMENT FOR BENDING MOTION

	Electromagnetic force components (N)			**Resultant Displacement (cm)**
	x	***y***	***z***	
Right Bend	0.07	0.02	0.27	1.13
Left Bend	0.03	-2.94	7.37	1.18

C. Three directional movement

For three directional movement of the tube, all the three coils are given the same input voltage of 9.2V but one at a time. It is noticed that the tube moved to the coil which had magnetic field. Figure 9 (a) and (b) show the force direction and deflection of the tube when the coil placed at the middle is given 9.2V. Similarly, when the coil placed on the left side is given with 9.2V, the force and deflection of the tube to the left direction is shown in Fig. 9 (c) and (d). Again, when the coil placed on the right side in given with 9.2 V, It attracts the tube towards it. The force direction and the deflection is shown in the (e) and (f) of Fig. 9 respectively.

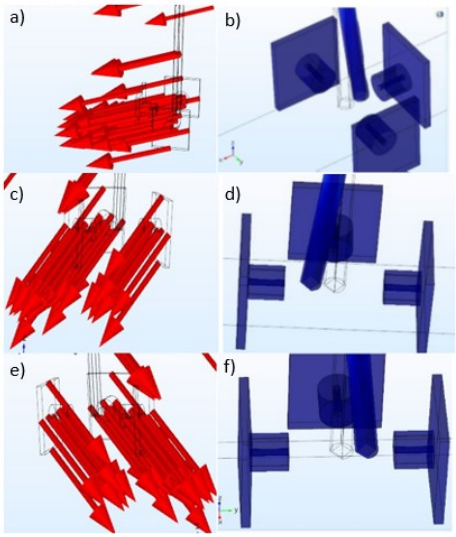


Fig. 9. a) Direction of the magnetic force vectors when middle coil is turned on, b) Position of the tube when coil placed at the middle is turned on, c) Direction of magnetic force vectors when left hand coil is turned on, d) Position of the tube when coil placed on the left side is turned on, e) Direction of the magnetic force vectors when right hand coil is turned on, f) Position of the tube when coil placed on the right hand side is turned onx

The electromagnetic force and the displacement of the tube for three directional movement are shown in Table V.

TABLE V. ELECTROMAGNETIC FORCE AND DISPLACEMENT OF THE TUBE FOR THREE DIRECTIONAL MOVEMENT OF THE TUBE

	Electromagnetic force components (N)			**Resultant Displacement (cm)**
	x	***y***	***z***	
Right Position	-0.04	-1.49	-0.38	1.18
Left Position	-0.05	1.49	-0.22	1.18
Back Position	-1.46	-0.06	0.49	1.18

VIII. CONCLUSION

Ferrofluid is widely used in many applications, starting from frictionless seals to maintain a vacuum while still allowing for moving or rotating parts. But using voltage control of electromagnet to control the vertical movement of the ferrofluid tube is completely a new approach. In this paper the simulation approach was able to calculate the electromagnetic forces and the displacement of the tube for the corresponding forces together with the various motion of the tube. This approach shows the flexibility of the ferrofluid tube when controlled magnetic field is applied to it from different directions. The simulation results obtained from this study will help refinement of a practical model and optimization of the parameters needed for building a prototype.

VIII. FUTURE WORK

From the results of this simulation approach, a practical implementation of the ferrofluid soft actuator can be built in near future that will have various medical applications as well as can be considered as soft robotic finger that has more degree of freedom and flexibility to handle tasks that rigid robots may unable to perform.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

ACKNOWLEDGMENT

The authors would like to show their deepest gratitude to Asian Institute of Technology for providing support to perform this simulation.

REFERENCES

[1] A. Chiolerio and Marco B. Quadrelli. Smart Fluid Systems: The Advent of Autonomous Liquid Robotics. Adv. Sci.2017. DOI: 10.1002/advs.201700036, 2017, pages 1-18.

[2] F. Ilievski, A.D.Mazzeo, R.F.Shepherd, X. Chen, and G. M. Whitesides, "Soft robotics for chemists," Angewandte Chemie, vol.123, 2011, pp.1930-1935.

[3] E. Brown, N. Rodenberg, J. Amend, "Universal robotic gripper based on the Jamming of granular material," Proceedings of the National Academy of Sciences, vol.107, no.44,2010, pp.18809- 18814, doi:10.1073/pnas.1003250107.

[4] Jr. J. H. Long, S. Combes, J. Nawroth, et al, "How Does Soft Robotics Drive Research Animal Locomotion," Soft Robotics, vol.1, no.3, pp.161-168, doi:10.1089/sero.2014.1502.

[5] R. Deimel and O. Brock, "A novel type of compliant, underactuated robotic hand for dexterous grasping," Robotics: Science and Systems, Berkeley, CA, 2014, pp.1687-1692.

[6] P. Polygerinos, Z. Wang, J. T. B. Overvelde, "Modeling of Soft Fiber-Reinforced Bending Actuators," IEEE Transactions on Robotics, vol.31, no.3, 2015, pp.778-789.

[7] N. W. Bartlett, M. T. Tolley, J. T. B. Overvelde, "A 3D-printed, functionally graded soft robot powered by combustion," Science, vol.349, no. 6244, 2015, doi:10.1126/science.aab0129.

[8] Sharon Glotzer, "Rise of the Colloidal Machines," physics.aps.org/ videos/2015-glotzer, 2017.

[9] Fionnuala Connolly, Conor J. Walsh, and Katia Bertoldi, "Automatic design of fiberreinforced soft actuators for trajectory matching," PNAS Early Edition, 2016.

[10] Hui Wang, Jungong Ma, Ziyu Ren, Zheyuan Gong, Yufei Hao, Tianmiao Wang, Li Wen, "Fiber-reinforced soft robotic anthropomorphic finger," International Conference on Robotics and Automation Engineering, 2016.

[11] Evrim Kurtoglu, Alp Bilgin, Muhsincan Sesen, Burc¸ Mısırlıoglu, Mehmet Yıldız, Havva Funda Yagcı Acar, Ali Kos ar, "Ferro fluid actuation with varying magnetic fields for micropumping applications," Springer-Verlag ,2012.

[12] Bangxiang Chen, Yanhe Zhu, Jie Zhao, and Hegao Cai, "Design of a prototype of an Adaptive soft robot based on ferrofluid," IEEE Conference on Robotics and Biomimetics,Zhuhai, China, 2015.

[13] Iliana Marinova, Valentin Mateev, Aneliya Tezova, "Electromagnetic Systems and Devices with Nano Ferrofluid,"2013.

[14] Valentin Mateev, Aneliya Terzova and Iliana Marinova, "Design Analysis of Electromagnetic Actuator with Ferrofluid," 18th International Symposium on Electrical Apparatus and Technologies (SIELA), 2014.

[15] Hernando Leon-Rodriguez, Viet Ha Le, Seong Young Ko, Jong-Oh Park, and Sukho Park, "Ferrofluid Soft-robot Bio-inspired by Amoeba Locomotion," 15th International Conference on Control, Automation and Systems (ICCAS 20 \5),2015.

[16] Dae Woong Oh, Dong Yoon Sohn, Doo Gyoon Byun, and Young Sun Kim, "Analysis Of Electromotive Force Characteristics and Device Implementation for Ferrofluid based Energy Harvesting System,"17th International Conference on Electrical Machines and Systems (ICEMS), Oct. 22-25, 2014.

[17] Camelia Petrescu, Radu Olaru and Radu Hertanu, "Ferro fluid Actuator with Levitating Nonmagnetic Disc," 7th International Symposium on

Advanced Topics in Electrical Engineering, 2011.

[18] B. Assadsangabi, M.H. Tee, and K. Takahata "Ferrofluid-assisted Levitation Mechanism for Micromotor applications," IEEE transducers, 2013.

[19] Ion Valentin Gurgu, Marius Giorgian Ionita, Ion Vasile Dinu Coltuc, Ioan Alexandru Ivan, Andrei "Simulation method and measurement system of electromagnetic force used in micromanipulation systems," ECAI 2017 - International Conference – 9th Edition Electronics, Computers and Artificial Intelligence 29 June -01 July 2017.

[20] Lee Jia Ji, Pua Chang Hong and Lee Poh Foong, "Investigation of feasibility of conventional magnetic directing ferrofluid targeting cancer treatment," 2nd International Conference on Biomedical Engineering (ICoBE), 30-31 March 2015.

[21] Mengfei Yang, Robert O'Handley and Zhao Fang, "Modeling of Ferrofluid Passive Cooling System," Excerpt from the Proceedings of the COMSOL Conference, 2010.

[22] M. S. Kamarudin, N. H. Radzi, A. Ponniran and R. Abd-Rahman , " Finite Element Modelling of Electric Field and Voltage Distribution on a Silicone Insulating Surface Covered with Water Droplets," IEEE Transactions on Dielectrics and Electrical Insulation Vol. 25, No. 2; April 2018.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Design of a Prototype of a Ferrofluid Based Soft Actuator

Tasnova Tanzil Khan
Mechatronics, Department of Industrial Systems Engineering, Asian Institute of Technology
Pathumthani, Thailand
tasnovatanzil@gmail.com

A. M. Harsha S. Abeykoon
Mechatronics, Department of Industrial Systems Engineering, Asian Institute of Technology
Pathumthani, Thailand
harsha@ait.ac.th

Tanujjal Bora
Nanotechnology, Department of Industrial and Systems Engineering, Asian Institute of Technology
Pathumthani, Thailand
tbora@ait.ac.th

Golam Fahad Mullick
Department of Electrical and Electronics Engineering, American International University Bangladesh
Dhaka, Bangladesh
fahadmullick89@gmail.com

***Abstract*—A ferrrofluid based soft actuator which shows various vertical motions has been introduced for the first time in this paper. The flexibility as well as the potential of the prototype designed, is tested in this study. The soft actuator used here is made by injecting ferrofluid inside a soft and thin rubber tube. The system also consists of some electromagnets placed at different positions along the side of the tube, whose magnetic field intensity is controlled by changing the duty cycle of the input voltage supplied to the electromagnets. The position of the coils as well as the input voltage are varied during the experiment which created various vertical motions of the ferrofluid tube. The simple design of the prototype described in this paper made this system more cost effective and easy to configure.**

Keywords—soft actuator, ferrofluid, electromagnet, voltage control

I. INTRODUCTION

Soft robots have shown its astounding motion through its various pattern of movements including stretching, twisting, bending and expanding in the multifarious environment. In comparison to the widely known rigid robots, soft robot has shown a remarkable adaptation to the surrounding environment together with more flexibility and can be readily made [1]. As a result of their flexible nature, soft robots can easily alter its shape when subjected to an outside force but of course assuring the safety of itself and the user [2]. Often, soft robots have one or more flexible parts that take on the vital part of the robot and deform actively or passively under external pressure or force to acquire the desired movements [3].

Soft and flexible materials with which some of these robots are made are called smart materials commonly called Shape Memory Alloy (SMA) or material made of wax and foam [4] [5]. In 2015, a bioinspired Octopus has been created with SMA [6]. In the year 2016, two major innovations on artificial soft finger have been sited, one is a pneumatic fiber-reinforced actuator where the air pressure and external fiber arrangement replicated complex motions of the index and thumb of a human being [7], another is a soft anthropomorphic finger made with a low cost soft hyperplastic material which made a breakthrough in the world of robotics [8].

Beside these approaches, magnetic method of actuation where found to be more effective and shows noticeable movements keeping the softness of the structure unaffected. Therefore a combination of magnetism together with soft materials can create an effective system that can undergo remote as well as harsh environment [3].

A magnetic fluid named as ferrofluid is used as the main structural element for the soft robot developed in this study. Ferrofluid falls in the category of smart material that contains magnetite particles suspended in a liquid or oil and can be magnetized in presence of any magnetic field around it [9]. When a magnetic field is applied to this smart fluid, the nano particles in the fluid together with the carrier fluid accumulate towards the area that has exposed to the strongest magnetic field, this is because the gathered layer of the ferrofluid experience an external pressure caused by the generated magnetic field around it [10]. This concept is used by the researchers in developing ferrofluid based soft robots. Previously, ferrofluid has been injected inside a polymeric soft ball to make a soft ball which could pass through various obstacles and change shape accordingly [11]. A pumping devices made with ferrofluid which changed its directions according to the change in magnetic field has also been developed using permanent magnet [12]. Moreover, controlling this magnetic property of the magnet together with the movement of the fluid, researchers could successfully make the robot transport things placed on the surface of a sheet from one place to the other [13]. In 2015,

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

a system using magnetic material to control the drug delivered to the human body made the use of these magnetic material more pronounced [14]. A microcontroller together with a L298N motor driver were used to control the motors placed on each side of the patient. Basics C programmed algorithm was constructed to control the magnet in the desired path so that the drug reaches the accurate disease area. Moreover, Ferrofluid assisted levitation mechanism [10] and the levitating nonmagnetic disc inside ferrofluid [15] are some major applications of ferrofluid during the last decades.

However, most of the studies that includes ferrofluid as the core material, either used permanent magnet [11, 12, 14, 16, 17] or both permanent and electromagnet [9, 10, 15] to move the fluid from one point to the other or moving the object containing the fluid forward and backward in horizontal direction. A very few research has been done on vertical deflection of ferrofluid by varying magnetic field strength of electromagnet. Therefore, the aim of this thesis is to make a ferrofluid based soft actuator that uses the concept of magnetism in creating vertical movements of a tube that is filled with ferrofluid.

In this study, a soft and flexible rubber tube of about 13 cm was filled with ferrofluid and subjected to a varying magnetic field. The magnetic field was created by the electromagnets kept at about 2 cm distance apart from the tube. As the magnetic field strength of the electromagnets are varied the tube shows different movement patterns in the vertical position. The electromagnets are controlled by a microcontroller module (Arduino Mega) together with a L298N motor driver that supplied the voltage to drive the electromagnets. A basics C programmed algorithm are also constructed to control the voltage supplied to the electromagnets so that the tube moves in the desired direction.

II. DESIGN& FABRICATION

A. *Thin Rubber Tube*

While designing the prototype the first thing that was taken into consideration was the ability of the robot to deform or change its shape like liquid, which leads to a noticeable softness and flexibility of the structure, when controlled magnetic force is applied to it. On the other hand, it will be undesirable if the robot when losing its control, sheds into droplets like liquid. Therefore, considering all the necessary conditions ferrofluid has been chosen as the core material of the robot as it can be controlled by external magnetic force and at the same time can attain the liquid characteristics. So, taking all the necessary requirements into consideration, the prototype has been designed. The shell of robot is made of the commonly known polymeric rubber which is mostly used in toy balloons. A rubber tube of 13 cm height, 0.6 cm inner and 0.7 cm outer radius is filled with the chosen ferrofluid which is stable colloidal suspensions of single-domain magnetic nanoparticles in 55-91% of distillate. The nanoparticles in ferrofluid used here are usually 3-5% of Magnetite (Fe_3O_4) particles, suspended in Petroleum as distillate.

B. *Electromagnet*

In order to make the tiny electromagnets of desired number of turns, Standard Wire Gauge (SWG 26) of 0.46mm diameter has been used in winding the ferrite core. The specification of wire and ferrite are shown in Table I. The ferrite was winded with 400 turns to give the optimal magnetic force needed to create the required motion of the tube.

The appearance of the electromagnet together with the flexible ferrofluid tube used in this study is shown in Fig. 1.

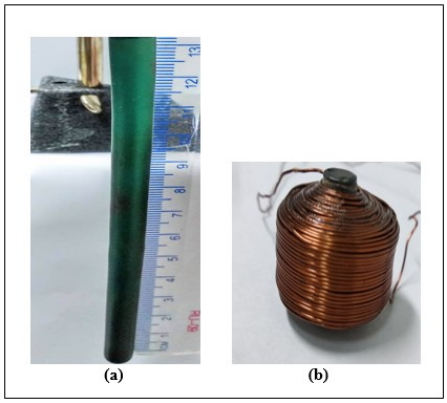


Fig. 1. a)The initial shape and dimension of the flexible tube, b) Electromagnet used in the model design

TABLE I. SPECIFICATIONS OF THE ELECTROMAGNET USED IN THE EXPERIMENT

Geometrical parameters	**Length (cm)**
Core material of electromagnet	Ferrite F 3000 mu
Height of the core in the electromagnet	2
Diameter of the core in the electromagnet	0.4
Diameter of electromagnet	2.3
Height of electromagnet	2

III. MODEL DESIGN

Three different models were designed for seven different movement patterns. The arrangement of the coils along with the input voltage are varied in such a way to give different motions to the ferrofluid tube.

A. *Placement of Electromagnets for Vertical Deflection Pattern*

Three electromagnets where placed on each side of the two stands placed on opposite of each other. The electromagnets are placed with 2 cm gap between the opposite one, keeping 4 cm gap between each consecutive electromagnets placed in the same column. The arrangement of the electromagnets are shown in Fig. 2.

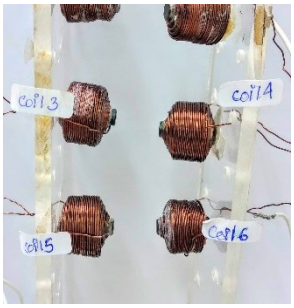


Fig. 2. Placement of electromagnet for the vertical deflection pattern of the tube

B. *Placement of Electromagnets for Bending Pattern*

The same arrangement of coil as that used for the deflection pattern is used here but coil 1, 3 and 2, 4 have 2 cm distance between them. But coil 5 and 6 are placed 4 cm away from the other two coils of the same column.

C. *Placement of Electromagnets for three directional movement*

For the three directional motion, three coils of the same number of turns as before are placed at the three corners of the ferrofluid tube and about 2 cm away from each other. Shown in Fig. 3.

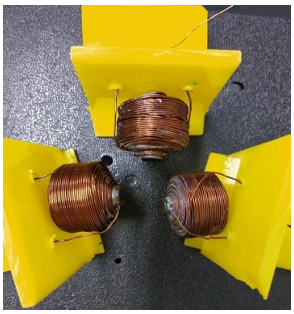


Fig. 3. Placement of electromagnet for the three directional movement pattern

IV. VOLTAGE CONTROLLER

The arrangement of coils are connected to a voltage controller circuit that includes three L298N motor drivers which is connected to an Arduino Mega microcontroller. The voltage of each coil are varied by varying Pulse Width Modulation (PWM) of the input signal which correspondingly causes different motions of the tube containing the ferrofluid. The arrangement of the circuit and the connection diagram are shown in Fig. 4.

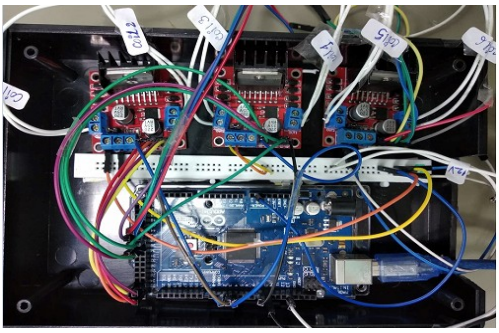


Fig. 4. Circuit design for Voltage control

The duty cycle of the input PWM signal for each of the electromagnet are varried in such a way that gives a controlled automatic movement of the soft tube containing the ferrofluid. The Logic control used to achieve the desirable movement of the tube in various direction is potraied in Fig. 5.

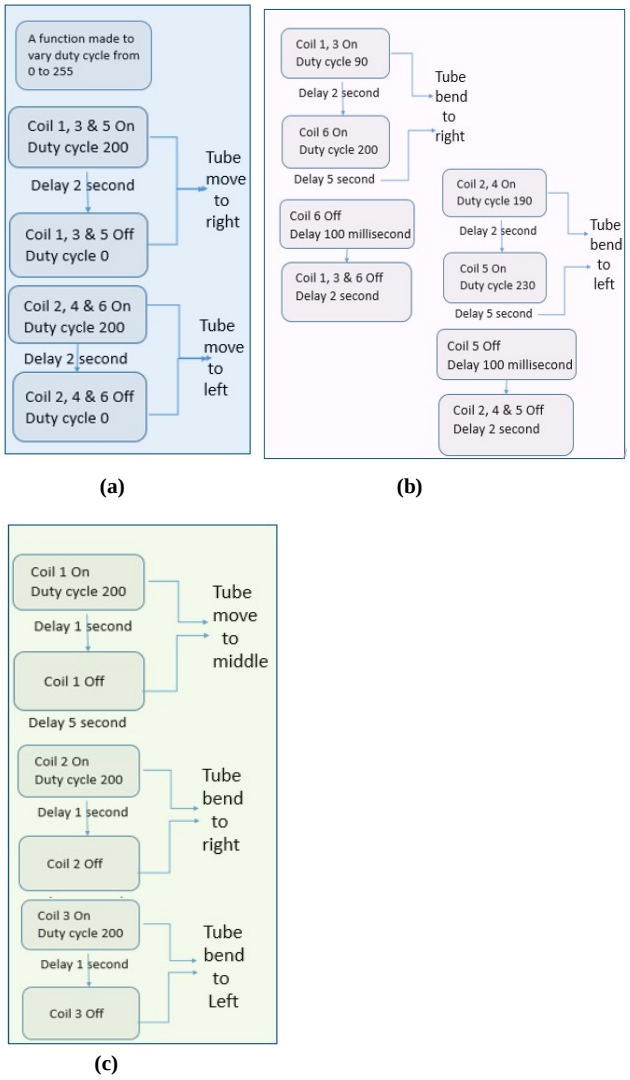


Fig. 5. Flow chart showing each steps used in the logic control of voltage used in the microcontroller

V. EXPERIMENT AND DISCUSSION

This section will investigate the performance of the ferrofluid tube when varied magnetic field is applied to it.

A. *Vertical Deflection*

When 78% duty cycle of PWM signal i.e. is 9 V of supply is given to Coil 2, Coil 4 and Coil 6 the tube showed a vertical motion on the right hand side towards the direction of the magnetized coils, shown in Fig. 6(a). When the same supply is given to all the electromagnets placed on the left hand side of the tube, i.e. to Coil 1, Coil 3 and Coil 5, the tube shows a vertical deflection but now to the left hand side, towards the magnetized coils which is shown in Fig. 6(b). The same procedure is repeated but with varying the amplitude of the PWM signal from a maximum value to zero, the tube shows the same motion but with more smooth movement but shows a longer reaction time. The displacement of the tube from its original direction are measured approximately to be 1cm toward each direction.

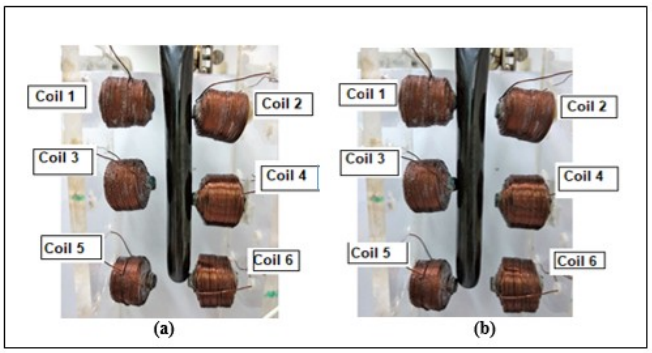


Fig. 6. Vertical deflection of the tube containing ferrofluid

B. *Bending motion*

When coil 2 and coil 4 are fed with 4.2V i.e. 35% of duty cycle of PWM signal, whereas Coil 5 are given 78% (9V) duty signal, the tube shows a bending motion towards left hand side shown in Fig. 7(a). Alternatively, when coil 1 and coil 3 are given 4.2V i.e. 35% of duty cycle of PWM signal, whereas Coil 6 is given 78% (9V) duty signal, the tube bend towards the right hand side shown in Fig. 7(b) The tip of the tube shows a displacement of approximately 1 cm from its original position towards the magnetized coils.

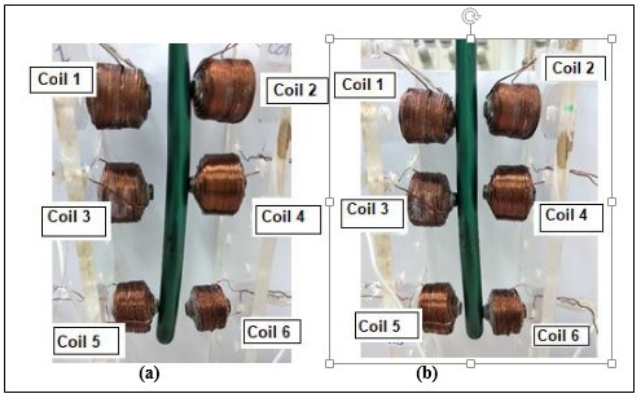


Fig. 7. Bending of the flexible tube in both direction

C. *Three directional motion*

When all three coils placed on the three corners of the tube are given an input voltage of 9.2V i.e. 78% duty cycle but one at a time, the tube shows movement to the direction of the electromagnet which is magnetized. The Fig. 8 shows the three different movement pattern of the tube.

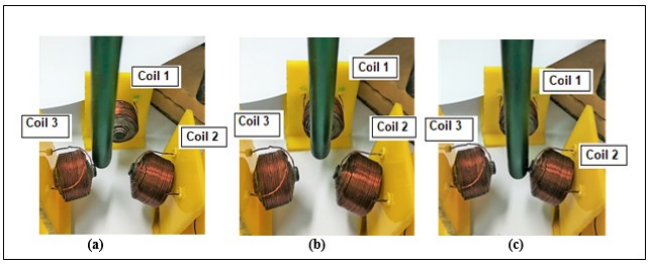


Fig. 8. Three directional movement of the flexible tube

VI. CONCLUSION

This concept of a ferrofluid soft actuator showing vertical motions is a new approach in the field of soft robotics. The simplicity of the design make the system more cost effective and readily made. The fabrication of the system is easy and the ability of the robot showing various vertical motions is fantastic. The different experiments performed in this study shows that intensity of the magnetic field greatly affects the shape of this soft robot and the deforming characteristics of the tubing material determines its flexibility. Therefore, it can be concluded that the magnetic field intensity together with the flexibility of the structure itself contributes in the moving capability of the ferrofluid tube.

This prototype is an initial proof of concept that ferrofluid can show vertical motions when controlled by small electromagnets. It has been created for the future design of a soft finger by using ferrofluid as the core material. In this initial stage, the tube is not much rigid to resemble a soft finger which could grab or even apply force to objects; but it shows some waving and bending motions which is indeed a major contribution to the design of the future soft finger.

VII. FUTURE WORK

This paper mainly focuses on the realizability of a prototype of ferrofluid based soft actuator and its performance with varying magnetic field. For future investigational approach, the ferrofluid used in this study can be combined with a liquid monomer and form into a polymer with magnetic properties (for example, magnetic rubber/plastic) or can be incorporated with a non-toxic polymer, Polydimethylsiloxane (PDMS), commonly known as dimethylpolysiloxane or dimethicone. This incorporation will make the structure more rigid and flexible with more degrees of freedom. In case of soft robots or actuators, the flexibility and rigidness of the structure (PDMS embedded with the ferromagnetic particle) can create a wide range of motions and bending when a externally controlled magnetic field is applied, which can lead even to a soft robotic hand.

ACKNOWLEDGMENT

The authors would like to show their gratefulness to Asian Institute of Technology for providing their support to perform the experiments needed for this study.

REFERENCES

[1] R. F. Shepherd, F. Ilievski, W. Choi, S. A. Morin, A. A. Stokes, A. D. Mazzeo, X. Chen, M. Wang, and G. M. Whitesides, "Multigait soft robot," Proceedings of the National Academy of Sciences, vol. 108, pp. 20400-20403, 2011

[2] H. Lipson, "Challenges and opportunities for design, simulation, and fabrication of soft robots," Soft Robotics, vol. 1, pp. 21-27, 2014

[3] T. S. Zhang, A. Kim, M. Ochoa, and B. Ziaie, "Controllable 'somersault'magnetic soft robotics," in Micro Electro Mechanical Systems (MEMS), 2015 28th IEEE International Conference on, pp. 1044-1047, 2015

[4] J. C. Nawroth, H. Lee, A. W. Feinberg, C. M. Ripplinger, M. L. McCain, A. Grosberg, J. O. Dabiri, and K. K. Parker, "A tissue-engineered jellyfish with biomimetic propulsion," Nature biotechnology, vol. 30, pp. 792-797, 2012

[5] N. G. Cheng, A. Gopinath, L. Wang, K. Iagnemma, and A. E. Hosoi, "Thermally Tunable, Self-Healing Composites for Soft Robotic Applications," Macromolecular Materials and Engineering, vol. 299, pp. 1279-1284, 2014

[6] M. Cianchetti, M. Calisti, L. Margheri, M. Kuba, and C. Laschi, "Bioinspired locomotion and grasping in water: the soft eight-arm OCTOPUS robot," Bioinspiration & biomimetics, vol. 10, p. 035003, 2015

[7] Fionnuala Connolly, Conor J. Walsh, and Katia Bertoldi, "Automatic design of fiberreinforced soft actuators for trajectory matching",PNAS Early Edition, 2016

[8] Hui Wang, Jungong Ma, Ziyu Ren, Zheyuan Gong, Yufei Hao, Tianmiao Wang, Li Wen, " Fiber-reinforced soft robotic anthropomorphic finger" , International Conference on Robotics and Automation Engineering, 2016

[9] Dae Woong Oh, Dong Yoon Sohn, Doo Gyoon Byun, and Young Sun Kim, "Analysis Of Electromotive Force Characteristics and Device Implementation for Ferrofluid based Energy Harvesting System", 17th International Conference on Electrical Machines and Systems (ICEMS), pp. 22-25, October 2014

[10] B. Assadsangabi, M.H. Tee, and K. Takahata "Ferrofluid-assisted Levitation Mechanism for Micromotor applications", IEEE transducers, 2013

[11] Bangxiang Chen, Yanhe Zhu, Jie Zhao, and Hegao Cai**, "**Design of a prototype of an Adaptive soft robot based on ferrofluid", IEEE Conference on Robotics and Biomimetics,Zhuhai, China, 2015

[12] Evrim Kurtoglu, Alp Bilgin, Muhsincan Sesen, Burc¸ Mısırlıoglu, Mehmet Yıldız, Havva Funda Yagcı Acar, Ali Kos ar, **"**Ferro fluid actuation with varying magnetic fields for micropumping applications". Springer-Verlag ,2012

[13] Tadayuki Tone, Francesco Visentin and Kenji Suzuki **, "**Sheet Type Soft Robot with Magnetic Fluid for Object Transportation", IEEE International Conference on Automation Science and Engineering (CASE) Taipei, Taiwan, pp. 18-22, August 2014

[14] Lee Jia Ji, Pua Chang Hong and Lee Poh Foong**, "** Investigation of feasibility of conventional magnetic directing ferrofluid targeting cancer treatment", 2nd International Conference on Biomedical Engineering (ICoBE), 30-31 March 2015

[15] Camelia Petrescu, Radu Olaru and Radu Hertanu, "Ferro fluid Actuator with Levitating Nonmagnetic Disc", 7th International Symposium on ADVANCED TOPICS IN ELECTRICAL ENGINEERING, 2011

[16] Hernando Leon-Rodriguez, Viet Ha Le, Seong Young Ko, Jong-Oh Park, and Sukho Park, " Ferrofluid Soft-robot Bio-inspired by Amoeba Locomotion", 15th International Conference on Control, Automation and Systems (ICCAS 20 \5), 2015

[17] Mengfei Yang, Robert O'Handley and Zhao Fang, "Modeling of Ferrofluid Passive Cooling System",.Excerpt from the Proceedings of the COMSOL Conference, 2010

Stable walking of an underactuated bipedal robot with unbalanced masses using one actuator

*

Mohammad Farhan Ferdous
Founder and Chairman
Japan-Bangladesh Robotics and Advanced Technology Research Center (JBRATRC)
Fukui, Japan
jbratrc2018@gmail.com

Jannat Binta Alam
Technical Member
JBRATRC
Dhaka, Bangladesh
jannat.cse.ewu@gmail.com

***Abstract*—The objective of the research is a stable gait generation on the level ground and some sloping ground of the 3-link biped robot with unbalanced masses using one actuator. We know the biped robot need actuator for stable walking on level ground. A biped robot with a torso model at least two actuators to achieve a stable walk. In the research we have investigated the only one motor which is very much challenging to achieve stable walking gait without the robot falls down, also there is used unbalanced masses where is very much difficult to achieve balanced walking. In this research, the numerical simulations show the robot with unbalanced masses can walk efficiently on level ground by only using one actuator. The robot can generate 2-period and 4-period gait without fall down. First, we introduced the 5-DOF (Degree of freedom) unbalanced masses model where used one actuator that walks the level ground. Second, describes the control system of Proportional-derivative (PD) controller between stance leg and torso, while the swing leg does not use any actuator and it does remain free. Third, we perform numerical simulations and show that the robot can generate a stable walk on level ground and some sloping ground by using one actuator. Furthermore, we conduct some parametric study to investigate the more natural human-like motion of the robot and achieved different results.**

***Index Terms*—Unbalanced masses, bipedal robot, one actuator, underactuated.**

I. Introduction

Since several years ago, humans have tried to make robots similar to humans. To build a humanoid robot is not an easy task due to the complexity in the design and control. Recent research has been shown that the human and animals structure is the best option to use for legged robots [1]. There is a tendency in mobile robotics to develop robots transported by legs; an advantage of using them is that legged robots may easily move in places designed for humans and it is better than wheeled robots. In this paper, we consider the unbalanced masses upper body biped robot which could better be understanding the biped locomotion in both human and robot. In 1989, McGeer started systematic research into "Passive dynamics walking mechanism", which walks down a shallow slope and need no motors and no control [2]. This research is very much inspired by walking the robotic field. The planer biped robot with torso can walk efficiently on level ground over a wide range of speed by only using hip actuators [3]. The rolling model was developed for semi-circular feet with hip and leg masses [4]. Narukawa [3] shows that swing leg control to enable the robot to walk on level ground. Furthermore, the biped robot also assists in the development of new rehabilitation devices with some control input. The energy efficiency is very much necessary for the biped robot. However, energy efficiency is a problem for many biped robots [5]. Besides the torso control; we investigate the change of parameters value and achieved different types of output through numerical simulation. Besides the torso control, Hobbelen [6] also used ankle control to allow the biped robot on level ground. The underactuated robot with the torso model can achieve stable gait on the slippery road surface [7]. Harata studied two asymmetric models in 2013. This model introduced a wheel without frame and the last model was a compass-like model with a curved tip. He showed the effect of asymmetries efficiency and energy consumption [8].

In this paper, we have developed a simple control strategy for an underactuated biped robot with unbalanced masses. In this technique, there is used only one PD controller which worked between the torso and stance leg, that time while the swing leg remains totally free. This controller utilizes both the torso different state states such as angle, angular velocity)and those of the stance leg. The numerical simulations display that, by applying this controller, a biped robot is able to walk stably on level ground and also slope ground. Thereafter, we searched for the differently desired angle and get the different types of result with suitable control parameters.

II. Mathematical Model of the Underactuated Bipedal Robot with Unbalanced Masses

A. Modeling of the robot

Fig.1 demonstrates the model of underactuated planar 5-DOF biped unbalanced masses robot with a humanoid torso body link. This is a straight vertical connection whose center of mass is arranged at the hip joint as appeared in Fig 1. Motion

is obliged to the sagittal plane, is made out of two identical rigid legs. The robot strolls on level ground by the symmetric walking approach. The two legs are associated deliberately to the hip joint. The model of the robot walks on level ground with unbalanced masses. Here, (x, z) is the end position of the stance leg θ_1 is the angle of the stance leg, θ_2 is the angle of the thigh of the swing leg and θ_3 is the angle of the torso. There are two legs different masses m_1 [kg] and m_2 [kg] and m_H [kg], m_t [kg] is the mass of position leg, swing leg and hip joint, torso individually. L is the complete length of every leg where $L = (a + b)$ [m], torso length L_t and g $[m/s^2]$ is the increasing speed because of gravity. The robot can walk on level ground since it utilized one actuator. Fig. 2 demonstrates the control model of the biped robot. We expect that this model can apply one torque u which is connected between the torso and stance leg.

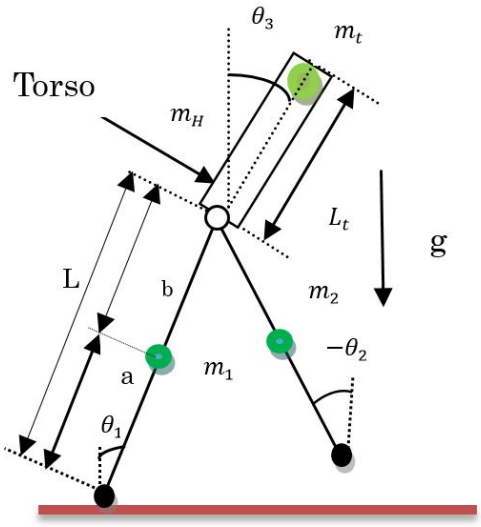


Fig. 1. Underactuated biped robot with unbalanced masses

B. Equations of motion

We pick two legs angles θ_1 and θ_2 which the two angles are appreciation to vertical to the dimension ground. Here, θ_H = θ_1 - θ_2 is the relative hip joint point, θ_3 is the vertical angle with torso. It is important to observe that unbalanced masses requires two individual equation of motions, one for walking around left stance leg and other walking around right stance leg. Walk lopsided of each improvement make unmistakable bearings, and two conditions developments have handled this issue. We choose the stance leg introductory advance as the left leg, which transforms into the swing leg of the second step and other is the right stance leg. The distinction in different masses between two legs have then again happened.

Let $q = \begin{bmatrix} x & z & \theta_1 & \theta_2 & \theta_3 \end{bmatrix}^T$ be the generalized coordinate vector. Where we can see there are 5-DOF of the robot.

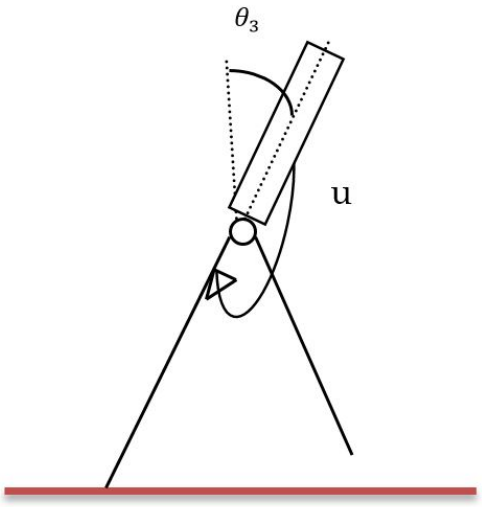


Fig. 2. Control model of the robot

The equations of motion can be derived using the Lagrangian dynamics. The underactuated biped robot with unbalanced masses equation of motion then becomes

$$\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \boldsymbol{S}\boldsymbol{u} + \boldsymbol{J}^{\mathrm{T}}\boldsymbol{\lambda} \tag{1}$$

where left side indicated that $\boldsymbol{M}(\boldsymbol{q})$ is 5x5 inertia matrix, $\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ is the centrifugal matrix and gravitational force. The right hand side indicated that Su is control input vector. $\boldsymbol{J}\epsilon R^{2\mathrm{x}5}$ is a Jacobian matrix and $\boldsymbol{\lambda}$ is ground reaction force.

The details of the left-hand side for left stance leg are as follows.

$$\boldsymbol{M}_L(q) = \begin{bmatrix} M_{11} & 0 & M_{13} & M_{14} & M_{15} \\ 0 & M_{22} & M_{23} & M_{24} & M_{25} \\ M_{31} & M_{32} & M_{33} & M_{34} & 0 \\ M_{41} & M_{42} & M_{43} & 0 & M_{45} \\ M_{51} & M_{52} & M_{53} & 0 & M_{55} \end{bmatrix},$$

$$\boldsymbol{h}_L(q, \dot{q}) = \begin{bmatrix} h_{11} \\ h_{21} \\ h_{31} \\ h_{41} \\ h_{51} \end{bmatrix}, S = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

,

$$M_{11} = m_1 + m_2 + m_b + m_h,$$
$$M_{13} = M_{31} = (am_1 + L(m_2 + m_h + m_b))cos\theta_1,$$
$$M_{14} = M_{41} = -bm_2cos\theta_2,$$
$$M_{15} = L_h m_b cos\theta_3,$$
$$M_{22} = m_1 + m_2 + m_b + m_h,$$
$$M_{23} = M_{32} = -(am_1 + L(m_2 + m_h + m_b))sin\theta_1,$$
$$M_{24} = M_{42} = bm_2 sin\theta_2,$$
$$M_{25} = -L_h m_b sin\theta_3,$$
$$M_{33} = a^2 m_1 + L^2(m_2 + m_h + m_b),$$
$$M_{34} = M_{43} = -bLm_2 cos(\theta_1 - \theta_2),$$
$$M_{35} = LL_h m_b cos(\theta_1 - \theta_3),$$

$$M_{45} = L_h^2 m_b,$$
$$h_{11} = -\dot{\theta_1}^2(am_1 + L(m_2 + m_h + m_b)sin\theta_1 + b\dot{\theta_2}^2 m_2 sin\theta_2 - b\dot{\theta_2}^2 m_2 sin\theta_2,$$
$$h_{21} = 4g(m_1 + m_2 + m_h) - \dot{\theta_1}^2(am_1 + L(m_2 + m_h)cos\theta_1 + d\dot{\theta_3}^2 L_b m_b cos\theta_3,$$
$$h_{31} = -g(am_1 + L(m_2 + m_b + m_h))sin\theta_1 - b\dot{\theta_2}^2 Lm_2 sin(\theta_1 - \theta_2),$$
$$h_{41} = bm_2\dot{\theta_1}^2 Lsin(\theta_1 - \theta_2) + gsin\theta_2,$$
$$h_{51} = -L_h m_b \dot{\theta_1}^2 Lsin(\theta_1 - \theta_3) + gsin\theta_3.$$

The other dynamic equation with right stance leg can be solved in a similar way

The velocity constraint condition for stance leg end is given by

$$\dot{x} = 0, \dot{z} = 0. \tag{2}$$

The Jacobian matrix $J \in \Re^{2\times 5}$ in Eq. (1) then becomes

$$J = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{3}$$

J is the constant vector and $J\ddot{q}$ = 0 holds. The Lagrange undetermined multiplier $\lambda \in R$ Eq. (1) can be solved as

$$\lambda = (JM^{-1}J^T)^{-1}(JM^{-1}(h - Su)) \tag{4}$$

C. Combined hybrid system

The hybrid system used when the two legs masses are different In this system consists one walking steps has two phases: a continuous phase and a discontinuous impact phase. In these walking, there are two different phases: a continuous swing and a separated collision phase. This type of system called a combined hybrid system. The two separate impact or collision equations necessary because of the two different masses in the two legs. The robot can continue walking in trajectories and its flows Combined hybrid dynamic system.

D. Impact or collision equation

The inelastic impact or collision equation with left stance leg described by

$$M(q)\dot{q}^+ = M(q)\dot{q}^- + J_I(q)^T\lambda_I \tag{5}$$

$$J_I(q)\dot{q}^+ = 0_{2\times 1} \tag{6}$$

Where $J_I(q) \in \Re^{2\times 5}$ is the Jacobian matrix determined by the horonomic constraint condition immediately after impact and λ_I is the Lagrange undetermined multiplayer vector that describe the impulse.

Let $(\bar{x}, \bar{z})$ the end position of swing leg.

Then we can $(\bar{x}, \bar{z})$ specify by

$$\begin{bmatrix} \bar{x} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} x + Lsin\theta_1 - Lsin\theta_2 \\ z + Lcos\theta_1 - Lcos\theta_2 \end{bmatrix} \tag{7}$$

After Eq. (7) time derivative then becomes

$$= \begin{bmatrix} \dot{x} + L\dot{\theta}_1 cos\theta_1 - L\dot{\theta}_2 cos\theta_2 \\ \dot{z} - L\dot{\theta}_1 sin\theta_1 + L\dot{\theta}_2 sin\theta_2 \end{bmatrix} \tag{8}$$

Rearranged the Eq. (8) then the Jacobian matrix becomes

$$J_I(q) = \begin{bmatrix} 1 & 0 & Lcos\theta_1 & -Lcos\theta_2 & 0 \\ 0 & 1 & -Lsin\theta_1 & Lsin\theta_2 & 0 \end{bmatrix} \tag{9}$$

The Lagrange multiplyer $\lambda_I \in \mathbb{R}^2$ after solve the Eqs. (5) and (6)

$$\boldsymbol{\lambda}_I = -\boldsymbol{X}_I(\boldsymbol{q})^{-1}\boldsymbol{J}_I(\boldsymbol{q})\dot{\boldsymbol{q}}^- \tag{10}$$

Where

$$X_I(q) := J_I(q)M(q)^{-1}J_I(q)^T \tag{11}$$

. By putting the Eq. (11) into Eq. (5), we get

$$\dot{q}^+ = (I_5 - M(q)^{-1}J_I(q)^T X_I(q)^{-1}J_I(q))\dot{q}^- \tag{12}$$

where $\dot{q}^+$ means velocity vector immediately after impact or collision.

We can solve another impact or collision equation with right stance leg by a similar procedure.

III. Control system

In the robot model, the torso as like an inverted pendulum, which is stabilized is needed for stable walking motion. In this research, a PD controller is used for stable gait generation of the robot where using only one actuated.

The legs are assumed to keep the vertical which follow the condition θ_1=θ_2=0 and $\dot{\theta}_1$=$\dot{\theta}_2$=0. With the controller mentioned the above equation for the torso, θ_3=0, θ_d=0

The control input system developed by following PD controller,

$$u = K_p(\theta_d - (\theta_3 - \theta_1)) - K_d(\dot{\theta}_3) - (\dot{\theta}_1)) \tag{13}$$

Where θ_d is desired angle, K_p and K_d are proportional and derivative gain. u is the torque applied between the torso and stance leg. This controller behaves as like damped spring which is acting against the stance leg. Fig. 3 shows the control system of the underactuated biped robot where using only one actuator.

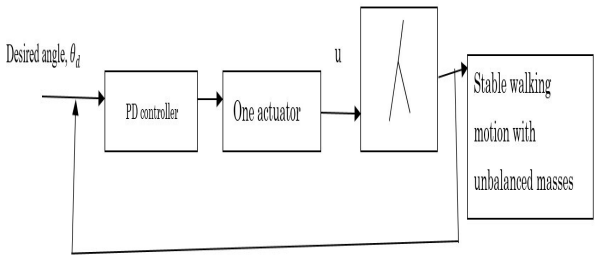


Fig. 3. Control system diagram for the biped robot

IV. Simulation Results

A. Typical walking gait of biped robot with unbalanced masses

The underactuated biped robot with unbalanced can be walking started from the initial condition

$$q(0) = \begin{bmatrix} 0 \\ 0 \\ -\alpha/2 \\ \alpha/2 \\ 5\pi/180 \end{bmatrix}, \dot{q}(0) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Table 1: Parameter Settings

a	0.5 [m]
b	0.5 [m]
L =	1.0 [m]
L_t	0.8 [m],
α	$30\pi/180$ [rad]
g =	9.81 $[m/s^2]$
m_1	5.0 [kg]
m_2	10.0 [kg]
m_h	15.0 [kg]
m_t	3.0 [kg]
K_p	490
K_d	20

Fig. 4 shows the plot of walking as a sequence of stick diagram of the biped robot when the robot walks on level ground. In this diagram, displacement 1 step each leg cycle has been shown. The red color identifies the right leg and blue color line indicated the left leg. The physical parameters showed Table 1 chosen for the underactuated robot with unbalanced masses walking. Where one leg carry double mass then other leg and hip mass also heavy m_h = 15.0 [kg]. In this case, the robot can stably walk on the level ground without fall down. The robot changes the movement each 2-period. α is desired inter leg angle at impact, the acceleration of gravity is g, Total leg length, L. The K_p and K_d are proportional gain and derivative gain and the value of K_p = 490 and K_d = 20 are used for this robot controller. The torso is controlled Or maybe delicately, with the goal that jerk at collision is trailed by an articulated forward influence and recuperation.

These parameters used for Fig.5 to Fig.10 and there are 2-period walking generated for Angular position, velocity, ground reaction force, control output, control input. The Fig.7 and 8 show that the torque input and output of the motor intercept are the same levels, which is very significant because here used only one motor and also unbalanced masses. Fig. 10 shows that the total mechanical energy, kinetic energy, and potential energy, where we can see every walking step it needs low energy and keep it the same level. We can see the two legs carry the different masses which are almost double one leg to another and hip joint carry heavy mass, torso mass m_t = 3.0 [kg] and generate 2-period steady gait on the level ground. For the period-2 gaits describe the progress of collision time has converged after two points. In other description, after some iterations, the state variables replicate

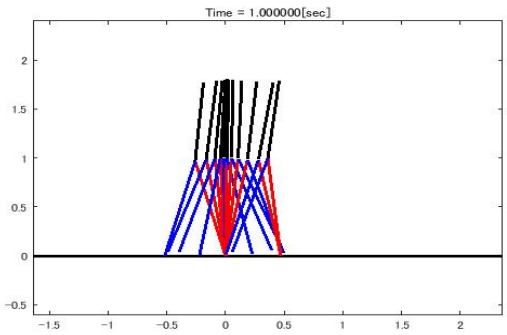


Fig. 4. The plot of walking as a sequence of stick diagram of the biped robot

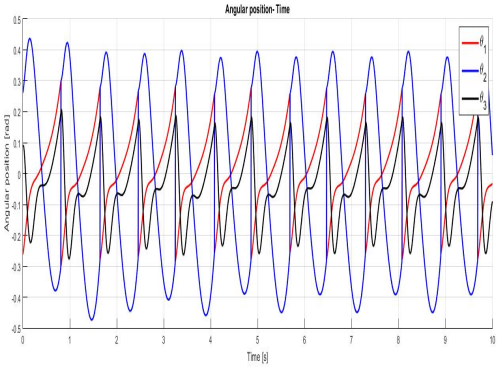


Fig. 5. Angular position for period-2

themselves after two continuous heel-strikes on the ground.

B. Gait analysis

Fig. 11 shows that the 2-period stable gait for three desired angles where used two negative and one positive values. The negative values step period is almost the same and for positive value after some disorder then step period walk 50 steps. Looking at Fig.12, it can see that when positive value used

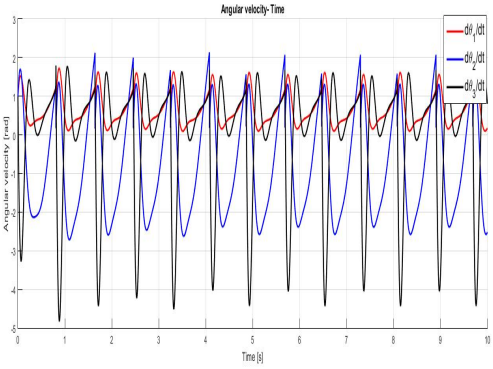


Fig. 6. Angular velocity for period-2

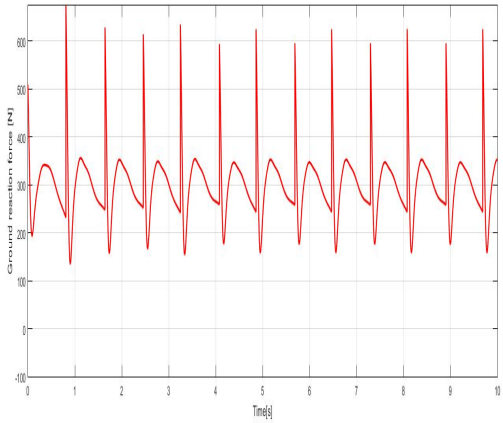


Fig. 7. Ground reaction force versus time for period-2

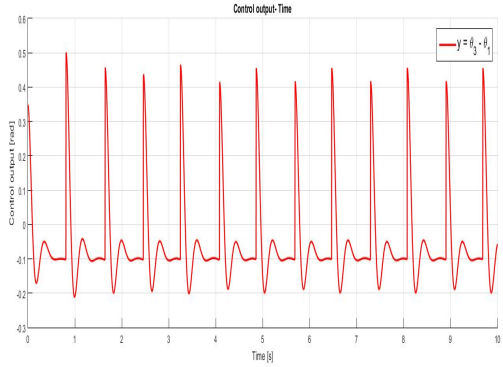


Fig. 8. The plot of output use to define the Controller versus time for period-2

for the desired angle than the walking speech low but when two negative values used then the walking speed is high and almost similar. In this case, the robot leg carries light mass instance leg which defined prosthesis leg and heavy mass carry in swing leg which defined physiological leg. In this simulation prove that light prosthesis leg design or modeling possible even other leg carry heavy mass. In this simulation shows robot can steady walk without unstable or fall-down.

Fig. 13 shows the step length for 3 different slope angles

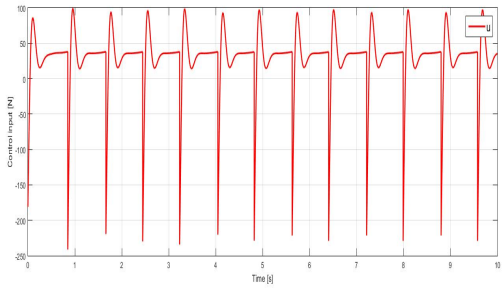


Fig. 9. The plot of applied torque input versus time for period-2

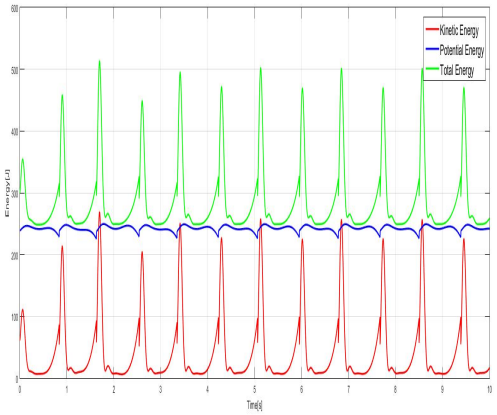


Fig. 10. Total Mechanical energy, kinetic energy and potential versus time for 2-period walking

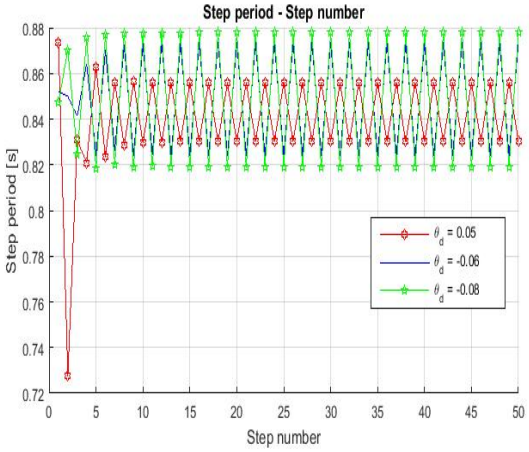


Fig. 11. Step period for different desired angle

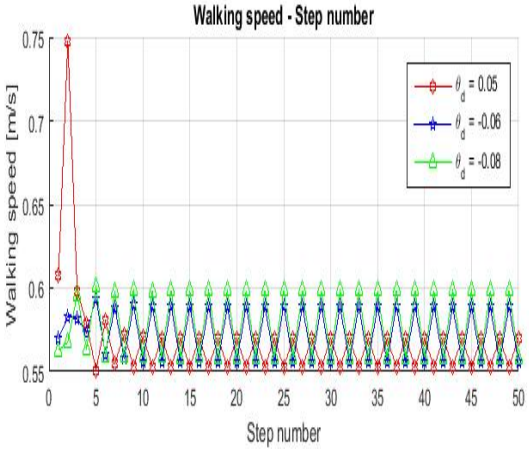


Fig. 12. Walking speed for different desired angles

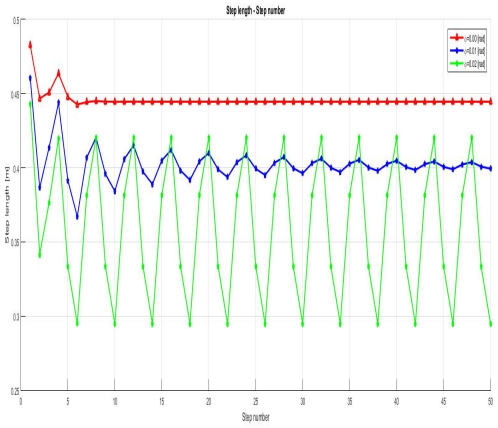


Fig. 13. Step length for different slope angles

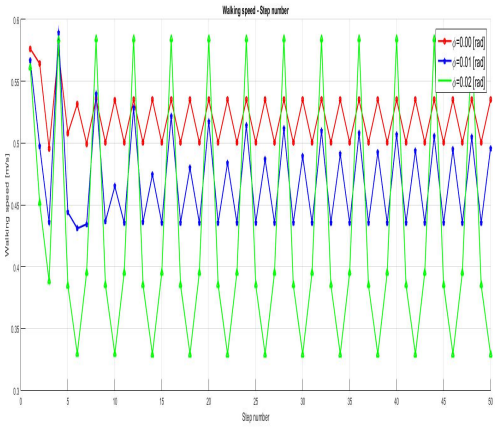


Fig. 15. Walking speed for different slope angles

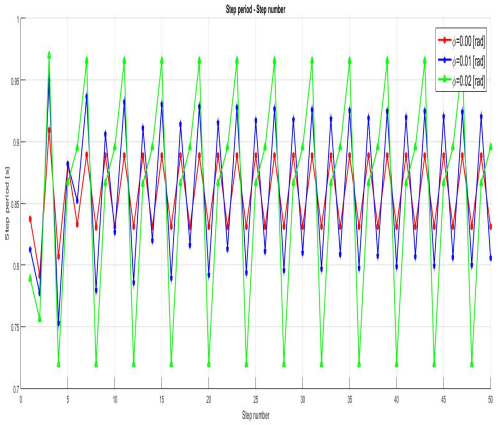


Fig. 14. Step period for different slope angles

when slope angle $\phi = 0$ [rad] then 1-period step length walking happened. Then if we used $\phi = 0.01$ and $\phi = 0.02$ the 4-period walking can generate the robot. Fig. 14 shows the step period for 3 different slope angles when slope angle $\phi = 0$ [rad] than walking step period is low. Then when we increased the slope angles $\phi = 0.01$ and $\phi = 0.02$ the 4-period walking step period can generate the robot. Fig. 15 demonstrates the walking speed of three different types of slops when slope angle $\phi = 0$ [rad] then 2-period walking cycle can generate. On the other hand, when the slope angle increased $\phi = 0.01$ and $\phi = 0.02$ then walking speed is high and 4-period walking cycle generated by the robot.

V. Conclusion

In this paper, the underactuated biped robot with unbalanced masses using one actuator was introduced. Then we have numerical simulation tested by different slope ground and desired angles. It was indicated that changing the different slope and desired angles of generated 2-period and 4-period stable gait for walking speed, step length, step period. The numerical simulations showed stable gait generation carrying with light and heavy masses is possible for the robot. The upper body robot walks are able to give accurate and simple behavior of the biped system. This model also a better understanding of human movement analysis who have central nervous system damage such as injury, stroke.

VI. Future Work

In the future, we should investigate more analysis of underactuated with unbalanced masses and develop the more advanced control system for biped robot upper body model with unbalanced masses. We will use the two actuators and control system for the upper body robot with unbalanced masses. Also, consider the different type of mathematical and simulation model the robot such as adding knees, semicircular feet and study different desired angles, downhill gentle slope and leg masses ratio [9].

References

[1] A. D Kuo. “Choosing your steps carefully”, IEEE Robotics & Automation Magazine., Vol. 14, pp. 18-29, June 2007.

[2] T. McGeer, “Passive dynamic walking,” Int. J. Robot. Res., vol. 9, no.2, pp. 62–82, Apr. 1990.

[3] Terumasa N, Masaki T, Kazuo Y. “Numerical simulations of level ground walking based on passive walk for planar biped robots with torso by hip actuators”. J Syst Des Dyn, 2008, 2: 463–474.

[4] F.Asano, L.Zhi-Wei,”On energy efficient and high-speed dynamic biped locomotion with semicircular feet,in :proceeding of 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006, pp.5901-5906.

[5] M. W. Gomes and A. Ruina, “A walking model with no energy”.

[6] Hobbelen D G E. “Limit cycle walking”. Dissertation for the Doctoral Degree. Delft: Delft University of Technology, 2008.

[7] Xuan Xiao, F. Asano ”Limit Cycle Walking of Underactuated Bipedal Humanoid on Slippery Road Surface”. 2014 14th IEEE-RAS International Conference on Humanoid Robots (Humanoids) November 18-20, 2014. Madrid, Spain.

[8] Harata, Y., et al., Efficiency analysis of two-period asymmetric gaits. International Journal of Dynamics and Control, 2013: p. 1-10.

[9] A. TEHRANI SAFA, M. NARAGHI, ”The role of walking surface in enhancing the stability of the simplest passive dynamic biped”, ROBOTICA, Robotica 33.1, 2015, 195-207.

Artificial Intelligence Humanoid BONGO Robot in Bangladesh

A S M Ahsanul Sarkar Akib
Computer Science and Engineering & Technical Director
Bangladesh University of Business and Technology (BUBT) and Japan-Bangladesh Robotics and Advanced Technology (JBRATRC)
Dhaka, Bangladesh
ahsanulakib@gmail.com

Mohammad Farhan Ferdous
Founder and Chairman
JBRATRC
Fukui, Japan
jbratrc2018@gmail.com

Milon Biswas
Assistant professor
BUBT
Dhaka, Bangladesh
milonbiswas4702@gmail.com

Hossain Mohammad Khondokar
Computer Science and Engineering
BUBT
Dhaka, Bangladesh
hossainkhondokar78@gmail.com

***Abstract*—This paper presents artificial intelligence BONGO humanoid robot which is totally made with minimal cost in Bangladesh. The full meaning of BONGO is Bangla Robot. BONGO is a multilingual edifying robot arranged and impelled robot perspective in Bangladesh. This robot can appreciate Bangla and English the two lingos and prepared to talk with people and give the proper reaction to any sorts of the request. This robot sends irregular state voice and picture update advancement that why BONGO can know the front of the all inclusive community and it can talk with people by various sorts of explanation. For instance, Welcome to people and handshake with them, move the head etc. This robot can recognize an item as one of a kind individual. At that point, It will Be coordinating with our capacity organizers picture record. We use a unique mark hashing picture method to identify an image uniquely. Moreover, BONGO can be prepared to walk around shown by people directly. These days we are attempting to greater improvement of BONGO.**

***Index Terms*—BONGO; Sophia; humanoid robot; Bangla Robot, ASIMO.**

I. Introduction

The humanoid robot works as like Sofia [1] is particularly difficult however trying for building up the nation in Bangladesh. The confinement of monetary help and not accessible all materials for making a robot. The accomplished outcome is especially confused, yet we are endeavoring to research and make the BONGO robot as like human conduct. The BONGO robot media announced by the voice of America [2]. This robot likewise shapes like as Honda humanoid robot ASIMO [7].

Humanoid robot arrangement isn't equivalent to a customary present day diagram. A humanoid robot needs personality and character since it buries acts earnestly with individuals. Moreover, the layout procedure should fluctuate from that of machines, for example, vehicles and planes since it moves self-rulingly by seeing its condition. As the enhancement of human-PC participation additionally, interface design attributable to the advancement of the PC business, humanoid robot arrangement should be considered extensively and be recognized as another arrangement asks about the zone. The Sophia is a social humanoid robot made by Hong Kong-based association Hanson Robotics. Sophia was established on April 19, 2015 [1] and appeared at South by Southwest Festival (SXSW) in mid-March 2016 in Austin, Texas, United States [3]. She can appear more than 62 outward appearances. Sophia has been anchored by media around the globe and has shared in some conspicuous gatherings. While examiners around the world have been roused by the refinement of an expansive number of Sophia's responses to their request, the fundamental piece of Sophia's critical enunciations is acknowledged by masters to be somewhat scripted. KAIST configuration, control and built up their humanoid robot name KHR-2 [4]. The BONGO robot has some potential utilization of the distinctive sort of place, for example, security robot, instructive robot, amusements robots, Intelligent robot, military robot, human correspondence robot and so forth. The BONGO recently achieved the champion award of innovation challenge category on an international competition Ureckon 19 uem, kolkata [5].

II. Architectures of BONGO

Human-like shape and improvement, lightweight, moderate size, and kickback free actuator, independent structure, low power use As shown by above arrangement thought, we sketched out a youth assessed robot which resembles a human shape and has satisfactory joints to copy a human in Fig. 1. The BONGO robot front view and side view photos appear in Fig. 2. We used the front ultrasonic sensor for perceiving impediment in front of the robot. For autonomous development, all controllers, tangible gadgets, engines, and batteries, etc are set up inside BONGO. The BONGO robot full body and hand developed of Aluminum. BONGO has

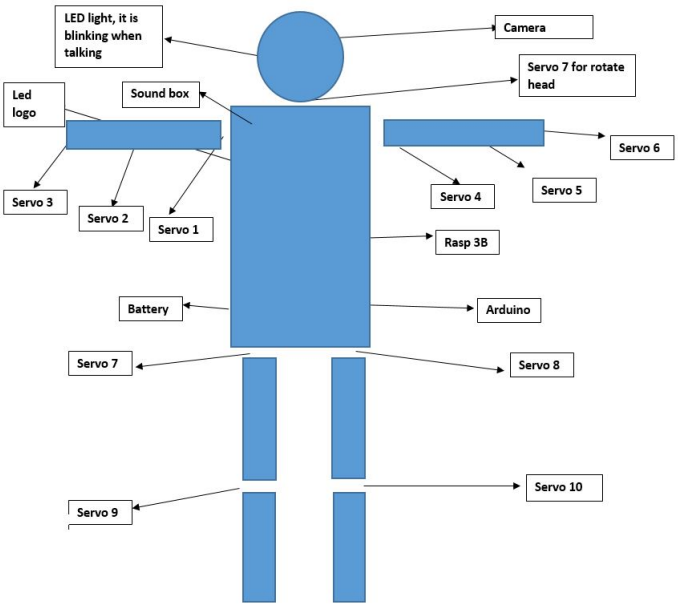


Fig. 1. Model of the BONGO robot

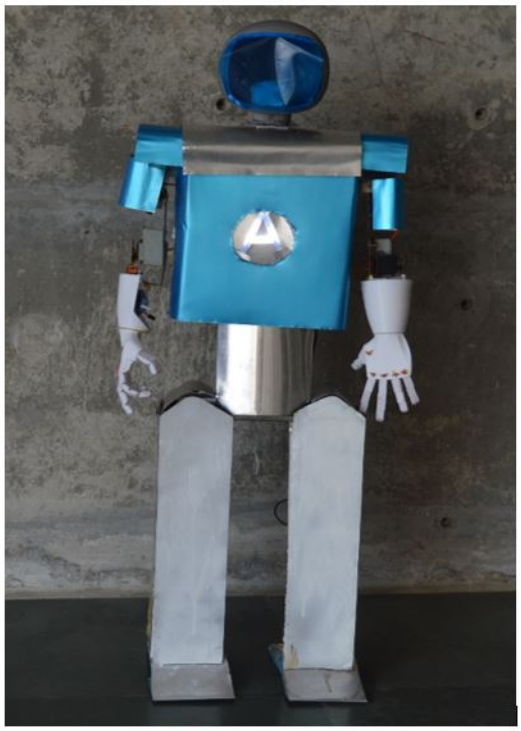


Fig. 2. Humanoid BONGO robot

Height and weight is 4.6 fit and 20 kg separately. We make the leader of the robot by a 3D printer. The body parts made by an aluminum sheet. We utilized the legs for the development of the robot. Later on, we will utilize legs for strolling the robot.



Fig. 3. MicroController



Fig. 4. Servo Motor



Fig. 5. Camera

III. METHODOLOGY

A. Hardware components

1) Microcontroller: The microcontroller shows in Fig.3. It controls any fringe device. like servo motor, DC motor, Ultrasonic sound Sensors and so on. We control here 18 Servo motors, 2 Ultrasonic sound Sensors, and some LED.

2) Servo motor: Servo motor to a blunder detecting input control which is utilized to remedy the execution of a framework. Fig. 4 shows the picture of the servo motor. Servo or RC Servo Motors are DC machine furnished with a servo system for exact control of the precise position. The RC servo engines, as a rule, have a revolution confine from 90° to 180°. A few servos additionally have turn farthest point of at least 360°. In any case, servos don't turn constantly. Their pivot is limited in the middle of the settled points. We are utilizing servo Motor here for control robot hands.

3) Camera: The camera shows in Fig. 5 which is a picture sensor gadget. So we utilized the camera to take face or protest picture. This photo perceives or process by the microprocessor.

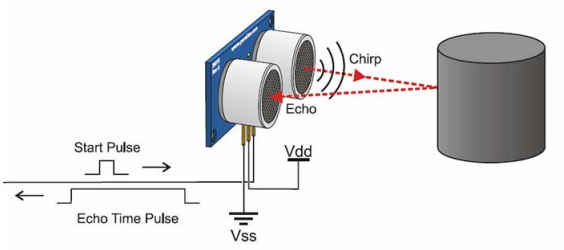


Fig. 6. Ultrasonic Sensor



Fig. 7. Raspberry Pi 3B

4) Ultrasonic sensor: Ultrasonic sensors show in Fig.6 which is a high-recurrence sound sensor. At the point when any protest results in these present circumstances Sensor. It detecting by the question and it sends information to the microcontroller. The BONGO detect an obstacle in front of the sensor.

5) Raspberry Pi 3B: The Raspberry Pi 3 device is the third epoch Raspberry Pi. It dislocates the Raspberry Pi 2 design B in February 2016. Contrasted with the Raspberry Pi 2 it has A 1.2GHz 64-bit quad-center ARMv8 CPU 802.11n Wireless LAN. Fig.7 shows the Raspberry Pi 3B device. Bluetooth 4.1 and Bluetooth Low Energy (BLE). We interface the camera with raspberry pi for preparing the picture to recognize human and items.

6) Arduino Uno R3: Arduino Uno device is a microcontroller board serving on the ATmega328p. We utilize this microcontroller to control servo and drove light. We have associated the Bluetooth module with it. Fig.8 shows Raspberry the Raspberry Pi 3B.

The Fig.9 shows hard ware block diagram of BONGO. The camera and Raspberry Pi 3B input connected with Andrino device and sound box, microphone is output device which used for robot communicate with people. The Arduino device input and output connected to the ultrasonic sensors and output with LEDs and Servo motors.

B. *Software tasks*

We used the Android operating System on android studio with the programming language Java. There are two main parts used for the BONGO robot. One part is Artificial Intelligence and another control of Microcontroller. The Artificial Intelligence part describes voice recognition. The robot also

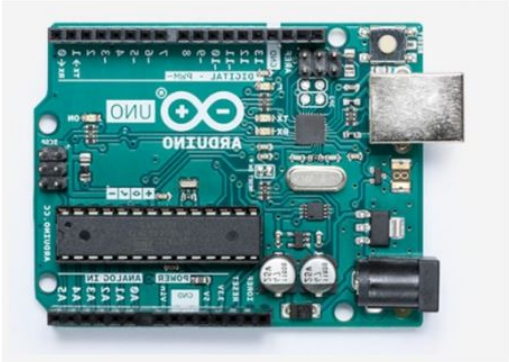


Fig. 8. Arduino Uno R3

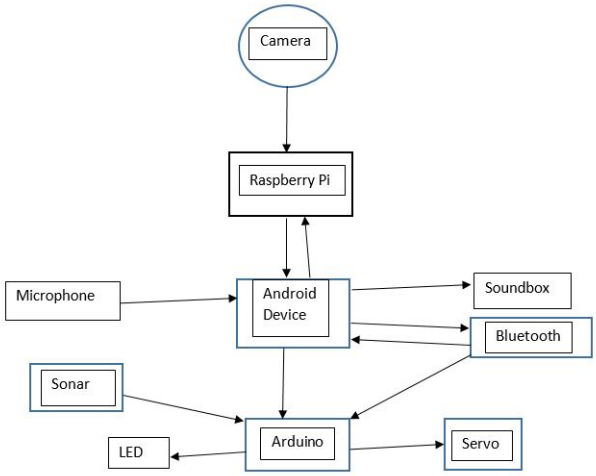


Fig. 9. Hardware block diagram

able to answer any types of questions. Microcontroller control movement of the robot, the handshake of the people, move head, hands and it can control by the microcontroller.

Right off the bat, we gather discourse question content, Then utilizing point, we deliver a watchword based scanning for the ideal answer, on the off chance that not discovered, it hunting on Google down an answer. For instance, We set an inquiry 'what is your name' and answer is 'my name is BONGO' .at that point we ask 'your name' .it produce key your and name and it scans our application framework for answer and afterward it discovered its watchword with our inquiry and return answer content and the content go to content to discourse framework. Tragically, when asking content watchword don't coordinate, at that point return 0 and when it discovered 0 it looks on Google for reply .even it not discovered, at that point, it stores learning framework, for learn for its answer. There are described some important parts of the control of the BONGO robot.

- Artificial Intelligence part code

```
public class Main {
public Main() {}
public static void main(String[] args)
{MagicStrings.root\_path
= System.getProperty("user.dir");
System.out.println("Working Directory
= " + MagicStrings.root\_path);
AIMLProcessor.extension
 = new PCAIMLProcessorExtension();
mainFunction(args);}
public static void mainFunction
(String[] args) {
String botName = "super";
String action = "chat";
System.out.println
(MagicStrings.programNameVersion);
String[] timer = args;
int bot = args.length;
for(int i$ = 0; i$ < bot; ++i$) {
String s = timer[i$];
System.out.println(s);
String[] splitArg =
s.split("=");    if(splitArg.length >= 2)
{String option = splitArg[0];
String value = splitArg[1];
if(option.equals("bot")) {
botName = value;
                                }
.....................
.....................
```

- Describe Artificial Intelligence part

1) Our speech AI part contains below part- speech collects, train data, predict the requested data with trained data, predict the high accuracy present answer.

2) First, we collect human speech text using Google API. Then we compare with this text within our train data using neural technic. If this is founded with high accuracy, then we collect this answer. Otherwise, we try to collect within Google and Wikipedia using summarizing question keyword.

3) Our image processing part-collect image, detect as unique human or animal, store in our system file storage.

4) First, when a human stand in front of this robot. if a human trying to question him or introduce to him, the robot takes an image of this human and store this picture. using OpenCV and analysis this picture with our previous train picture or data, if it corresponds with our train data. identify this human. Else it recognizes as an unknown person. Collect these personal data for future analysis.

Example: A human face in front of him, and tell it, do you know me, it takes pictures to frame in a different angle, then analysis it for identify this person, answer according to train data information, otherwise answer is random, no, you are an unknown person, etc.

- Some examples of robot command part code

```
if(data.get(0).toString() != " ")
{String command = data.get(0).toString();
String response =
chat.multisentenceRespond(command);
muteAudio(false);
String res = new EngSearchAndResult().
ProcessResult(data.get(0).toString());
MainActivity.t2.setText
(data.get(0).toString());
if(command.indexOf("handshake") != -1 ||
command.indexOf("hand shake") != -1||
command.indexOf("hand shake with me")
!= -1||
command.indexOf("can you hand shake")
!= -1){
bt.sendMessage("A");}
else if(command.indexOf("head") != -1 ||
command.indexOf("Head") != -1){
if(command.indexOf("turn") != -1){
if(command.indexOf("right") != -1) {
bt.sendMessage("C");}
if(command.indexOf("left") != -1) {
bt.sendMessage("D");
}       }}
```

```
else{
if (command.indexOf("tell") != -1) {
if (command.indexOf("about") != -1) {
if (command.indexOf("yourself") != -1 ||
command.indexOf("your self") != -1) {
bt.sendMessage("B");
}}}}
.....................
.....................
```

We can see from Fig. 10 speech recognition flow chart. The robot can take the speech of the people from the microphone then process by the microprocessor. The people speech will match by specific programming command if the specific speech match then the robot can do the work, otherwise, the robot can tell "sorry, please repeat again".

IV. RESULTS

The BONGO robot can execute the following work :

- The BONGO can understand both Bengali and English of any questions and given the answer.
- The BONGO can response when it gets instructed such as a handshake, catch object by hand, move fingers by the voice command.

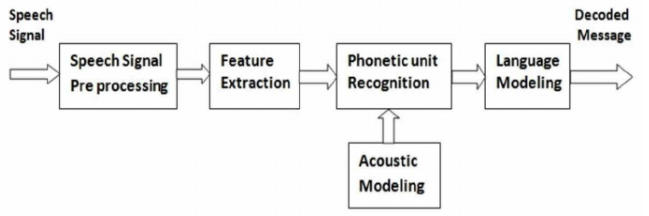


Fig. 10. Speech recognition

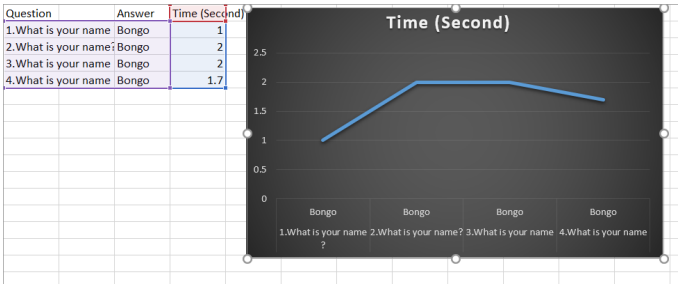


Fig. 11. The experiment of how much time the robot needs to answer the question

- BONGO can interact or communicate with a person as like Sophia robot.
- BONGO can move the head by direction of human command.
- BONGO can recognize the human voice.
- BONGO can calculate general mathematics and answer it.
- BONGO can answer any question searching from google. So the robot can answer everything that sometimes people does not know it answers.

Fig. 11 shows a sample question experimental result where we observed that there are different time need for same question answer of the robot. This answer response time generally depends on internet speed and how much noise of the environment.

Now, We have done this robotics camera experimental shot when collecting the object detection picture as a unique person. Then this picture can be matching with our storage folders image file successfully. We use fingerprint hashing image technic to detect an image uniquely. Fig. 12 shows the Image recognition experiment of two persons, where it can indicate a specific id of each person.

V. CONCLUSION

In this paper, we demonstrate that the utilization of the BONGO in reality. The structure and development of the BONGO utilize the exceptionally incorporated inserted framework and Artificial Intelligence. This robot accomplished some sort of new highlights which not present the Sophia till now. It can likewise compute general arithmetic and reply. This robot can move all heading as like left, right, in reverse and forward by the guidance of human voice. This robot camera detects to collect the image face reorganization of persons or animals.

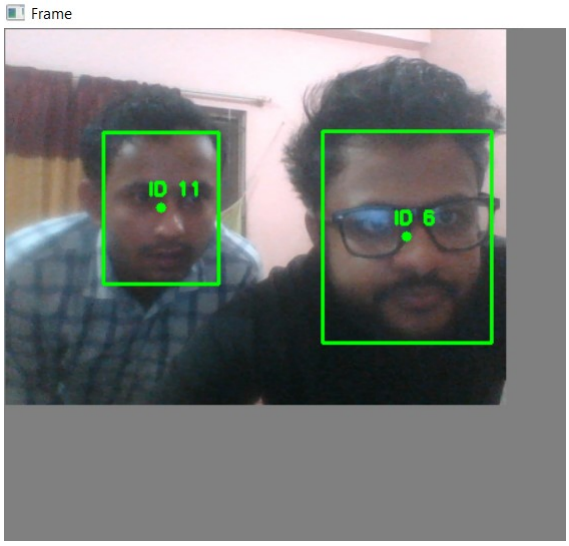


Fig. 12. Image recognition experiment

The power utilization of the robot up to 5 hours keeps on providing power by DC 12 v Battery.

VI. FUTURE WORK

We will attempt to headway BONGO robot as like progressively human lead and agreeable with people. The looking great structure will make for the robot. There need to consolidate the legs for passing on a dissent beginning with one place then onto the following and moreover effortlessly walk. We are additionally are attempting the robot will ready to keep away from the snags and easily stroll by 3D picture preparing [6]. The robot likewise causes individuals to discover an area and provide the guidance of wherever. It can likewise utilize his hand and do some basic work. We will contain the online structure by then trust that the robot will be prepared to answer the unpredictable request. We will in like manner endeavor to enhance the BONGO for therapeutic, incitement, excitements, military explanation behind different sort's business application.

ACKNOWLEDGMENT

We are particularly appreciative to go with people who support us for fiscally makes the BONGO robot. Bangladesh University of Business and Technology and Japan-Bangladesh Robotics and Advanced Technology (JBRATRC) jointly financially helps for making the robot.

REFERENCES

[1] Taylor, Harriet (2016-03-16). "Could you fall in love with robot Sophia?". CNBC. Retrieved 2018-01-04.
[2] https://www.facebook.com/voabangla/videos/1851143538322877/
[3] "Meet Sophia, the female humanoid robot and newest SXSW celebrity". PCWorld. Retrieved 2018-01-04.
[4] Jung-Yup Kim, Ill-Woo Park and Jun-Ho Oh, "Design and Walking Control of the Humanoid Robot, KHR-2(KAIST Humanoid Robot - 2)" ICCAS2004 August 25-27, The Shangri-La Hotel, Bangkok, THAILAND.
[5] http://ureckon.org.

[6] https://www.youtube.com/watch?v=AO4In7d6X-c
[7] ”Honda humanoid robots development”, Phil. Trans. R. Soc. A (2007) 365, 11–19 doi:10.1098/rsta.2006.1917 Published online 17 November 2006.

Design & Implementation of Voice Command Based Bipedal Surveillance Robot

Himel Das
Department of Electrical and Electronic Engineering
University of Science and Technology Chittagong
Chittagong, Bangladesh
himeldas077@gmail.com

Himel Chakraborty
Department of Electrical and Electronic Engineering
University of Science and Technology Chittagong
Chittagong, Bangladesh
chimel709@gmail.com

Md. Sarwar Uddin Chowdhury
Department of Electrical and Electronic Engineering
University of Science and Technology Chittagong
Chittagong, Bangladesh
sarwarcuet@gmail.com

***Abstract*— Surveillance of international border areas, industries, banks, shopping malls, and any other suspected areas is a difficult task to be done. It is not possible to scrutinize everything at each and every moment. In this case, a requisite solution can be a robot which automatically detects any suspicious activity or condition and reports to the control unit. This paper is on a bipedal surveillance robot which can budge in an area, send images and videos even in caliginous condition by using night vision mode of the Wi-Fi camera as well as the installed sensor on this robot provides gaseous parameters of the area such as the amount of Methane, LPG, Carbon mono-Oxide. The robot has a wireless Bluetooth module for communication and collection of data. The whole process can be run on a voice command system. After the data collected from the robot, data are appraised and analyzed to get an understandable idea of the surveilled area.**

Keywords— Surveillance Robot; Bipedal Robot; Gas Sensor; Night-Vision camera; Voice Command; Wireless Communication.

I. INTRODUCTION

From the dawn of civilization, there were always problematic activities that imbalance the whole socio-economic stability of a country. Terrorist attacks or Trespassing of the border is one kind of these problems. In many times, it is seen that the surveillance of these situations get very condemnatory as law-enforcers cannot get to the place at the right time. Even if they can reach, they may face a deadly counter-attack there. Some such kind of incidents occurred in Dhaka when some militants took over a café named as Holey Artisan café [1-3] or Rohingya trespassing from Myanmar into Bangladeshi border [4-6]. The multifaceted employability of robot has brought the chance to send them in a place where human life can be at stake. However, with the modernization of robotic technology, the appeal towards the human look-alike robot is increased. Bipedal Robot with the advantages like more energy efficiency [7-8], better handling capacity of the rough terrains, more maneuverability and better Human-Robot Interaction (HRI) [9], has emerged as one of the interesting research sectors in the robotics field. A chemical attack is another recent problems arising across the world [10-11]. Sometimes it is been treated more dangerous than nuclear weapons. So, it is now a crying need to sense the presence of the toxic gases in a place. Any chemical attack can be breached by avoiding the place by sensing the levels of toxic gases like Carbon-Di-Oxide (CO2), Carbon mono-Oxide (CO) and Methane (CH4) of the specific place with the use of the gas sensors. If the still pictures or the motion pictures can be found through any system then it would be more planned to counter these havocs. But the surveillance could have been done with perfection if there were any availability of surveillance robot system. With all these, there is another feature that can be developed and that is the Voice Control [12] which makes the whole system more efficient and encrypted as it gives the system a massive advantage which is to control it from a significant and safe distance. In our project, we will try to seek out these problems with the help of a Bipedal Robot which can go to a place without any human intercession where the gas sensor and night vision Wi-Fi camera are installed on it to get the data about the gaseous condition, images and videos even if in the obsessed situation. Bluetooth module is used to send the Voice Command and to transmit the data. The project also illustrates the ways to develop a Bipedal Robot and also dispenses the data transmitted from the robot during its exploration.

II. LITERATURE REVIEW

Above last decades, many researchers have designed different purpose soldier or surveillance robot. Soldier or surveillance robot is a robot which does the surveillance of International border areas, war field and terrorist activities [13]. Arduino controlled war field spy robot using night vision wireless camera and Android application is very muchly used nowadays [14]. For the implementation of this robot night, vision camera is used for monitoring and for operating Bluetooth module is used. Implementation of Spy Robot for a surveillance system using internet protocol of raspberry pi where sensors are used to detect the living object and night vision camera is used to capture the pictures of the object [15]. Wireless Controlled Military combat Robot system is a work where the micro wireless camera is fitted and where the video and voice transmitted to the control room by using the Bluetooth Module [16]. The bipedal structure is another recent conception in this sector to give pace to this specific robotic research [17]. It is done to avoid the obstacles on the way to the target. Rough terrain can be topped as like as human leg behavior by this structure. Along with this, Voice command is now playing a vital role in the establishment of a surveillance robot with more perfection [18]. Bipedal Structure has the advantages over the wheel based robot as it has the lighter weight, has higher efficiency as it has the less number of the actuator and better Human-Robot Interaction (HRI). Along with this Voice, command helps the user to control the surveillance robot from a distance which can be very significant. Android applications are made to provide the voice command by using Bluetooth Module

[19]. In addition, Arduino Movi-Shield, RF Transmitter Receiver Pair 433 MHz-RXB6, ZigBee communication modules are also used to transmit the voice command to the robot. The proposed robot can capture still images and videos by using night vision camera which also makes it perfect to send images and videos even in the night time. It can walk in unprepared terrain which is inaccessible for robots with wheels or caterpillars. It has the Gas sensor which is capable of providing data of the gaseous condition of the area under surveillance. The implemented robot is using Bluetooth module with Android platform for controlling purpose as well as to transmit the voice command.

.

III. Proposed Methodology

The upliftment of the research work is divided into 3 stages. They stand as mechanical, electronic and software portions after considering the plan for making a research work on the topic named Voice Command Based Bipedal Surveillance Robot. The upliftment procedure is represented in Figure 1. The Mechanical stage is divided into the parts as pursues:

- Design of Model
- Build the prototype
- Upliftment of Model Hardware

The Electronic stage is divided as pursues:

- Design of Circuit Diagram
- Circuit Simulation
- Upliftment of Circuit

In the software stage, an algorithm is designed as per as the plan and the programming is sorted according to the algorithm.

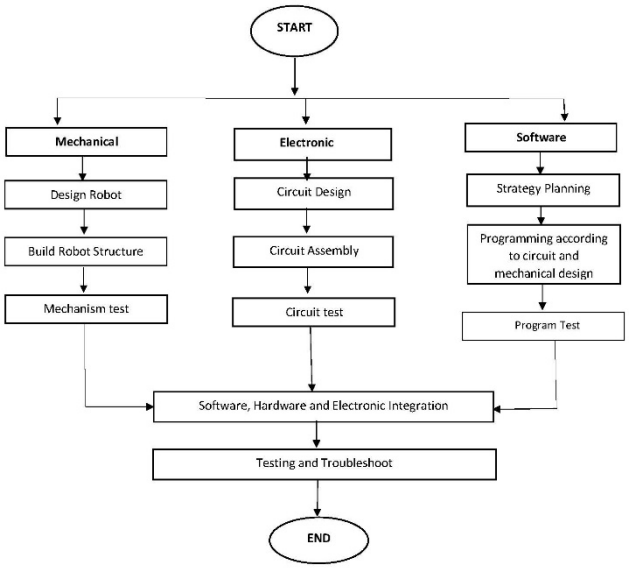


Fig. 1. Methodology of a Bipedal Surveillance Robot

IV. Proposed System Design

By taking into account its better performance over long distance data transmission from the working station the proposed robot is designed so that the user can take the idea under the surveilled area and can control the robot. Block diagram of the proposed system shown in Figure 2 & Figure 3.

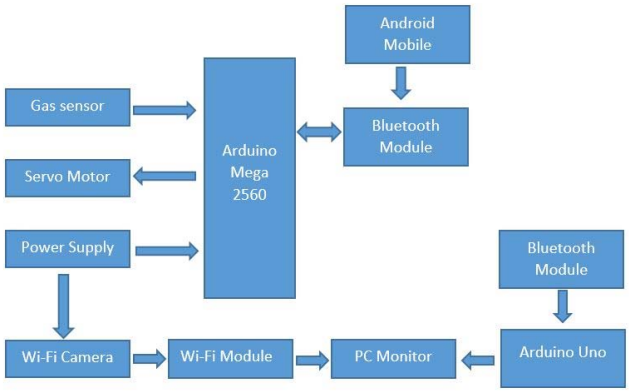


Fig. 2. Block Diagram of the Proposed System

The robot is controlled by voice command using a Bluetooth module. Voice command is transmitted by an android application. The robot collects gaseous data by a gas sensor and sends to the user with the help of the Bluetooth module as well as the robot sends the video using a smart wifi camera with night vision mode which enables it to operate even in the darkness.

In the process of building a surveillance robot, both mechanical structure and electronic components are needed

A. Mechanical Structure

In the proposed model, Acrylic plastic sheets which have 2mm thickness are used to build the structure of the model. On account of lightweight and better flexibility, this sheet is considered. In the purpose of movement, 6 metal gear servo motors are used in hip, knees, and ankles of both legs. Figure 3. shows the 3D model of the proposed system.

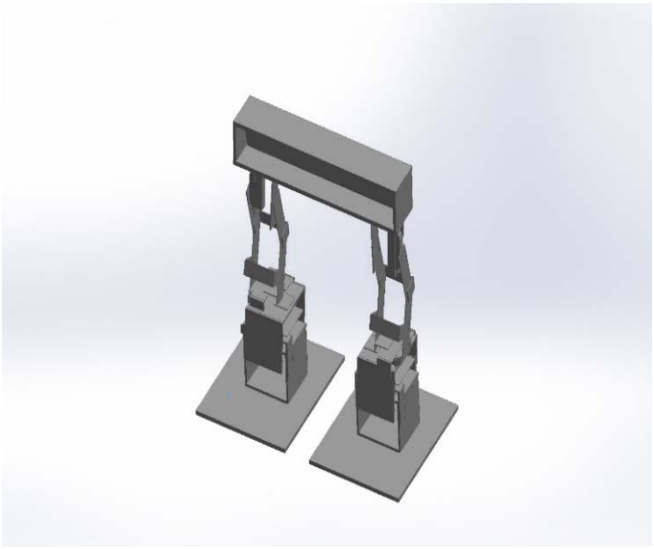


Fig. 3. 3D Model of the proposed system

B. Electronic Components

On the purpose of transmission, sensing and controlling motive, there are some components which are used and their names are given below:

- Arduino Mega 2560 and Arduino UNO
- Gas Sensor(MQ-5)
- Servo motor
- Bluetooth Module(HC-05)
- Wi-Fi Smart Camera

C. Design of Circuit Diagram

The circuit diagram of the proposed system exposes in figure 4. PROTEUS® software is used to design and partial simulation of the circuit. Since there was no way to simulate gas sensor with the presence of real gas, the simulation of gas sensor was do/-*
ne by imitative binary input to the Arduino. At the time of implementation, the circuit diagram had been modified which was designed in the simulation software. After the completion of the design, simulation and implementation of the circuit in the robot, the expected result has come out.

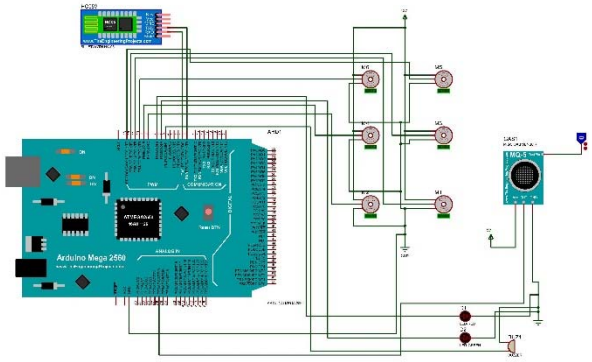


Fig. 4. Circuit Diagram of the Proposed System

D. Gas Detection

Toxic gases in the air are addressed in the microcontroller and the amplitudes of their concentration were measured by using gas sensors. To measure the specific gas concentration, the sensor indicates the result as the reference point or scale in a Concentration versus Time curve. When any toxic gas is found in the air, an alarm is rung by using a buzzer. There are two types of gas sensor have found. One is an electrochemical gas sensor which is used for detecting LPG, CO_2, CO, CH_4, etc. and another one is metal oxide semiconductor (MOS) which is used for inflammable gases.

V. Hardware Implementation

Figure 5 shows the final implementation of the robot after the development and combination of mechanical construction and circuit diagram. Hither Arduino Mega 2560 board and Bluetooth module are used for controlling the robot and transmitting and receiving data respectively. The gas sensor is used for detecting the toxic gases as like as LPG, CO, CH_4. In this project, the most complicated part is the movement of the robot accurately. Stabilization is done by using proper materials and following correct procedures of building it.

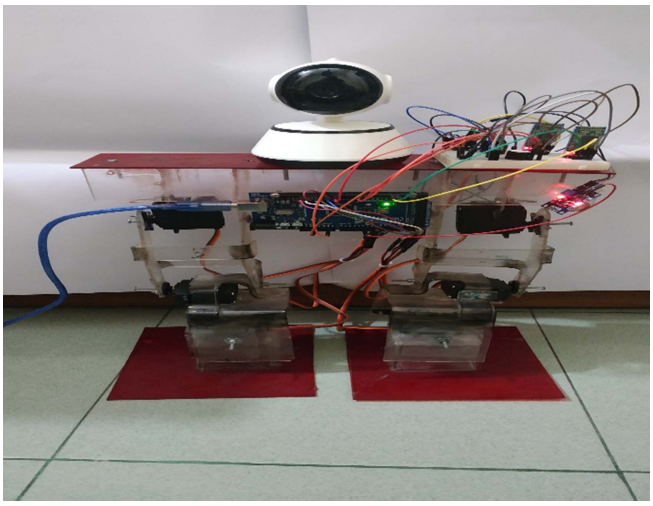


Fig. 5. Implemented prototype of the Robot

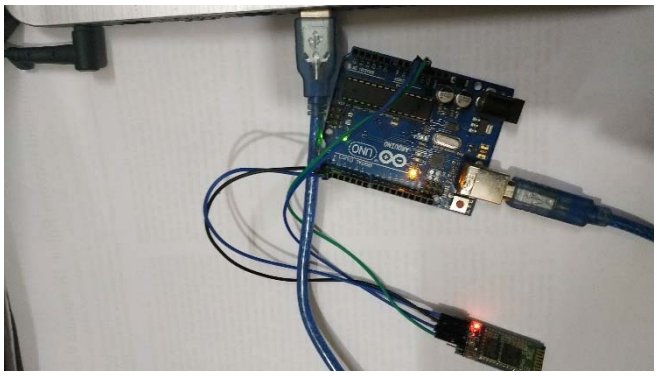


Fig. 6. Implemented Gas sensor Data Collecting station

Figure 6 shows the gas sensor's data collecting station. Here Arduino UNO and Bluetooth module are used to receive the data of the gas sensor.

VI. Software Description

Various types of software have been used for mechanical design, circuit design, real-time data analysis which are:

- Arduino: Arduino is an open source electronics platform by which many types of sensors, motors, lights, etc. can be controlled. Two types of Arduino Board have been used for the robot.
- Proteus: For Circuit Design and Simulation, it is a famous software. Since the software is easy to design and simulation, it's used for the simulation purpose.
- Blender: For 3D character design, Blender is a compatible software. Using this software, the 3D design of the robot is drawn.
- MATLAB: It is a multi-paradigm numerical computing environment and fourth-generation programming language. It is used in this project for the purpose of plotting of the collected data from the receiving module of the gas sensor.

- Android Application: An Android application is a smartphone software which runs in the Android OS platform. In this project, an android application is used to transmit the voice command with the help of Bluetooth connectivity. Figure 7 shows the android application which is used to transmit the voice command using google voice command.

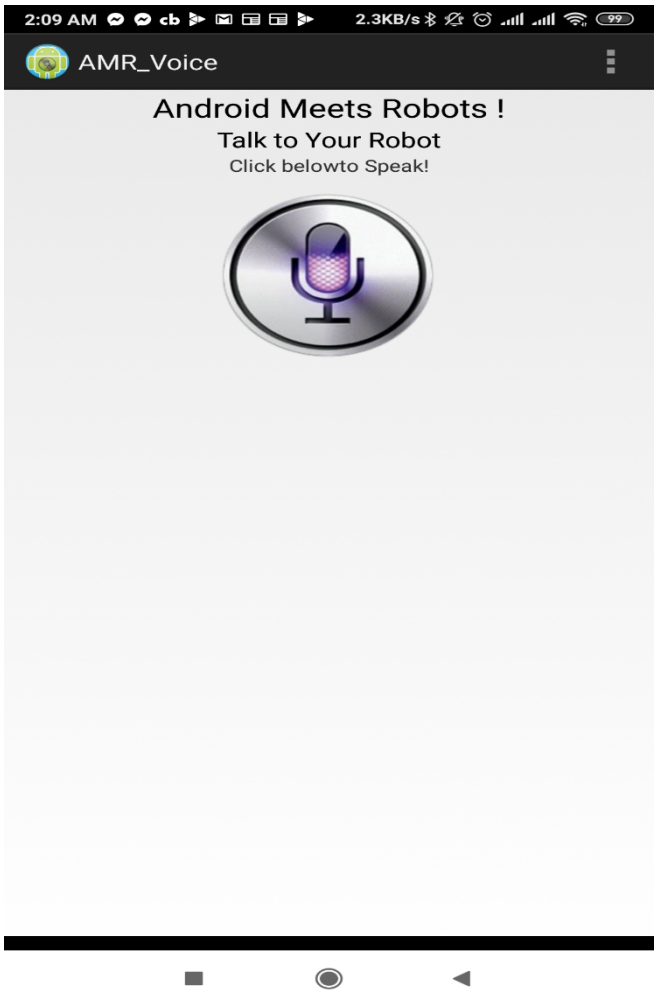


Fig. 7. Android application

VII. Experimental Result

The main goal of this project was to develop a robot which can be sent to a place where surveillance is needed and to know about the condition of this place by video streaming continuously and sensing the condition of toxic gases in the air of this place by it. It was a must to give it a capability to send data to the user by which the state of the place can be both analyzed and realized. It was not possible to conduct this environment in any real field area. So, we created an artificial environment to test this robot. The video streaming is shown in figure 8. It can send data in both bright and dark condition using night vision mode which makes the robot capable of operating in 24 hours of a day regardless of the light condition.

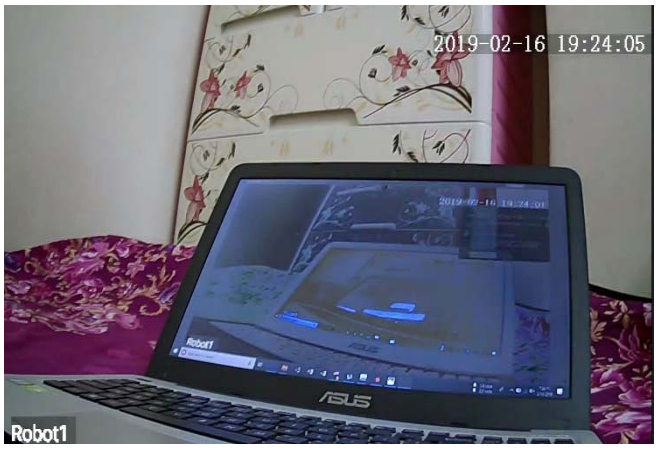


Fig. 8. Video streaming of the surveillance place

The Collected data from the gas sensor illustrated here:

a) Methane Gas (CH_4): The Methane gas variation is respected to time which is shown in figure 9. The amount of methane gas varied from 0-500 ppm. The result is shown in figure 8.

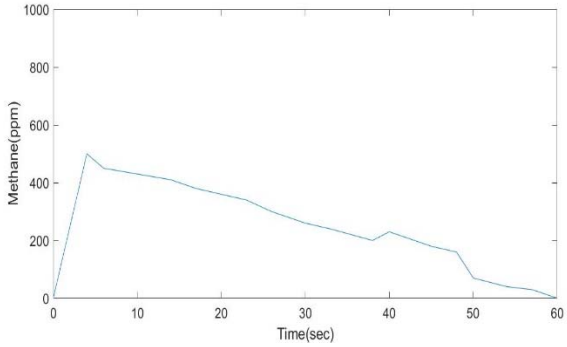


Fig. 9. Change the amount of Methane Gas over time

a) LPG: Figure 10, Shows the amount of LPG gas which was varied from 0-800 ppm. Release of the LPG gas and the decreasing rate of it can be analyzed from the plot.

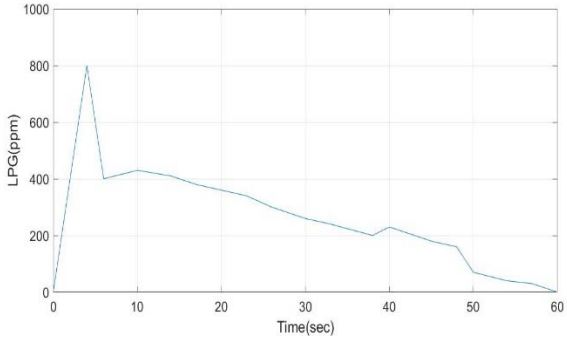


Fig. 10. Change of the amount of LPG Gas

b) Carbon mono-Oxide (CO): In figure 11, it shows the amount of CO gas over time. It varies from 0-430 ppm. By the figure, get a clear idea about the rate concentration of CO in the air.

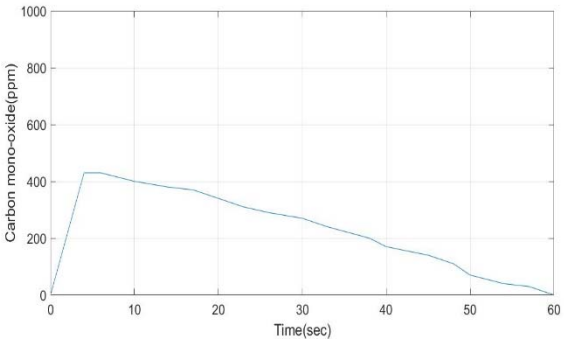


Fig. 11. Variation of the amount of CO over time

VIII. CONCLUSION

The experiment was taken in a room so the actual data is not collected. Without this, the project covers all types of design and development. It can detect LPG, CH_4, CO. through the sensor. To analyze this data get a clear idea about the environment. By the video streaming know the condition of the place. It might be a suitable one because of walking in unprepared terrain where wheel robot is faced with difficulties.

REFERENCES

[1] D. Attack and C. Correspondent, "Holey Artisan Attack: Key witness testifies", *The Daily Star*, 2019. [Online]. Available: https://www.thedailystar.net/city/dhaka-holey-artisan-bakery-attack-eyewitness-recounts-horror-1698355. [Accessed: 14- Apr- 2019].

[2] "Opinion | Under the Shadow of Terrorism in Dhaka", *Nytimes.com*, 2019.[Online].Available:https://www.nytimes.com/2017/08/09/opinion/under-the-shadow-of-terrorism-in-dhaka.html. [Accessed: 14- Apr- 2019].

[3] S. Hammadi, A. Yuhas and R. Scammell, "Dhaka cafe attack ends with 20 hostages among dead", *the Guardian*, 2019. [Online]. Available: https://www.theguardian.com/world/2016/jul/01/dhaka-bangladesh-restaurant-attack-hostages. [Accessed: 14- Apr- 2019].

[4] "Opinion | How the Rohingya Crisis Is Changing Bangladesh", *Nytimes.com*, 2019. [Online]. Available: https://www.nytimes.com/2017/10/06/opinion/rohingya-bangladesh-myanmar.html. [Accessed: 14- Apr- 2019].

[5] S. Report, "67,000 Rohingyas intruded into Bangladesh: Foreign minister", *The Daily Star*, 2019. [Online]. Available: https://www.thedailystar.net/country/67000-rohingyas-intruded-bangladesh-foreign-minister-1353877. [Accessed: 14- Apr- 2019].

[6] "Crisis update: Nearly one million Rohingya refugees are in Bangladesh now - Bangladesh", *ReliefWeb*, 2019. [Online]. Available: https://reliefweb.int/report/bangladesh/crisis-update-nearly-one-million-rohingya-refugees-are-bangladesh-now. [Accessed: 14- Apr- 2019].

[7] T. Hooi and M. Mahyuddin, "A study of walking gait stability and gait efficiency of a cost-effective small humanoid bipedal robot: Analysis, simulation and implementation", *2017 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)*, 2017. Available: 10.1109/iris.2017.8250109 [Accessed 14 April 2019].

[8] S. Han, S. Um and S. Kim, "Mechanical design of robot lower body based on four-bar linkage structure for energy efficient bipedal walking", *2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 2016. Available: 10.1109/ssrr.2016.7784334 [Accessed 14 April 2019].

[9] A. Bicchi, "Of Robots, Humans, Bodies and Intelligence: Body Languages for Human Robot Interaction - IEEE Conference Publication", *Ieeexplore.ieee.org*, 2019.

[10] "Syria 'chemical attack': What we know", *BBC News*, 2019. [Online]. Available: https://www.bbc.com/news/world-middle-east-43697084. [Accessed: 14- Apr- 2019].

[11] "U.S. Says Syria Has Used Chemical Weapons at Least 50 Times During War", *Nytimes.com*, 2019. [Online]. Available: https://www.nytimes.com/2018/04/13/world/middleeast/un-syria-haley-chemical-weapons.html. [Accessed: 14- Apr- 2019].

[12] P. Dubey, K. Pal, B. Champaty and D. Tibarewala, "Development of a wireless voice control system for rehabilitative devices", *2014 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2014]*, 2014. Available: 10.1109/iccpct.2014.7055028 [Accessed 14 April 2019].

[13] T. Sobh, R. Sanyal and B. Wang, "Remote surveillance via wireless-controlled mobile robots", *Proceedings World Automation Congress, 2004.*, 2005. [Accessed 14 April 2019].

[14] J. Park and K. Sim, "A design of mobile robot based on Network Camera and sound source localization for intelligent surveillance system", *2008 International Conference on Control, Automation and Systems*, 2008. Available: 10.1109/iccas.2008.4694586 [Accessed 14 April 2019].

[15] A. Bokade and V. Ratnaparkhe, "Video surveillance robot control using smartphone and Raspberry pi", *2016 International Conference on Communication and Signal Processing (ICCSP)*, 2016. Available: 10.1109/iccsp.2016.7754547 [Accessed 14 April 2019].

[16] S. Joshi, G. Aravalli, A. Vidyashree, S. Ranade and S. Badami, "Wireless controlled military combat robot system", *2017 2nd International Conference on Communication and Electronics Systems (ICCES)*, 2017. Available: 10.1109/cesys.2017.8321173 [Accessed 14 April 2019].

[17] A. Nath, G. Das, A. Mallick and S. Chowdhury, "Design, implementation and stabilization of a Bipedal robot", *AIP Conference Proceedings, 2017*, vol. 1919, no. 1, 2017. Available: 10.1063/1.5018541 [Accessed 14 April 2019].

[18] X. Lv, M. Zhang and H. Li, "Robot control based on voice command", *2008 IEEE International Conference on Automation and Logistics*, 2008. Available: 10.1109/ical.2008.4636587 [Accessed 14 April 2019].

[19] S. Zope, P. Muluk, R. Mohite, A. Lanke and M. Bamankar, "Voice Control Robot Using Android Application", *Imperial Journal of Interdisciplinary Research*, vol. 3, no. 2, 2017. [Accessed 14 April 2019].

Development of Myoelectric Robotic Hand using 3D Printer

Sivapong Nilwong
Graduate School of Science and Engineering
Hosei University
3-7-2 Kajinochō, Koganei,
Tokyo, Japan
sivapong.nilwong.46@stu.hosei.ac.jp

Delowar Hossain
Research and Development
Fairy Devices Inc.
Yushima Urban Bldg. 7F, Yushima 2-31-22, Bunkyo-ku,
Tokyo, Japan
hossain@fairydevices.jp

Eneo Petoku
Faculty of Information Technology
Polytechnic University of Tirana
Bulevardi "Deshmoret e Kombit", Sheshi "Nene Tereza", 4,
Tirana, Albania
petokueneo@gmail.com

Genci Capi
Department of Mechanical Engineering
Hosei University
3-7-2 Kajinochō, Koganei,
Tokyo, Japan
capi@hosei.ac.jp

***Abstract*—This paper presents a low-cost myoelectric robotic hand which was developed by a 3D printer. The purpose of the developed system is to be used as a low-cost prosthesis. It is controlled by electromyogram (EMG) signals from flexor and extensor muscles in the lower arm. Each part of the robotic hand is created using the 3D printer. Arduino MEGA microcontroller is used as the main controller. Experimental results show that the robotic hand can perform the objects grasping operations according to EMG signals from muscles without dropping the objects in the experiments.**

Keywords—robot hand, EMG signal, 3D printing, a prosthesis

I. INTRODUCTION

The limb loss is the problem that occurs on a large scale in every year. There was an estimation that 1.6 million people in the United States were living with the loss of limbs in 2005, and this number has the possibility to be doubled by 2050 [1]. Prostheses are devices which replace the lost limbs. Persons who suffer from the loss of upper limbs may require prosthetic arms or prosthetic hands to replace their missing limbs. There exists a significant number of researches on prosthetic limbs. Such research includes the classification and implementation of EMG signals from human arm on a prosthetic hand [2], the design and development of the prosthetic limb which controlled by EMG signals from muscles near the human shoulder [3], an implementation of the actor-critic reinforcement learning method for controlling prosthetic arm movements using EMG signals [4], the development of the hybrid prosthetic arm controlled by signals from the human brain [5], an application of the force sensor for sensing purpose on prosthesis [6] and the design of the prosthetic hand that can grab multiple types of objects using a single actuator [7]. Among the methods for controlling prostheses, myoelectric is one of the widely applied methods. Users of the myoelectric prostheses can use the EMG signals from their muscles to control their prostheses, which can be seen from [2], [3], [4], [8] and [9].

Despite the number of prostheses research available, the majority of the works have overlooked one problem: the cost. The cost problem can lead to less accessibility to the prosthesis. Recently, the application of 3D printing has spread throughout many fields of researches. The use of 3D printing allows the manufacturing of objects to be more simple, cheaper and more versatile. Varieties of 3D printing implementations can be found such as digital forensics [10], robot parts manufacturing [11-12], and prosthetic limbs manufacturing [13-14].

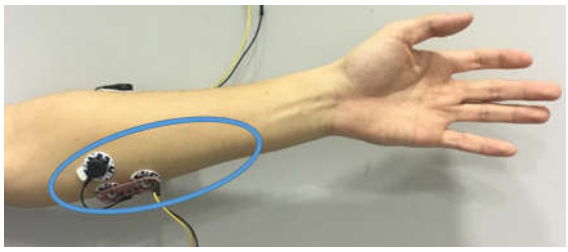
We propose the 3D printed myoelectric robot hand that controlled by EMG signals to achieve the goal of developing a controllable prosthetic hand with low production and low maintenance costs. Parts of the proposed robot hand are created from the 3D printer. For controlling the robotic hand, EMG signals are received from muscles in the human lower arm. The developed system was tested with two experiments on the grasping tasks: plastic bottle and small computer mouse. The initial tests of the robotic hand gave satisfactory results.

Remaining parts of this paper are divided into sections as follows: section 2 illustrates the EMG signals implemented to the robotic hand. Section 3 describes the components of the proposed robotic hand. Section 4 displays experiments and the experimental results of the robotic hand. Finally, section 5 concludes matters in this paper and discuss the future works on the proposed robotic hand.

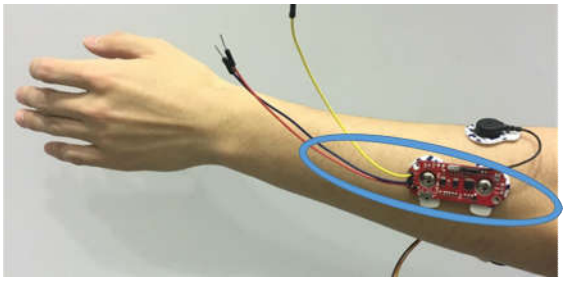
II. IMPLEMENTED EMG SIGNALS

EMG signals are emitted from the contraction of muscles. The proposed robotic hand utilizes EMG signals from two muscle groups in the lower arm: flexor muscle and extensor muscle. Muscle sensors need to be placed around the specified muscles, one sensor for each muscle. In this paper, two MyoWare muscle sensors are applied for receiving EMG signals. Fig. 1 shows the locations of muscle sensors on the human arm, where locations of sensors are marked with circles. Fig. 1(a) indicates the sensor location for the flexor muscle and Fig. 1(b) for the extensor muscle.

We tested the muscle sensors with two simple operations: hand closing and hand opening. During the tests, EMG signals were received from the human arm and displayed as graphs. Fig. 2 shows graphs of the received EMG signals, where the blue line refers to the signal from flexor muscle and the red line is from the extensor muscle. The green square in Fig. 2 marks the EMG signals at the normal level when the human arm is relaxed. The blue circle and red circle areas show the EMG signals when the human hand is closed and stretched open, respectively. From the graphs, the EMG signal from flexor muscle is significantly stronger than the normal level when the hand is closed. On the other hand,



(a)



(b)

Fig. 1 Muscle sensors locations on human arm muscles: (a) flexor muscle (b) extensor muscle

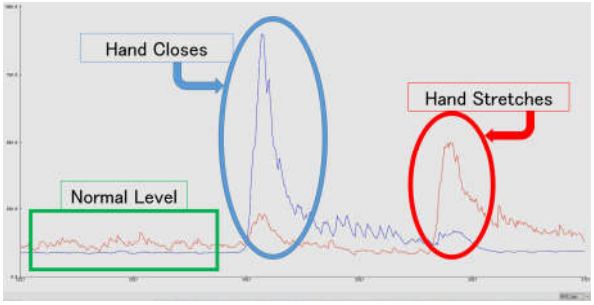


Fig. 2 EMG signals from two arm muscles (blue: flexor muscle, red: extensor muscle)

the amplitude of the signal from extensor muscle is greatly increased from the normal level when the hand is stretched.

These characteristics of EMG signals can be used for controlling the robot hand to perform simple actions. In this work, EMG signals are used as the trigger for two actions: hand closing and hand opening. The activities of the signal from flexor muscle are applied for the hand closing action, while the hand opening action is performed regarding the signal from extensor muscle.

III. Proposed Myoelectric Robotic Hand

The overview mechanism of our myoelectric robotic hand is displayed in Fig. 3. The system flow starts from the muscles in the human arm where muscle activities emit EMG signals. The EMG signals are detected by the MyoWare muscle sensors, then sent to the main controller. We used the Arduino MEGA as the main controller of the robotic hand. After receiving the signals, the main controller moves small servo motors according to the EMG signals. These small servo motors are located inside the robotic hand. Once the human user closes their hand, the robotic hand also performs the hand closing action. In contrast, the robotic

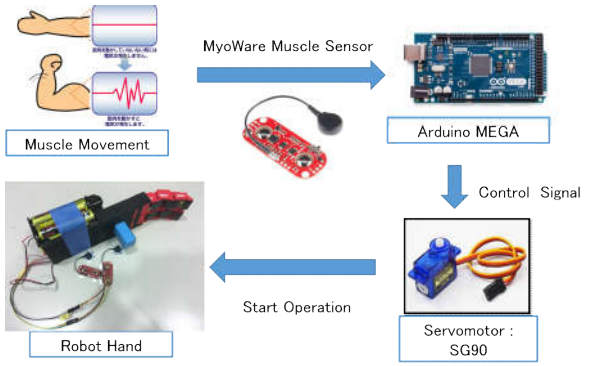


Fig. 3 Overview system flows of the robotic hand

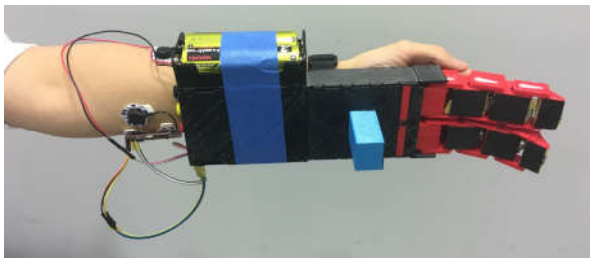


Fig. 4 The proposed robotic hand

TABLE I. Robot Hand Specifications

Attribute	Specifications
Dimensions	326.0×84.5×77.0 mm
Weight	732 g
Smallest Grasp	20×20 mm
Sensors	2 MyoWare Muscle Sensors
Motors	6 SG90 Servo Motors
Controller	Arduino MEGA

hand performs the opening operation when the user stretches open their hand. The hardware and software of our robotic hand are described in following subsections.

A. Mechanical Design and Components

We designed all parts of the robotic hand on SolidWorks 2015. The parts were printed from XYZprinting da Vinci 2.0A Duo 3D printer. The printed parts include the frame for the main controller, two movable fingers, and an immovable thumb part. Fig. 4 illustrates the appearance of the robot hand. Table 1 describes the specifications of the robot hand. The robotic hand is 326.0 mm in length, 84.5 mm in width and 77.0 mm in height, with the weight of 732g. The smallest object that the robot hand can grasp is at the size of 20×20 mm. Arduino MEGA is installed inside the frame part to work as the main controller. There is one battery socket installed along the Arduino MEGA to power the controller. Two MyoWare muscle sensors are connected to the main controller through the wired connection. Six SG90 servo motors are placed in six joints of two movable finger parts. All six servo motors are connected to the main controller through the wires. These servo motors are directly powered by four 1.5V batteries from the external battery

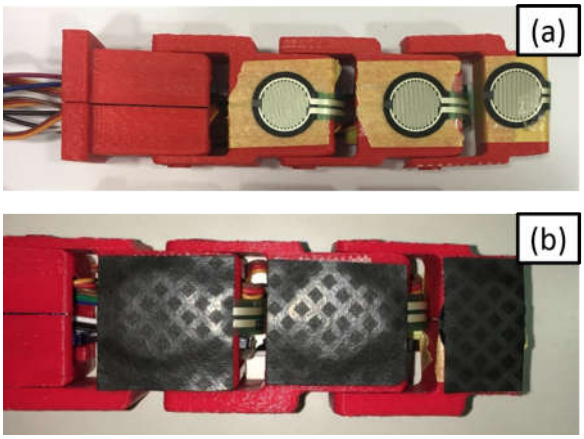


Fig. 5 Moveable finger parts: (a) with pressure sensors, (b) with anti-skid rubbers

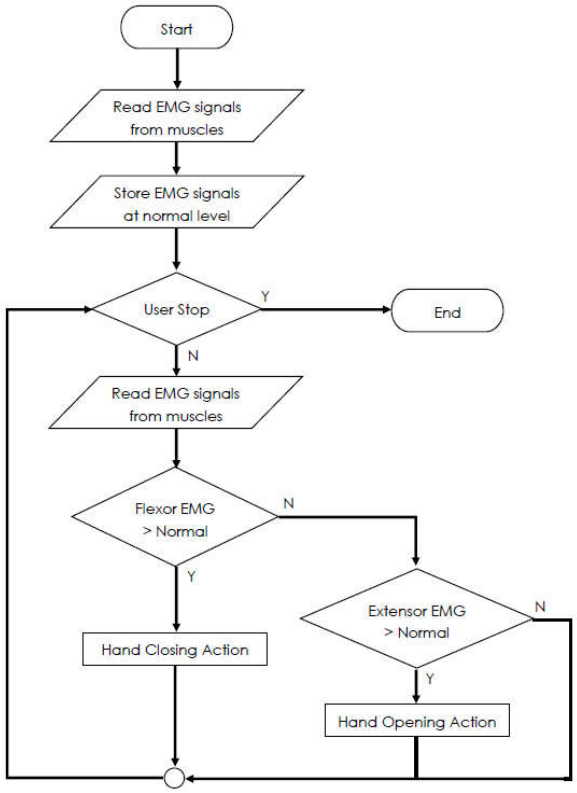


Fig. 6 Flowchart of the robotic hand software

sockets. To hold the robot hand with the human arm, the robot hand is also attached with two adhesives nylon tapes. There are three switches installed to control the hand: the power button, start button and stop button.

Fig. 5 shows two movable finger parts. There are two fingers available in the robot hand. The SG90 servo motor is attached to each joint of the finger part. The surface of each finger is attached with a set of pressure sensors shown in Fig. 5(a). These pressure sensors give the robot hand an ability to grasp an object with the proper amount of force. An anti-skid

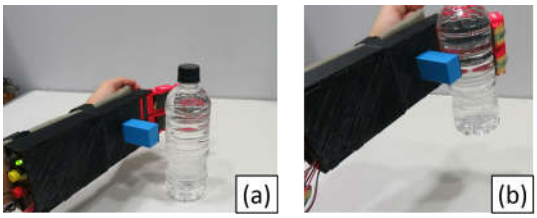


Fig. 7 Grasping of the plastic bottle: (a) before grasping, (b) during grasping

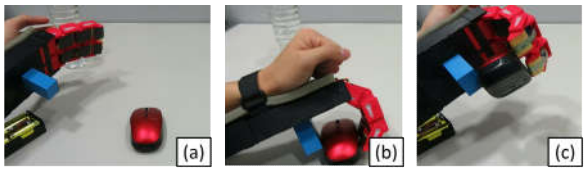


Fig. 8 Grasping of the computer mouse: (a) before grasping, (b) grasping (top), (c) grasping (bottom)

rubber sheet is placed on top of each pressure sensor to prevent the grasped object from slippering, as in Fig. 5(b)

B. Control Software

Software for controlling the proposed robotic hand is written as a single Arduino program, in which the core algorithm relies on the activities of EMG signals. Fig. 6 describes flows of the robotic hand software. The process of software part starts from the initialization, in which the EMG signals at the normal level are recorded. The EMG signals at the normal level are stored to be used as the reference. The robotic hand then progresses to the main process. In the main process, the robot hand continuously retrieves EMG signals from flexor and extensor muscles. If the signals from flexor muscle are stronger than the stored reference, the main controller triggers servo motors for the hand closing action. On the contrary, the hand opening action is triggered when the retrieved signals from extensor muscle are stronger than their reference. The process stops once the stop button on the robotic hand is pressed.

IV. Experimental Results

The robotic hand was tested with two object grasping tasks: 500 ml plastic bottle grasping and small computer mouse grasping. Experimental results are shown in Fig. 7 and Fig. 8. The result of the bottle grasping task is displayed in Fig. 7. The robot can grasp the 500 ml plastic bottle immediately after receiving the EMG signals from muscles. The bottle was not damaged during the grasping and also held in the robot hand without slipping. For the computer mouse grasping in Fig. 8, the robotic hand performed the hand closing action to grab the computer mouse after receiving the EMG signals. The computer mouse was also not damaged nor slipped during grasping. Therefore, the proposed robot hand passed the tests.

Despite the initial success in controlling the robot hand to perform the simple tasks, the robotic hand has some notable difficulties during the tests. The first problem is the design of the hand. The robot hand can be seen grasping small objects with difficulties. As seen in Fig. 8(b) and Fig. 8(c), the computer mouse was held by the fingertip of the robotic

hand. Another problem is the robotic hand movements. The robotic hand moves according to the programmed movements and pressure sensors. This can be unfavorable for tasks which require the specific amount of force such as handling extremely fragile objects.

CONCLUSIONS

This paper presented the myoelectric hand developed through 3D printing. The robotic hand is controlled by EMG signals from muscles in the human arm. Each part of the robotic hand was printed from the 3D printer. The main controller is the Arduino MEGA. EMG signals are received from the flexor and extensor muscles located in the human arm. Experimental results showed that the robotic hand can perform simple actions in terms of closing and opening the hand, regarding the EMG signals. The robotic hand was tested initially with two object grasping tasks. Objects in the tests were grabbed and held with the proper force without slipping from the robotic hand. The future works include the improvements of the hand design and further applications with deep learning approaches.

REFERENCES

[1] K. Ziegler-Graham, E. J. MacKenzie, P. L. Ephraim, T. G. Travison, R. Brookmeyer, "Estimating the Prevalence of Limb Loss in the United States: 2005 to 2050," Archives of Physical Medicine and Rehabilitation, vol. 89, issue 3, March 2008, pp. 422–429.

[2] G. Lisi, P. Belluco, D. Cattaneo, G. Gini, "From the Classification of EMG Signals to the Development of a New Lower Arm Prosthesis," IFAC Proceedings Volumes, vol. 44, issue 1, January 2011, pp. 6493–6498.

[3] S. Sudarsan, E. Chandra Sekaran, "Design and Development of EMG Controlled Prosthetics Limb," Procedia Engineering, vol. 38, 2012, pp. 3547–3551.

[4] G. Vasan, P. M. Pilarski, "Learning from demonstration: Teaching a myoelectric prosthesis with an intact limb via reinforcement learning," International Conference on Rehabilitation Robotics (ICORR), London, UK, July 17–20, 2017, pp. 1457–1464.

[5] A. H. Moreira, et. al., "Development of a Hybrid Arm Prosthesis Controlled by EEG Signals", International Conference on Cybernetics, Robotics and Control (CRC), Chengdu, China, July 21–23, 2017, pp. 203–207.

[6] T. Agcayazi, M. McKnight, P. Sotory, H. Huang, T. Ghosh, A. Bozkurt, "A scalable shear and normal force sensor for prosthetic sensing," IEEE SENSORS, Glasgow, UK, October 29–1 November 2017, pp. 1–3.

[7] P. Wattanasiri, P. Tangpornprasert, C. Virulsri, "Design of Multi-Grip Patterns Prosthetic Hand With Single Actuator," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 26, issue 6, April 2018, pp. 1188–1198.

[8] I. Ruhunage, C. J. Perera, K. Nisal, J. Subodha, T. D. Lalitharatne, "EMG signal controlled transhumerai prosthetic with EEG-SSVEP based approch for hand open/close," IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, October 5–8, 2017, pp. 3169–3174.

[9] R. Tamilselvi, et. al., "EMG activated robotic arm for amputees," International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, January 19–20, 2018, pp. 456–461.

[10] V. Garcia, C. Varol, "Digital forensics of 3D printers," International Symposium on Digital Forensic and Security (ISDFS), Antalya, Turkey, March 22–25, 2018, pp. 1–8.

[11] S. Pan, S. Guo, L. Shi, Y. He, Z. Wang, Q. Huang, "A spherical robot based on all programmable SoC and 3-D printing," IEEE International Conference on Mechatronics and Automation, Tianjin, China, August 3–6, 2014, pp. 150–155.

[12] Y. E. Tan, S. K. Moon, "Inflatable wing design for micro UAVs using indirect 3D printing," International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Kuala Lumpur, Malaysia, November 12–15, 2014, pp. 545–546.

[13] M. Tavakoli, J. Lourenço, A. T. de Almeida, "3D printed endoskeleton with a soft skin for upper-limb body actuated prosthesis," IEEE 5th Portuguese Meeting on Bioengineering (ENBENG), Coimbra, Portugal, February 16–18, 2017, pp. 1–5.

[14] J. Koprnický, P. Najman, J. Šafka, "3D printed bionic prosthetic hands," IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics (ECMSM), Donostia-San Sebastian, Spain, May 24–26, 2017, pp. 1–6.

Revised Scheme for Antimagic Graph Labeling for Splitting Graph and Shadow Graph Associated with Cycle

Zareen Tasneem
Computer Science and Engineering (CSE)
Military Institute of Science and Technology (MIST)
Dhaka, Bangladesh
mishmazareen@gmail.com

Farissa Tafannum
Computer Science and Engineering (CSE)
Military Institute of Science and Technology (MIST)
Dhaka, Bangladesh
tafannumfarissa4@gmail.com

Md Mahbubur Rahman
Computer Science and Engineering (CSE)
Military Institute of Science and Technology (MIST)
Dhaka, Bangladesh
mahbubcse@yahoo.com

***Abstract*—Antimagic graph labeling is the allotment of positive numbers to the edges of a graph such that the sum of graph edge labels incident to each node varies. The paper emphasizes on strong antimagic graph labeling for splitting graph and shadow graph of cycle. Modified theorems and their proofs on splitting and shadow graphs of cycle have been presented. The validity of the theorems has been illustrated through numerous examples.**

Keywords—antimagic graph, splitting graph, shadow graph, cycle, strong antimagic graph, weak antimagic graph

I. INTRODUCTION

Graph labeling can be defined as placing labels on the edges and nodes of a graph. These labels usually consist of integers. It's a very popular topic for research and many informative papers on different types of graph labeling have been published till now. Antimagic graph labeling is one of such type which was founded by Hartsfield and Ringel in 1990 [4]. This research basically focuses on antimagic graph labeling for splitting graphs and shadow graphs of cycle.

In [1] Vaidya and Vyas have introduced some theorems on antimagic graph labeling of total graph, middle graph, splitting graph and shadow graph of cycle and path. The theorems based on splitting graph and shadow graph of cycle from [1] have been modified in the paper by introducing some functional equations which have been narrated in the following sections with proper illustrations.

All the graphs used in the work are finite, connected, undirected and labeled. Here each graph is represented by G where $|V(G)|$ and $|E(G)|$ are its number of nodes and edges respectively and C_n is the cycle formed by the graph.

In this article Section II holds the literature review. All necessary definitions are stated in Section III. Section IV holds in our work and the outcomes of the work are briefly narrated in Section V. Lastly, Section VI brings a conclusion to the article.

II. LITERATURE ANALYSIS

In 1963 the notion of magic graph labeling was used by Sedlacek in his paper [4].

The theory of antimagic graph labeling was founded by Hartsfield and Ringel, where they verified that paths P_n where $n < 3$, cycles, wheels and complete graphs K_n where $n<3$ are antimagic graph in [3].

In [5] it has been demonstrated that graphs attained by switching of a pendant node in path P_n, switching of a node in the cycle C_n, switching of a rim node in wheel W_n, switching of apex node in helm H_n and switching of a node of degree 2 in fan F_n are antimagic graphs.

In [1] the authors investigated antimagic graph labeling for middle, total, splitting and shadow graphs of path and cycle respectively.

In [2], the authors have presented revised theorems on antimagic labeling for middle and total graph associated with a path that have been introduced in [1].

In [6], Michael Jackanich defined the concepts of strong and weak antimagic graph labeling and showed that any bipartite graph has weak ant-magic labeling.

In [7], C.Meenakshi and KM.Kathiresan have worked on magic and antimagic labeling of graphs and discussed k – magicness of complete bipartite graph $K_{p,p}$ where $1<k<2p$.

In [8] the authors defined degree splitting graph of a graph and studied some properties of it.

In [9], P. Lawrence Rozario Raj, S.Koilraj worked on splitting graphs of cycle C_n, path P_n, complete bipartite graph $K_{m,n}$, matching M_n, wheel W_n and $< K^{(1)}_{l,n} : K^{(2)}_{l,n} : \ldots : K^{(k)}_{l,n} >$ and proved all of them are cordial.

In [11], S. K. Vaidya and R. M. Pandit did analysis on shadow graphs, middle graphs, and total graphs of paths and cycles and determined the edge domination number.

The research mainly features on splitting and shadow graphs of cycle. The theorems introduced in [1] on splitting and shadow graphs of cycle have been improved in this paper with proper illustrations.

III. PRELIMINARIES

A. Antimagic Graph

It is a graph where positive integer values are used for labeling the edges in a manner, where the summation of the edge labels incident to each node varies from one another and its labeling is known as antimagic graph labeling. These graphs can be classified into strong and weak antimagic graph.

The first one is a type, where each edge label is completely different from one another. Such labeling is called strong antimagic graph labeling [2]. On the contrary, the second one

is a type where each edge label is not necessarily different from one another. Such type of labeling is weak antimagic graph labeling [2].

B. Splitting Graph

Splitting graph S'(G) of graph G is such a graph which is constructed in a way where a new node v' is added correlating with each node v of graph G and $N(v) = N(v')$. Here, $N(v)$ and $N(v')$ are the neighborhood sets of v and v' [1].

C. Shadow Graph

Let us first take two copies of a cycle G, say G' and G''. Now if we connect each node v' of G' to the neighbors of the correlating node v'' in G'', the graph that forms is defined as shadow graph $D_2(C_n)$.

D. Cycle

Cycle graph (C_n) having n nodes is a graph which is composed of a single cycle that is, the connected nodes form a closed chain.

IV. Our Work

A. For Splitting Graph

In [1], the authors used two cases, Case 1: $n \equiv 1 \pmod 2$ and Case 2: $n \equiv 0 \pmod 2$ whereas Case 2 contains two sub-cases to prove that splitting graph of cycle C_n is antimagic. But it deprives us from gaining a strong antimagic one as output. So, to gain a strong antimagic graph, we have revised the theorem and come up with a new theorem and evinced it.

Theorem 1.1: **Splitting graph of cycle C_n is antimagic.**

Proof: Let $v_1,v_2,v_3,.....v_n$ represent the nodes and $e_1,e_2,e_3,.....e_n$ the edges of cycle C_n. Let v_1', v_2', v_3',....,v_n' represent the newly added nodes to form the splitting graph of cycle C_n. Let G be $S'(C_n)$, the splitting graph of cycle C_n. $V(S'(C_n))=\{v_i, v_i' / 1 \le i \le n\}$ and $E(S'(C_n)) = \{v_i'v_{i+1};\ 1 \le i \le n-1,\ v_nv_1,\ v_1'v_n,\ v_i'v_{i-1};\ 2 \le i \le n,\ v_iv_{i+1};\ 1 \le i \le n-1,\ v_nv_1\}$. Here $|V(G)| = 2n$ and $|E(G)| = 3n$. To define $f:E(G) \rightarrow \{1,2,3,.....3n\}$ we consider following two cases.

Case 1: $n \equiv 1 \pmod 2$

For $1 \le i \le n-1$:

$f(v_i\ v_{i+1}) = 2n+1+i;$
$f(v_n\ v_1) = 2n;$
$f(v_n'\ v_1) = 2n+1;$

$$f(v_i'v_{i+1}) = \begin{cases} 2i+1, & i \equiv 1 \pmod 2 \\ 2i, & \text{otherwise.} \end{cases}$$

For $2 \le i \le n$:

$$f(v_i'v_{i-1}) = \begin{cases} 2i-1, & i \equiv 1 \pmod 2 \\ 2i-2, & \text{otherwise.} \end{cases}$$

$f(v_1'v_n) = 1;$

Case 2: $n \equiv 0 \pmod 2$

For $1 \le i \le n-1$:

$f(v_i\ v_{i+1}) = 2n+1+i;$
$f(v_n\ v_1) = 3n + 1;$
$f(v_n'\ v_1) = 2n;$

Sub Case 1: $n \equiv 0 \pmod 4$

$f(v_2'\ v_1) = 4;$
$f(v_2'\ v_3) = 2;$
$f(v_1'\ v_n) = 1;$

For $1 \le i \le n-1$, $(i \ne 2)$:

$$f(v_i'\ v_{i+1}) = \begin{cases} 2i+1, & i \equiv 1 \pmod 2 \\ 2i, & \text{otherwise.} \end{cases}$$

For $3 \le i \le n$:

$$f(v_i'\ v_{i-1}) = \begin{cases} 2i-1, & i \equiv 1 \pmod 2 \\ 2i-2, & \text{otherwise.} \end{cases}$$

Sub Case 2: $n \equiv 1 \pmod 4$

$f(v_1'\ v_n) = 3;$
$f(v_1'\ v_2) = 1;$

For $2 \le i \le n$:

$$f(v_i'\ v_{i+1}) = \begin{cases} 2i+1, & i \equiv 1 \pmod 2 \\ 2i, & \text{otherwise.} \end{cases}$$

$$f(v_i'\ v_{i-1}) = \begin{cases} 2i-1, & i \equiv 1 \pmod 2 \\ 2i-2, & \text{otherwise.} \end{cases}$$

These portrayed edge labeling functions will lead to all the distinct node labels fulfilling the state for antimagic labeling. Consequently, $S'(C_n)$ is antimagic.

Analogy 1.1 The theorem is illustrated through some examples. Splitting graph of cycle C_4 is antimagically labeled in Figure 1. C_4 has four nodes and four edges. So, its splitting graph has eight nodes and twelve edges. According to our proof, Sub-Case 2 of Case 2 is relevant for $n=4$ and implementation is done accordingly.

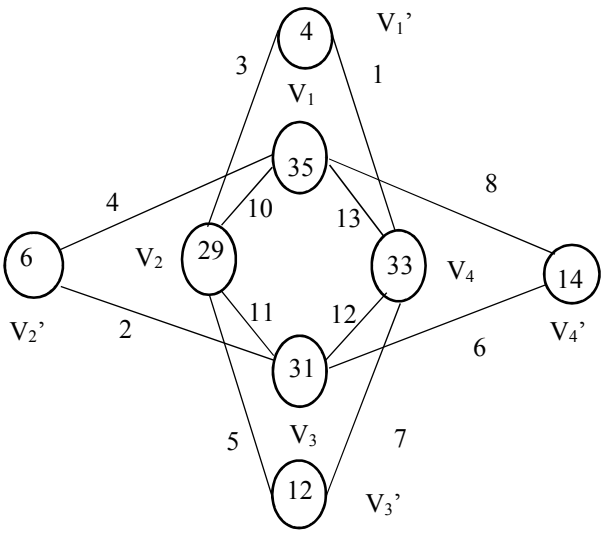


Fig. 1. Antimagically labeled splitting graph of cycle C_4.

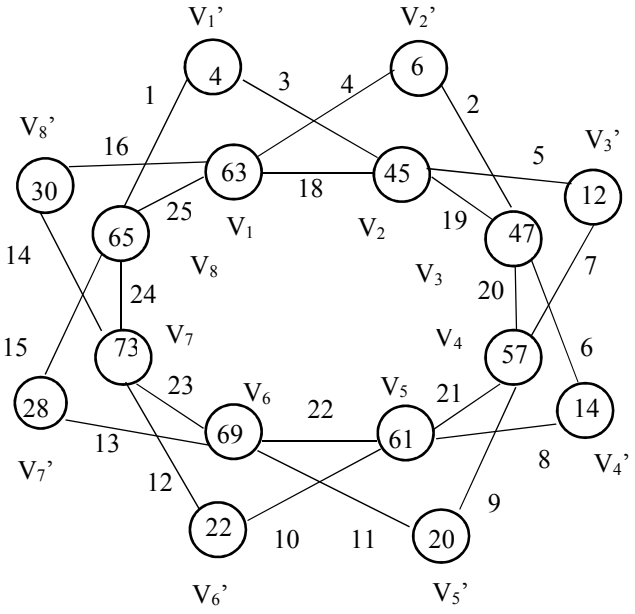


Fig. 2. Antimagically labeled splitting graph of cycle C_8.

Figure 2 displays us antimagically labeled splitting graph of cycle C_8. C_8 has eight nodes and eight edges. So, its splitting graph has sixteen nodes and twenty-four edges. According to our proof, Sub-Case 2 of Case 2 is relevant for $n=8$ and implementation is done accordingly.

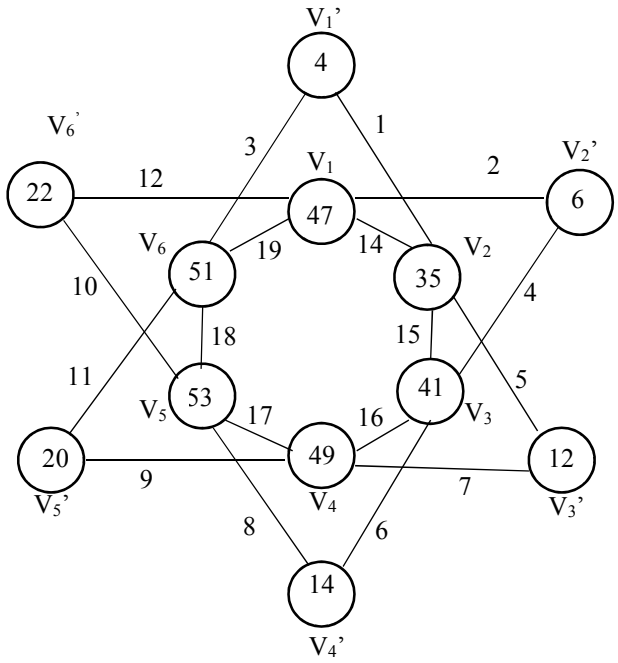


Fig. 3. Antimagically labeled splitting graph of cycle C_6.

Figure 3 displays us, antimagically labeled splitting graph of cycle C_6. C_6 has six nodes and six edges. So, its splitting graph consists of twelve nodes and eighteen edges. According to our proof, Sub-Case 1 of Case 2 is relevant for n=6 and implementation is done accordingly.

B. For Shadow Graph

In [1], the authors used three cases, Case 1: $n \equiv 1 \pmod 2$, Case 2: $n \equiv 0 \pmod 2$, $n \neq 6$ and Case 3 to demonstrate that shadow graph of cycle C_n is antimagic, which doesn't give a strong antimagic graph as an output. So, to gain a strong antimagic graph, we have revised the theorem and come up with a new theorem and evinced it.

Theorem 1.2: **Shadow graph of cycle C_n is antimagic.**

Proof: Let C'_n ,C''_n be two copies of cycle C_n . We indicate the nodes of first copy of C_n by $v_1', v_2', v_3', \ldots, v_n'$ and second copy by $v_1'', v_2'', v_3'', \ldots, v_n''$. Let G be $D_2(C_n)$ with $|V(G)|=2n$ and $|E(G)| = 4n$. In order to define $f:E(G)\rightarrow\{1,2,3,\ldots..4n\}$ the two cases given below have been considered.

Case 1: $n \equiv 1 \pmod 2$

For $1 \leq i \leq n-1$:

$f(v_1'' v_{i+1}'') = i;$

$f(v_n'' v_1'') = n;$

$f(v_i' v_{i+1}') = 3n + i;$

$f(v_n' v_1') = 4n;$

$f(v_i'' v_{i+1}') = n + 2i - 1;$

$f(v_n'' v_1') = 3n - 1;$

$f(v_i' v_{i+1}'') = n + 2i;$

$f(v_n' v_1'') = 3n;$

Case 2: $n \equiv 0 \pmod 2$,

$f(v_n'' v_1'') = 1;$

$f(v_n' v_1') = 4n;$

$f(v_n' v_1'') = n + 1;$

$f(v_1' v_n'') = n + 2;$

For $1 \leq i \leq n-1$:

$f(v_i'' v_{i+1}'') = 3n + i;$

$f(v_i' v_{i+1}'') = n + 1 + 2i;$

$f(v_i' v_{i+1}') = i + 1;$

For $2 \leq i \leq n$:

$f(v_i' v_{i-1}'') = n + 2i;$

These portrayed edge labeling functions will lead to all the distinct vertex labels satisfying the state for antimagic labeling. Consequently, $D_2(C_n)$ is antimagic.

Analogy 1.2 The theorem is illustrated through some examples. Figure 4 displays us, antimagically labeled shadow graph of cycle C_5. C_5 has five nodes and five edges. So, its shadow graph has ten nodes and twenty edges. According to our proof, Case 1 is relevant for $n=5$ and implementation is done accordingly.

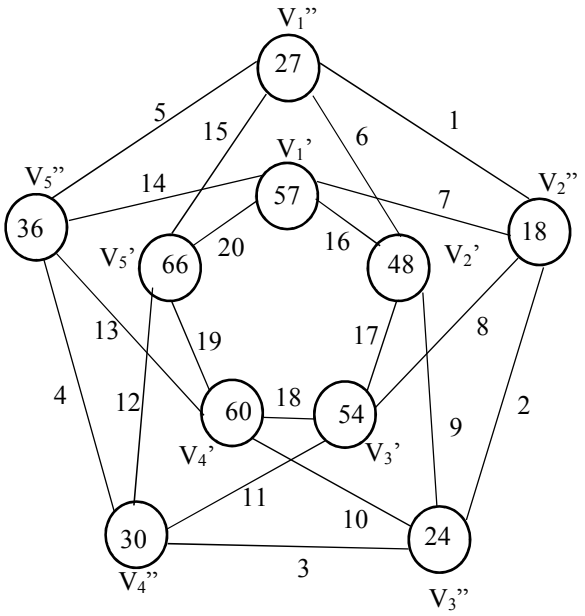


Fig. 4. Antimagically labeled shadow graph of cycle C_5.

Shadow graph of cycle C_6 is antimagically labeled in Figure 5. C_6 has six nodes and six edges. So, its shadow graph consists of twelve nodes and twenty-four edges. According to our proof, Case 2 is relevant for *n=6* and implementation is done accordingly.

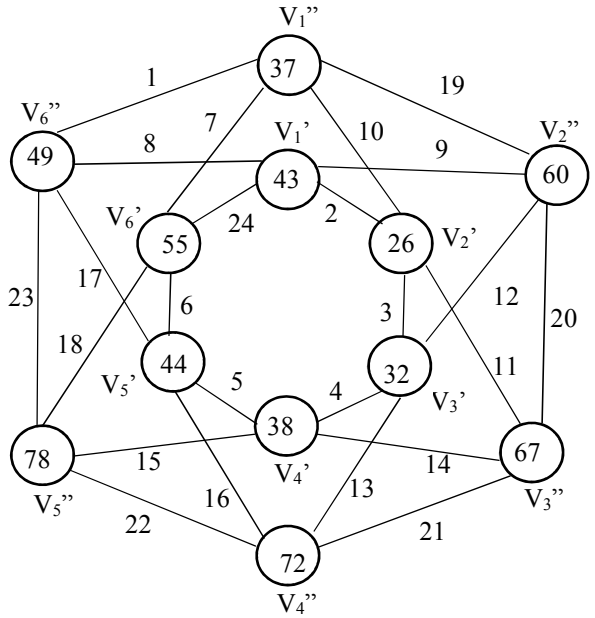


Fig. 5. Antimagically labeled shadow graph of cycle C_6.

Figure 6 displays us antimagically labeled shadow graph of cycle C_8. C_8 has eight nodes and eight edges. So, its shadow graph has sixteen nodes and thirty-two edges. According to our proof, Case 2 is relevant for *n=8* and implementation is done accordingly.

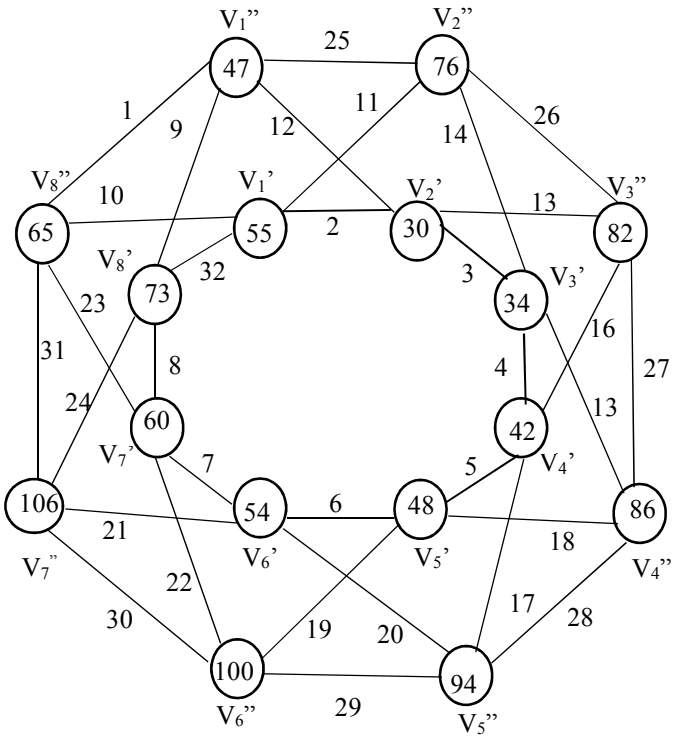


Fig. 6. Antimagically labeled shadow graph of cycle C_8.

V. RESULTS

A. For Splitting Graph

We improved the validation of the theorem which is unlike the one conveyed in [1].

In [1], the authors demonstrated two cases, Case 1: $n \equiv 1 \pmod 2$ and Case 2: $n \equiv 0 \pmod 2$ whereas, Case 2 consists of two sub-cases, Sub Case 1: $n \equiv 0 \pmod 4$, $n \neq 4$ and Sub Case 2: $n = 4, n \equiv 1 \pmod 4$. After implementing the theorem, we saw that Case 1 works perfectly providing a strong antimagic graph as an output. But, by using the edge label equations of Case 2, we don't achieve a strong antimagic graph, in fact, there remain some nodes with no values.

According to the equations of [1], when we used $f(v_n v_1) = 3n$, then we got a weak antimagic graph as an output. For instance, when $n=6$, we got value 18 for both v_6v_1 and v_5v_6 edges which made the graph weak.

For solving this problem, we used the function equation $f(v_n v_1) = 3n + 1$. When we applied this, we got a strong antimagic graph. Suppose, when $n=6$, we got value 19 and 18 for v_6v_1 and v_5v_6 edges which made the graph a strong antimagic.

In Case 2 provided in [1], there was no equation for $f(v_n' v_1)$. So, there was no way to get any value for the edge $v_n' v_1$. As a result, we couldn't get any value for nodes v_n' and v_1. Hence, it couldn't lead to any complete antimagic graph.

In order to resolve this issue, we provided $f(v_n' v_1) = 2n$. As a result, we got the value for the edge $v_n'v_1$ as well as the values of v_n' and v_1 nodes. All these can lead to a strong antimagic graph.

In Sub-case 1 of [1], when we used the condition, $1 \leq i \leq n, (i \neq 2)$ then we got an extra vertex and an extra unnecessary edge. For instance, when, $n=8$, we got an extra vertex v_9 and an extra edge $v_8'v_9$ whereas none of these exist in the graph.

For solving this issue, we used the condition, $1 \leq i \leq n - 1, (i \neq 2)$. Hence, we got rid of getting any extra and unnecessary vertex and edge which didn't exist in the graph.

Also, according to [1], Sub-case 1 was applicable for $n \equiv 0$ (mod 4) but not for $n=4$, Sub-case 2 was applicable for $n \equiv 1$ (mod 4) and $n=4$. But, after all the modification has been done to the theorem, Sub-case 1 worked for $n \equiv 0$ (mod 4) as well as $n=4$. The corrected implementation is provided in Figure 1.

B. For Shadow Graph

The proof of the theorem has been improved such that it differs from the proof provided on [1].

In [1], $f(v_i'v_{i-1}'')=n+2i$ was included in Case 1. But if that function equation was used then we would get values for some edges double times. Also, there was no function equation provided for $v_i'v_{i-1}''$. For example, when $n=5$, using $f(v_i'v_{i-1}'')=n+2i$, we would get double values for the edges $v_2'v_1''$, $v_3'v_2''$, $v_4'v_3''$ as there was another function equation given, which provided values for those same edges, whereas we couldn't get any value for the edges $v_1'v_2''$, $v_2'v_3''$, $v_3'v_4''$, $v_4'v_5''$.

For solving this problem we provided $f(v_i'v_{i+1}'')=n+2i$. It helped to find the values for the edges $v_i'v_{i-1}''$ and so, there was no more double values for $v_i'v_{i+1}''$.

In [1], the authors have given $f(v_1'v_n'')=3n$ for Case 2. But using this function equation, it would result into a weak antimagic graph. For instance, for $n=8$, we got the value 24 for both $v_1'v_8''$ and $v_8'v_7''$ which lead us to a weak antimagic graph.

In order to resolve this issue, we provided $f(v_1'v_n'') = n+2$. Using that, we obtained a strong antimagic graph. For instance, for $n=8$, we got the value 10 and 24 for $v_1'v_8''$ and $v_8'v_7''$ respectively which lead us to a strong antimagic graph.

Vaidya and Vyas showed Case 3 whereas there was no function equation to find the values for the edges of the graph in [1] and that Case 2 was not applicable for n=6. But, the modification that we provided didn't require any Case 3 and that Case 2 supported $n=6$. The corrected implementation is provided in Figure 5.

VI. CONCLUSION

We have appraised the antimagic labeling for splitting and shadow graph of cycle. The result and the examples reveal that our modification of the theorem provided in [1] produces better performance than those explained in [1]. We have presented the proofs of the theorem along with some examples in section IV and in section V we have unfolded the scope of improvements found in [1] and elaborately discussed our solution that we figured out. Magic and antimagic labeling are feasible for many more graphs and there is a huge scope to adopt further researches in this sector.

REFERENCES

[1] S.K. Vaidya and N.B. Vyas, "Antimagic Labeling of Some Path and Cycle Related Graphs," Annals of Pure and Applied Mathematics, vol. 3 , No. 2, July 2013, pp. 119- 128.

[2] Zareen Tasneem, Farissa Tafannum, Maksuda Rahman Anti, Wali Mohammad Abdullah and Md. Mahbubur Rahman, "A Modified Anti-Magic Graph Labeling for Middle Graph and Total Graph Related to Path," 2016 19th International Conference on Computer and Information Technology (ICCIT), Dhaka, 2016, pp. 339-34.

[3] N.Hartsfield and G.Ringel, "Pearls in Graph Theory: A Comprehensive Introduction," Academic Press, Boston, 1990, pp. 108-110.

[4] J.Sedlacek , "Problem 27. In Theory of Graphs and its Applications," Proc. Symposium Smolenice, June 1963, pp. 163-167.

[5] S.K. Vaidya and N.B. Vyas, "Antimagic labeling in the context of switching of a vertex," Annals of Pure and Applied Mathematics, vol. 2, December 2012,pp. 33 – 39.

[6] Michael Jackanich, "Antimagic Labeling of Graphs," version of April 13, 2011.

[7] C.Meenakshi, KM.Kathiresan, "A generalization of magic and antimagic labelings of graphs," AKCE International Journal of Graphs and Combinatorics, 2018.

[8] Raja Ponraj and S Somasundaram, "On the degree splitting graph of a graph," National Academy Science Letters. 27 (7-8), January 2004, pp. 275-278.

[9] P. Lawrence Rozario Raj, S.Koilraj, "Cordial labeling for splitting graph of some standard graphs," International Journal of Mathematics and Soft Computing, vol. 1, no. 1, 2011, pp. 105 – 114.

[10] Tao-Ming Wang, Cheng-Chih Hsiao, "On anti-magic labeling for graph products," vol. 308, Issue 16, 2008, pp. 3624-3633.

[11] S. K. Vaidya and R. M. Pandit, "Edge Domination in Some Path and Cycle Related Graphs," ISRN Discrete Mathematics, vol. 2014, Article ID 975812, 5 pages, 2014.

[12] Wikipedia [Online]. Available: https://en.wikipedia.org/wiki/ Magic graph.

[13] Anti-magic Graph [Online]. Available: http://mathworld.wolfram.com/ AntimagicGraph.html.

[14] R. Balakrishna and K. Ranganathan, A Textbook of Graph Theory, Springer, 2012.

[15] J. A. Gallian, "A Dynamic Survey of Graph Labeling," The Electronics Journal of Combinatorics, 19 (#DS6), 2012.

[16] L. W. Beineke and S. M. Hegde, "Strongly Multiplicative Graphs," Discuss. Math. Graph Theory, 21, (2001), 63-75.

Icons For the Mass: An Approach Towards Text Free Smart Interface

Eshtiak Ahmed
Dept of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
eshtiak.cse@diu.edu.bd

Md. Mahade Hasan
Dept of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
mahade.cse@diu.edu.bd

Mirza Omar Faruk
Dept of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
faruq15-518@diu.edu.bd

Muhammad Farhad Hossain
Dept of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
farhad15-5434@diu.edu.bd

Md. Arifur Rahman
Dept of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
rahman15-5830@diu.edu.bd

Ashraful Islam
School of Computing and Informatics
University of Louisiana at Lafayette,
Louisiana, USA
ashraful.islam1@louisiana.edu

***Abstract*— A major population of 3rd world countries have problems reading out texts because of the low literacy rate. While smartphones have given them a hope to lead a better life, interacting with them needs a certain level of literacy which a majority of them lack. Given the rapid growth of variation in this user base, it is evident that interactive graphical representations can give a better understanding of contents especially to the children and to the illiterate people. This work emphasizes on making the Google play app store more usable as this is the most used app for smartphone users. In this study, new icons have been designed and then compared side by side with 16 existing Google play app category icons to initially find out the legitimacy of the new icons. Further, a survey between smartphone users has been conducted and the results show that the new icons fare better while representing app categories. Based on the survey results and more widely accepted icons, some key factors to design icons have been identified.**

Keywords—Human Computer Interaction, Graphical Representation, Icons.

I. INTRODUCTION

The technological improvement in the field of mobile devices is a landmark of this decade. A smartphone is one of its significant achievement with a more powerful package than its predecessors. According to Statista [1], over 2.53 billion people have used a smartphone up to the year 2018. Statistical information suggested that this number will increase up to 2.87 billion by this year 2020. As the smartphone is reaching to more people, it is a challenging job to maintain a smart interface which can be easily understandable for its massive user groups.

Another survey of Statista states that, the Google Play store currently has a total of around 2.1 million applications while the Apple store has around 2 million applications [2]. Dividing or labeling these large number of apps in different categories or genres are used to help people browsing applications, finding what they need and what they can expect from a particular application. If considered different domains of mobile applications, merely 50 categories can be found at most in any app store. These categories are described by using textual labels such as, for example, "Education" or "Lifestyle" on the Google Play store. However, this kind of representation may not be suitable for people who do not know how to read or still developing reading skills such as children. Although there is some graphical representation associated with the text labels, those are not user-friendly. People often find it difficult to reach the app they wanted [3] [4]. To facilitate these user groups, work has been done in this study towards text free smart interface where icons can solely describe a particular application category.

Text is always the best way to interact with a system [5], but sometimes they are less effective when the person or user is not very familiar with the language. Graphical representation is the best possible alternative of this problem. In order provide a better understanding through graphical representations, new icons were designed and put alongside with the existing icons. Then people were asked to tell which icons they prefer most, and a total of 664 responses were collected. It was found that the people participating in the surveys were able to understand the new set of icons with accuracies over 80% of times for most of the categories. The result of this survey suggests that the existing icon sets are confusing and misleading people every time and using a new set of icons can better describe an application category which will improve their content searching experience and increase the usability of app stores for more users. Finally selection and analysis have been conducted for identifying the characteristics of more widely accepted icons.

II. RELATED WORK

Human-computer interaction is a very significant field of research as it investigates the way humans interact with computer or computerized machines to ensure better acceptability and convenience [6]. While human refers to people of every age, several fragments have been introduced in order to differentiate and specify works for different age groups. A very specific study on games category icons design has been done where different sets of icons were designed for different mobile gaming categories to investigate if kids can distinguish them as specific categories [7]. This particular study emphasizes on whether children can understand the type of a game solely using icons. This is a very narrow domain and does not cover a significant part of the smartphone users. However, it shows that children and adults differ cognitively and physically, as a result different technology and interaction techniques need to be developed for them [8]. While working for different age groups, it's very important to look at their different types of requirements and behavior [9]. Users' preferences and time to needed to process information from the icons were analyzed to determine the aesthetic appeal of different types of icons from different design eras in [10]. This study goes very theoretical with the time factor and also the preference. However, the objective of developing a set of icons that can be identified without the help of any kind of text

has not been achieved here. In [11], 315 icons were developed for 45 functions specifically for smart building. This work is also a very domain specific which in this case is smart buildings. There has not been much work in terms of designing icons for the mass people especially on a platform like the Google play store which is used by all the smartphone users all over the world.

There has been some studies where the factors of designing icons has been discussed. In [12], six influencing factors were found out regarding aesthetic designs of icons. Although it defines a way of designing aesthetically pleasing icons, a proper analysis has always been demandable. In this study, an analysis has been done from the results obtained from the human survey for newly designing icons by comparing with the existing ones from Google play store keeping the design interaction in mind.

III. App store category selection

In the Google play store, there are more than 35 app categories to choose from. The list of categories have a specific icon beside the category name to represent the category graphically. Primarily during the feasibility study of this research work, a public survey was conducted to know the opinion of users about these category defining icons. From the survey, it became evident that the current icons are very much abstract and do not exactly represent the whole idea of the category. Hence, designing better icons for representing app categories became a viable research goal. Also, every category icon is not poorly designed, so categories were chosen having more useful and widely used apps. Primarily to start the investigation, 16 different app categories were chosen. They are Arts & Design, Beauty, Business, Communication, Entertainment, Finance, Health & Fitness, Lifestyle, Maps & Navigation, News & Magazines, Personalization, Photography, Social, Sports, Travel & Local and, Weather. The primary goal of the study was to design new graphical representations (icons) for these 16 categories. The categories along with their definition is given in the table 1 which can help understand the type of apps that are under a specific category.

IV. Icons and User Survey

A. Icons Design

The primary objective of this study is to design better icons for better graphical representation and furthermore, analyze the icon adaptation behavior of user to understand specifics about the likes and dislikes. The app category icons were tried to be designed in a way which would be more understandable for illiterate, older adults and children with special focus on making them understandable without any help from text. It has been always kept in concern to design fair icons, representing them with real world objects as people can relate to them more easily. The table 2 represents the investigated icons which are kept side by side in order to find the actual icons.

B. Survey Design

After the icons for the 16 selected categories are designed, a survey was conducted among smartphone users to determine if the newly designed icons are any better than the existing ones from the Google play store. A Google form was designed for this purpose. For each category, both icons (1 new and 1 existing from Google play store) were kept side by side and the users were asked to choose the icon that represents that category better.

It was made sure that the participants of the survey were all smartphone users. Also, efforts were made to introduce varieties of users to get a more diversified result. To facilitate this, three extra parameters were introduced such as gender, occupation, and age. Total 664 responses collected from the

TABLE 1: SELECTED APP CATEGORIES AND DEFINITIONS

Category	Definition
Art & Design	Applications to create designs combination of text, colors, and pictures.
Beauty	Image-based makeup guidelines, makeover tools, and simulation applications.
Business	Applications having the functionality of editing or reading documents, tracking of packages, email management, and job searching.
Communication	Applications that help to engage people, share and communicate with each other.
Entertainment	Streaming TV shows, movies, music and many more.
Finance	Applications that help to manage banking, payment, taxes, fetching financial news and done complex financial calculations.
Health & Fitness	Exercise tracking, health & nutrition information, safety tips.
Lifestyle	Applications to guide your style, plan your wedding & party.
Map & Navigation	Applications that can help you to reach your desired locations.
News & Magazines	Blogging applications, newspapers, magazines.
Personalization	Applications that help to decorate your devices wallpapers and ringtones, alarm tones etc.
Photography	Taking photographs, editing, and sharing applications.
Social	Apps to socialize contents.
Sports	Sports coverage, commentary, scores, and news.
Travel & Local	Application to manage your tour, information about places you must visit, city guides etc.
Weather	Weather report and news applications.

participants of the different professions where the majority of them were students. About 513 (77.3%) participants were male while the rest 151 (24.7%) were female. Above 90% of the applicants were students while the others were either job holders or self-employed. According to our data, 81.3% participants were between the age of 21 and 25 as they are university students. 10.1% of them were aged between 16 and 20 while there were participants from age groups 26 to 30, 31 to 35 and, more than 36.
After receiving an adequate number of responses, they were carefully evaluated. The evaluation criteria is simple as all the questions were closed and could only take 2 responses.

V. Results Analysis

From the 664 responses that were recorded via the online survey, a very good idea can be obtained whether the newly designed icons represent better values or not. As mentioned earlier, there were only 2 options, one being the newly designed icons and the other being the existing category icons from Google play store. The Fig. 1 shows the overall survey result on the whole. It shows all the 16 chosen categories along with the acceptance percentages of both set of icons for them.

As there have been only 2 options for the users to choose from, the results percentages from new icons and existing icons add up to the full results. The newly designed icon for "Art and Design" was preferred by 88% participants while the existing one was preferred by the rest 12%. The new "Beauty" icon was preferred by 93.4% participants while the "Business" icon was preferred by 90.8% participants compared to the 6.6% and 9.2% for existing icons respectively. New icons for "Communication", "Entertainment" and "Finance" were preferred by 84.3%, 95.5%, and 95.9% participants respectively. New "Health and Fitness" icon was preferred by 82.1% participants while the new "Lifestyle" icon was preferred by 93.1% participants. The new icon for "Maps and Navigation" was preferred by 92.8% participants and the new "News and Magazines" icon was preferred by 82.4% participants.

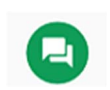
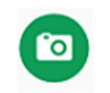
There were some newly designed icons that received mixed responses and have been challenging to make people understand by a small piece of art. Among these type of categories, "Personalization" was preferred by 64.8% participants, "Photography" was preferred by 78.3% participants and "Social" was preferred by 71.2% participants. There were some other icons like "Sports" which was preferred by 82.7% participants while newly designed "Travel and Local" and "Weather" were preferred by 72% and 91.1% participants respectively.

According to the result, most people have preferred the newly designed icons to the existing ones except from a couple of categories like personalization. This happened because of the abstractness of the category as it cannot be represented by a specific object.

VI. Challenges And Discussion

At the phase of designing the new icons there have been several challenges. As example, designing the icon for the category "Beauty", there was a dispute whether it should be represented by beauty products or anything else. Another challenging scenario was the icon of "Lifestyle" as it cannot be represented by a specific object, so summing a daily life picture in a single icon is very difficult. In the end a suitable

Table 2: Side by Side Comparison of Icons

Category Name	Designed Icons	Google play Icons
Arts & Design		
Beauty		
Business		
Communication		
Entertainment		
Finance		
Health & Fitness		
Lifestyle		
Maps & Navigation		
News & Magazines		
Personalization		
Photography		
Social		
Sports		
Travel & Local		
Weather		

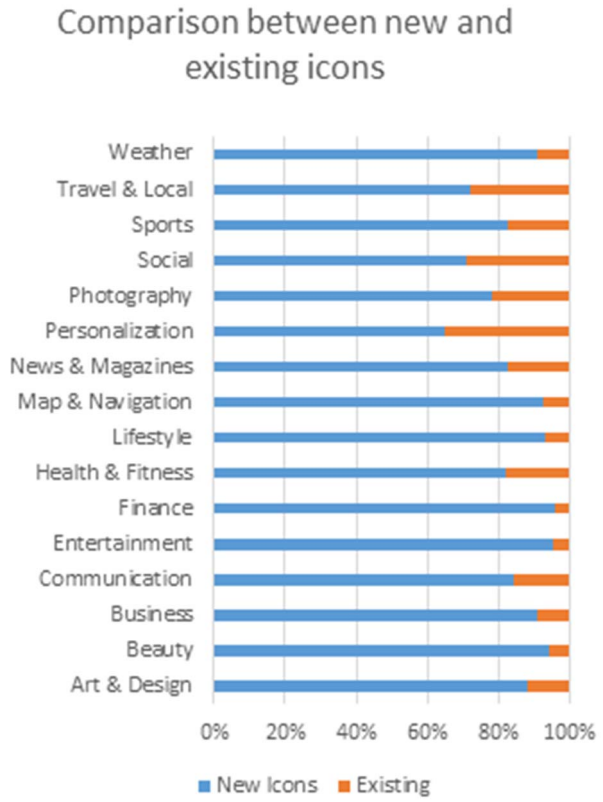


Fig. 1: Analysis of results obtained from the survey

icon was designed but when seen in a smaller format, it might seem too crowded. While designing an icon, the primary approach taken was representing the category by objects belonging to them. Such as the "Sports" icon has been designed by combining three of the most popular sports equipment.

The icons of "Finance" and "Business" were also very challenging as these categories are quite similar and related to each other. Similarly, "Communication" and "Social" were also challenging because of their connections with each other. From the fact of preference given to icons by users, specific points can be noted by analyzing them to find out the actual components of a successfully designed icon. As example, the "Arts and design" logo is represented by a pallet and brushes while the existing one only shows a pallet. Adding the brushes made the pallet more obvious and made a great difference. For beauty, the existing icon is a shape of hairdryer which does not nearly represents beauty.

For the new icon, beauty is represented by a beautiful face. The "Business" icon is represented by two figures of human and a money sign as business means deals between people. "Entertainment" has been represented by a couple of forms of entertainment like video and audio both which makes it easier to understand. The most challenging icon to design has been "Lifestyle" and "Personalization" as these are very abstract ideas. However the main components of people's daily lives have been cramped into one icon to represent "Lifestyle" while not much change could be done to the "Personalization" icon resulting in lower preference rate.

While designing the new icons, the main focus was to represent a specific idea using the objects or components that comes to mind specifically for the categories. It was evident from the results that users of smartphones prefer icons that have relevant objects in them relating to the category. While this is an encouraging start, there are still so much to do. This study has been conducted with a portion of the categories in Google play store. A store wide study can be done for more accurate and diversified result. Another shortcoming of this work is that the survey population is mostly young smartphone users. More thorough survey covering users from different age groups can create a dimension of results.

VII. Conclusion

Today people are using a smartphone beyond the boundary of languages and cultures, they often belong to different age groups, and they have a different level of education or might even be illiterate. This study aims to promote the use of designing interfaces based on graphical representations such as icons which could be the first step towards text free interfaces in mobile industry. The newly designed icons for app categories in Google play store have been appreciated by most users who have completed the survey comparing both sets of icons. Overall, this study is very encouraging for future works keeping text free interfaces in mind. Further extension of this work can be done in a larger scale to graphically represent not only icons but everything in a smartphone.

References

[1] Statista, Inc., "Number of smartphone users worldwide from 2014 to 2020." [Online]. Available: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/. [Accessed: 14-Apr-2019].

[2] Statista, Inc., "Number of apps available in leading app stores as of 3rd quarter 2018." [Online]. Available: https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/. [Accessed: 14-Apr-2019].

[3] M. Wang and X. Li, "Effects of the aesthetic design of icons on app downloads: evidence from an android market", Electronic Commerce Research, vol. 17, no. 1, pp. 83–102, 2017.

[4] K.-C. Hou and C.-H. Ho, "A Preliminary Study on Aesthetic of Apps Icon Design," IASDR 2013 5th Int. Congr. Int. Assoc. Soc. Des. Res., pp. 1–12, 2013.

[5] Google Material Design, "Material Design Guidelines," Material.io, 2017.

[6] E. P. Rozanski and A. R. Haake, "Human–computer interaction," in Systems, Controls, Embedded Systems, Energy, and Machines, pp. 18-1-18–12, 2017.

[7] M. Wiebe, D. Geiskkovitch, A. Bunt, J. Young, and M. Glenwright, "Icons for Kids: Can Young Children Understand Graphical Representations of App Store Categories?," in Proceedings of the 42Nd Graphics Interface Conference, pp. 163–166, 2016.

[8] M. Weiser, "The computer for the 21st Century," IEEE Pervasive Comput., vol. 1, no. 1, pp. 19–25, 2002.

[9] S. M. Huang, K. K. Shieh, and C. F. Chi, "Factors affecting the design of computer icons," Int. J. Ind. Ergon., vol. 29, no. 4, pp. 211–218, 2002.

[10] J. M. Silvennoinen and J. P. P. Jokinen, "Aesthetic Appeal and Visual Usability in Four Icon Design Eras," in Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16, pp. 4390-4400, 2016.

[11] C. F. Chi and R. S. Dewi, "Visual and auditory icons for intelligent building," in Proceedings of 2014 International Conference on Intelligent Green Building and Smart Grid, IGBSG 2014, pp. 1-5 2014.

[12] C. H. Ho and K. C. Hou, "Exploring the attractive factors of app icons," KSII Trans. Internet Inf. Syst., vol. 9, no. 6, pp. 2251–2270, 2015.

Efficient Compression Scheme for Large Natural Text Using Zipf Distribution

Md. Ashiq Mahmood and K.M. Azharul Hasan
Department of Computer Science and Engineering
Khulna University of Engineering & Technology (KUET)
Khulna, Bangladesh
Email: ashiqmahmoodbipu@gmail.com, azhasan@gmail.com

Abstract— **Data compression is the way toward modifying, encoding or changing over the bit structure of data in such a way that it expends less space. Character encoding is somewhat related to data compression which represents a character by some sort of encoding framework. Encoding is the way toward putting a succession of characters into a specific arrangement for effective transmission or capacity. Compression of data covers a giant domain of employments including information correspondence, information storing and database development. In this paper we propose an efficient and new compression algorithm for large natural datasets where any characters is encoded by 5 bits called 5-Bit Compression (5BC). The algorithm manages an encoding procedure by 5 bits for any characters in English and Bangla using table look up. The look up table is constructed by using Zipf distribution. The Zipf distribution is a discrete distribution of commonly used characters in different languages. 8 bit characters are converted to 5 bits by parting the characters into 7 sets and utilizing them in a solitary table. The character's location is then used uniquely encoding by 5 bits. The text can be compressed by 5BC is more than 60% of the actual text. The algorithm for decompression to recover the original data is depicted also. After the output string of 5BC is produced, LZW and Huffman techniques further compress the output string. Optimistic performance is demonstrated by our experimental result.**

Keywords— ***encoding; compression; decompression; 5-bit compression; compression ratio.***

I. INTRODUCTION

Data compression is broadly used and imperative issue in information science applications because of capacity limit and exchange speed of different systems [1][2]. There are undeniable advantages which may be accomplished incorporates diminishing capacity necessity and data transmission, encoding littler bit numbers, short time essential for broadcasting, powerful channel utilization etc. [3][5]. Having a decent packed scheme conspire remapping any text by shorter bit is vital [1][2][5]. So a new system of encoding can be important representing to a source of data precisely conceivable utilizing least bit numbers putting the actual meaning unaltered [1][2][11]. In many cases, to reestablish data into the original organization because of the compacted structure is vital [6]. Zipf distribution is used in this paper to construct the dictionary of characters. It actually came from Zipf's law. Zipf's law can be said to an experimental law formulated utilizing mathematical statistics that alludes to the way that numerous sorts of data studied in the physical and social sciences can be approximated with a Zipf distribution. Zipf's law expresses that given some corpus of natural language utterances, the frequency of any word is conversely corresponding to its position in the frequency table [4]. In this way the most frequent word will happen approximately twice as regularly as the second most frequent word, three times as regularly as the third most frequent word. The compression techniques of data are ordered extensively as lossy and lossless [8]. Lossy technique comprises towards a change to imperative against immaterial data, trailed by compression of lossless data in the essential case and disposing rest of the data. As a matter of fact, Lossy technique functions by decreasing amount of document by reducing superfluous text or data [1][11]. The 5BC algorithm which we proposed is lossless [1][2][8]. The technique deals with the mapping of both forward and backward. 5BC alters characters from 8 bit to 5 bits by parting the characters into 7 sets and utilizing them in a solitary table. By using Zipf distribution, the characters in the table i.e. dictionary are mapping [4][12]. Same characters in the corresponding set code are kept together since it produces shorter sequence of bit code. The 5BC can be able to compress data by more than 60% of the actual data. 5BC compression technique is a logical scheme that shows promising efficiency.

II. RELATED WORKS

The Huffman technique is supposed to be a lossless technique which relies on entropy [9]. It actually construct some binary trees so the codes can be found from the binary tree sequence. Huffman algorithm is said to be the variable length encoding as it mainly works with length of the variable. In the contrary, our 5BC algorithm constantly settled to work with 5 bits as it were. The rule used in this technique is utilizing progressively visit each data via a shorter amount of bits for encoding. In the records of JPEG, Huffman technique is used. JPEG 2000 (JP2) is supposed to be an image compression standard and coding system [13]. Static Huffman and Adaptive Huffman

are the classification of Huffman algorithm [9]. Static Huffman ascertain frequencies right off bat and afterward they create some typical tree for the forward and backward mapping [9][15] though Adaptive Huffman functions by building up trees by figuring frequencies thus constructs two trees in both of these running processes [14]. There exists some compression techniques which are based on dictionaries rather than factual organizations [13][15]. A well-known dictionary based technique is Lempel-Zev Welch or LZW technique [9]. In this algorithm, an individual string αK is constructed by a string α including a K which can be found through the dictionaries. α is additionally added into the dictionary [10]. At the end, the character of string has additionally supplanted again [5]. The existing dictionary functions in form of dynamic information thus it doesn't show to be static [10]. This crude dictionary can be recuperated through the decompressed information at the time of disentangling it. The LZW dictionary can be precisely worked amid the definite specific during Compression and decompression, and disposed of in the wake of packing or process of decompressing has been ended [10]. There is another technique named LZ Technique [5][13] which is additionally popular because of it's straightforwardness as well as It has effective compression proportions. Our technique utilizes a table look up which remains a static one and it will not change during the arrival of new data. The look up table i.e. the dictionary remains fixed and can be utilized for the purpose of compression and decompression as time required.

III. EFFICIENT COMPRESSION SCHEME FOR LARGE NATURAL TEXT

Generally, for the purpose of encoding a character in a memory, 8 bit space in a memory is required [1][2]. We introduce a character encoding scheme of 5 bit called 5BC representing a character via 5 bits in lieu of 8 bits. The algorithm works with all the general printable characters covering both English and Bangla characters that has in normal keyboard. A lookup table is built for representing the Characters by 5 bit. The table is illustrated in TABLE I.

Using 5 bit it can be $2^5 = 32$ combination possible. Among the 32 combinations 25 combinations are used for compression and the rest of the 7 are used for set representation. The character sets are splitted within the number of 7 sets. In the lookup table,

- Bangla characters are placed in Set1, Set2, Set3
- English characters are placed in Set4, Set5, Set6 & Set7
- 7 binary combination are taken for the 7 set.

The Unicode takes 16 bit of memory for representing the natural language character set. This modified 5 bit characters will be converted to 16 bit Unicode characters.

TABLE I. Lookup table for 5BC

Serial no.	Decimal value	Binary value	Set1	Set2	Set3	Set4	Set5	Set6	Set7
1	0	00000	অ	ণ	উ	A	Space	1	X
2	1	00001	আ	ত	ঊ	B	,	2	Y
3	2	00010	ই	থ	ঋ	C	a	3	Z
4	3	00011	ঈ	দ	এ	D	b	4	x
5	4	00100	া	ধ	ঐ	E	c	5	y
6	5	00101	ি	ঘ	ও	F	d	6	z
7	6	00110	ী	প	ঔ	G	e	7	'
8	7	00111	ু	ফ	।	H	f	8	!
9	8	01000	ে	ব	০	I	g	9	"
10	9	01001	ো	ভ	১	J	h	0	#
11	10	01010	ৌ	ম	২	K	i	+	\
12	11	01011	ক	য	৩	L	j	-	~
13	12	01100	খ	র	৪	M	k	*	^
14	13	01101	গ	ল	৫	N	l	/	|
15	14	01110	ন	শ	৬	O	m	=	$
16	15	01111	শ	ষ	৭	P	n	(	:
17	16	10000	স	হ	৮	Q	o	)	;
18	17	10001	ম	ঽ	৯	R	p	{	_
19	18	10010	জ	ড়	ঃ	S	q	}	`
20	29	10011	য়	ঢ়	ৎ	T	r	<	
21	20	10100	র	ঝ	ৃ	U	s	>	
22	21	10101	ট	ৄ		V	t	[	
23	22	10110	ল	ৢ		W	u	]	
24	23	10111	ড	ং		.	v	%	
25	24	11000	য	ঁ		?	w	&	
26	25	11001			Set-1				
27	26	11010			Set-2				
28	27	11011			Set-3				
29	28	11100			Set-4				
30	29	11101			Set-5				
31	30	11110			Set-6				
32	31	11111			Set-7				

Algorithm 1: Forward Mapping

Input: Normal string S,

Output: An encoded compressed string Sc

1: By adding set change,

2: Representing the string S by S`

3: By using the lookup table,

4: Representing the string S` by its corresponding 5 bits representation.

5: The representative string S' contain K bits.

if K%16=0 **then**
jump into end
end if
else if K%16!=0 **then**
Sum up m bits in the formation of the last 5 bit set code by taking every bit of the set
Code one by one and increasing in the same way until (K+m)%16=0
end else if
Represent the 16 bits from (K+m)/K to,
Compress string Sc by applying corresponding Unicode text.

Example 1:

Original Text (Input):

I am Bangladeshi

Set Representation: Set4 I Set5 space am space Set4 B Set5 angladeshi

Decimal Representation: 28 8 29 0 2 14 0 28 1 29 2 15 8 13 2 5 6 20 9 10

5 bit representation: 11100 01000 11101 00000 00010 01110 00000 11100 00001 …….

16 Bit representation: 1110001000111010 0000000100111000 ……

Compressed String: □腱 ·舵卯ĭ廷

Algorithm 2: Backward Mapping

Input: Compressed String Sc

Output: An encoded compressed string Sc

Representing the string Sc according to it's corresponding 16 bits value of binary number from the Unicode table.
Let, Sc contain K bits.
For determining 5 bit representation,
At first find $K' = int\ (K/5) * 5$.
Store the 1st set code of K' in C.
Take every 5 bit set code of the K' bit and compare it with C.
if C is not repeated **then**
which means there is an occurrence of set change so C is updated with the new 5 bit set code and compare this new C with the next set code, When C is repeated, the next bit stream will be Cropped
Including the bit code of the repeated C.
end if
From the remaining Kth bits stream, taking 5 bits and representing it according to the set of characters in the lookup table.
After excluding the number of set, the actual string can be generated.

Example 2:

Compressed String: ·腱 ·舵卯ĭ廷

16 Bit representation:

· --- 1110001000111010
腱 --- 0000000100111000
After dividing it by 16

The 5 bit Representation:

11100 --- 28 --- Set4
01000 --- 8 --- I
11101 --- 29 --- space
00000--- 0 ---- Set5 ……..

Decompressed String:

Set4 I Set5 space am space Set4 B Set5 angladeshi

Original Text: I am Bangladeshi

IV. ANALYTICAL ANALYSIS

In 5BC, each characters of 8 bits are represented via 5 bits hence 3 bits are saved. So the probable efficiency of 5BC is (8-5)/8= 37.5%. But after compressing it with Unicode scheme the best efficiency would be more than 60%.

We tried to develop a theoretical analysis to calculate the efficiency precisely. The parameters which are considered for analytical analysis are shown in following table.

TABLE II. PARAMETERS FOR ANALYTICAL EVALUATION

Parameter	Description
I	Original data
E	Encoded data
Ls	Length of the set
Li	Length of I
Ld	Length of E; $Ld = Li + Ls$
T	Total bytes for I
Te	Total bytes for E
η	Efficiency(Te/T)

Let Input data I has length Li. After encoding it becomes E and length become Ld

Where $Ld = L\,i + Ls$

Total bytes for I, $T = Li \times 8$

Total bytes for E, $Te = Ld \times 5$

$$= (Li + Ls) \times 5$$

$$Compression\ ratio = \frac{Compressed\ size}{Uncompressed\ size}$$

$$\eta = \frac{Te}{T}$$

$$= \frac{(Ld \times 5)}{(Li \times 8)}$$

$$= \frac{(Li + Ls) \times 5}{(Li \times 8)}$$

$$= \frac{(5Li + 5Ls)}{(Li \times 8)}$$

$$= \frac{5Li}{(Li \times 8)} + \frac{5Ls}{(Li \times 8)}$$

$$= \frac{5}{8} + \frac{5Ls}{(Li \times 8)}$$

$$= 0.625 + \frac{5Ls}{(Li \times 8)}$$

For usable compression, $\eta \leq 1$ then

$$=> 0.625 + \frac{5Ls}{(Li \times 8)} \leq 1$$

$$=> \frac{5Ls}{(Li \times 8)} \leq (1 - 0.625)$$

$$=> \frac{5Ls}{(Li \times 8)} \leq 0.375$$

$$=> 5Ls \leq (0.375 \times 8) \times Li$$

$$=> Ls \leq \frac{(0.375 \times 8) \times Li}{5}$$

$$=> \frac{Ls}{Li} \leq \frac{(0.375 \times 8)}{5}$$

$$\therefore \frac{Ls}{Li} \leq \frac{3}{5}$$

It is actually the proportion between the set length and original length. It demonstrates that this technique is properly usable until the length of set remains equal or less than 3/5th of the actual data length. And if so then we will get a positive efficiency. But if the ratio is greater than 3/5th the efficiency would be negative.

V. EXPERIMENTAL RESULT

Since 5BC produces a large Unicode string in a printable format. It can be further compressed after applying them in LZW and Huffman. We actually experiment 5 algorithms namely, 5BC, LZW, Huffman, 5BC+LZW and 5BC+Huffman. The experimental result shows the comparison between this 5 algorithms. We have done 3 kind of test comprising the Compression of files, Compression ratio and Compression time.

Compression of files: Here we can see the comparison between the compressions sizes achieved after applying the above mentioned compression techniques on the different datasets.

Compression ratio: Here we can see the comparison between the compressions ratios achieved after applying the above mentioned compression techniques on the different datasets

Compression time: Here we can see the comparison between the compressions times achieved after applying the above mentioned compression techniques on the different datasets.

A. Compression of files:

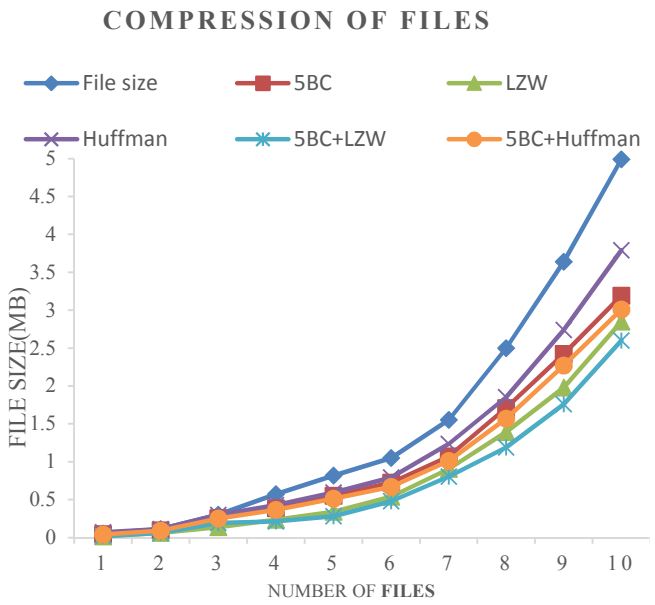


Fig. 1. Compression of file for different data set

From Fig. 1. The result shows that LZW provides the best result in terms of compression of different size file. But our proposed 5BC algorithm also provides very promising result. Our proposed technique performs far better than the well-known Huffman algorithm. And another fact that in terms of increasing data or increasing the file size our technique will converge the compression result of LZW. And if we combine these two techniques with 5BC, both the algorithm 5BC+Huffman as well as 5BC+LZW improves their performance.

B. *Compression Ratio:*

Fig. 2. Demonstrates the compression ratio for different datasets. The result shows that LZW provides very good compression ratio at the initial level when the file size is small but as the size of data is increasing, the efficiency rate will be poor Because LZW provides bad compression ratio for less dictionary size and as the character length increases the dictionary will not easily be made of. So performance reduced. But our proposed idea provides a constant efficiency

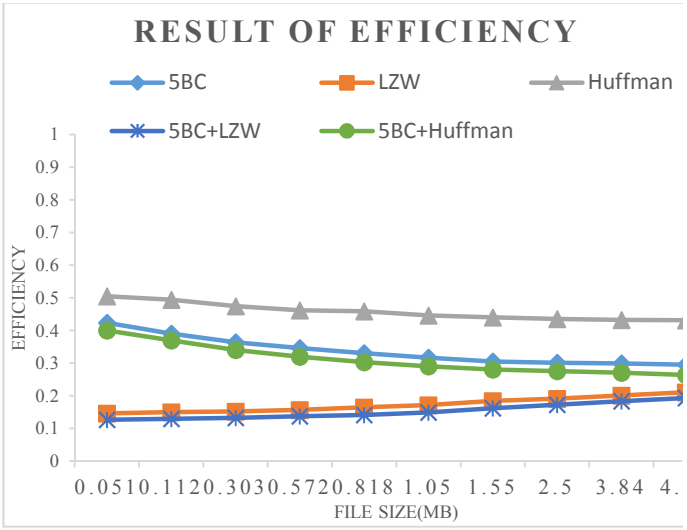


Fig. 2. Compression ratio for different data set

rate at any file size and the performance improves for the increasing number of datasets. It provides a quite better efficiency rate than Huffman. And at some point it will converge the efficiency rate of LZW also. The reason Huffman provides poor efficiency that Huffman technique derives the data from the probability or frequency of occurrence of the possible values in the source symbol [13]. So if the data set is quite large, the frequency generation would be more difficult and a large number of individual symbol will be created. As a result, it provides compression ratio which is average. Another important scenario is that after combining this both algorithm with 5BC, the performance of Huffman will be dramatically improved because this time it works with the compressed data by the 5BC which is much smaller than the original data. The performance of LZW is also improved.

C. *Compression time:*

Figure 3 provides the compression times for different datasets. The result shows that our proposed idea shows average result time when after combining it with LZW and Huffman the performance improves drastically. Though Huffman individually provides the time quite similar with our proposed idea.

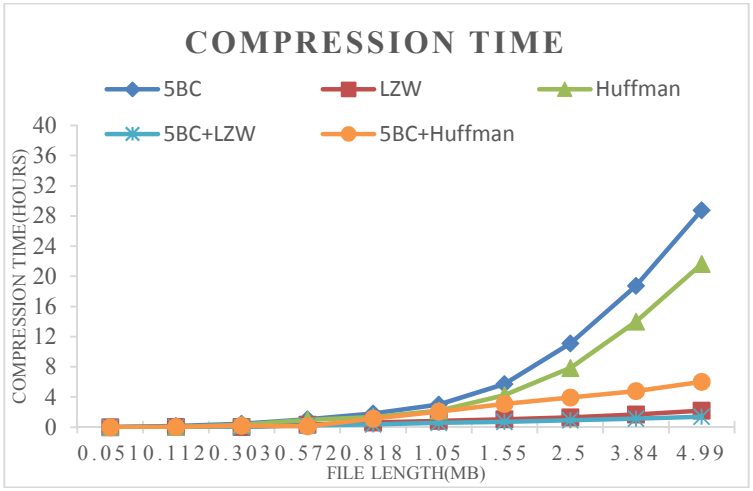


Fig. 3. Compression time for different data set

VI. CONCLUSION

In this research, we try to propose an efficient encoding and decoding technique for data compression which provides a fast and promising performance. This technique is a 5 bit compression scheme and it's character set i.e. dictionary is produced by the Zipf distribution. The proposed technique provides a promising efficiency rate which is more than 60%. This technique can be easily utilized compressing a huge amount of natural datasets. Both the mapping forward and backward make this compression mapping complete as it shows i.e. while decomposing we can achieve the original text. This technique can be implemented in building an innovative and efficient software which can be applied in communication of data, storage of data and database engineering. It can be utilized to parallel processing environment as well as load balancing technique to achieve promising encoding time.

REFERENCES

[1] Ashiq Mahmood, Tarique Latif, and K. M. Azharul Hasan, "An Efficient 6 bit Encoding Scheme for Printable Characters by table lookup (6BE)", International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 468-472, 2017.

[2] Md. Ashiq Mahmood, Tarique Latif, K. M. Azharul Hasan and riadul Islam, "A Feasible 6 Bit Text Database Compression Scheme with Character Encoding (6BC)", 2018 21st International Conference of Computer and Information Technology (ICCIT), pp. 1-6, 2018

[3] Nieves R. Brisaboa, Antonio Fariña Gonzalo Navarro "Lightweight natural language text compression", Information retrieval 10(1), 1-33, 2007.

[4] Wentian Li "Random Texts Exhibit Zipfs-Law-Like Word", Ieee Transactions On Informationtheory, Vol. 38, No. 6,November 1992

[5] Md. Abul Kalam Azad, Rezwana Sharmeen, Shabbir Ahmad and S. M. Kamruzzaman, "An Efficient Technique for Text Compression" The 1st International Conference on Information Management and Business, pp. 467-473, 2005.

[6] M. M. Kodabagi, M. V. Jerabandi, Nagaraj Gadagin, "Multilevel security and compression of text data using bit stuffing and huffman coding", 2015 International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), pp. 800 - 804, 2015.

[7] Muthukumar Murugesan, T. Ravichandran, "Evaluate Database Compression Performance and Parallel Backup," International Journal of Database Management Systems (IJDMS) Vol.5(4), 17-25, 2013.

[8] Vu H. Nguyen, Hien T. Nguyen, Hieu N. Duong and Vaclav Snasel " n-Gram-Based Text Compression, " Computational Intelligence and Neuroscience, Volume 2016.

[9] Senthil Shanmuga Sundaram, and Robert Lourdusamy, "A Comparative Study Of Text Compression Algorithms," International Journal of Wisdom Based Computing, Vol.1 (3), pp. 68-76, 2011

[10] A Carus, and A Mesut, "Fast text compression using multiple static dictionaries," Information Technology Journal, 1013-1021, 2010.

[11] Replacement And Bit Reductionn", SIPM, FCST, ITCA, WSE, ACSIT, CS & IT 06, pp. 407–416, 2012.

[12] N. Miladinovic, "Minimum-cost encoding of information. Coding with digits of unequal cost", 4th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services. TELSIKS'99 (Cat. No.99EX365), pp. 600 – 603, vol.2, 1999

[13] M. Sangeetha, P. Betty, G.S. Nanda Kumar, "A biometrie iris image compression using LZW and hybrid LZW coding algorithm," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), pp. 1-6, 2013.

[14] Wenjun Huang, Weimin Wang, Hui Xu, "A Lossless Data Compression Algorithm for Real-time Database,". 2006 6th World Congress on Intelligent Control and Automation, , pp. 6645-6648, 2006

[15] Ahmed Mokhtar, A. Mansour, Mona A. M. Fouad, "Dictionary based optimization for adaptive compression techniques", 2013 36th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 421 – 425, 2013

Design and Optimization of AlGaAs/InP Multi-junction Solar Cell

Md. Masud Rana
Department of EEE, Jashore University of Science and Technology
Jashore-7408, Bangladesh
mmrana.3012@gmail.com

S. Md. Firoz Khan
Department of EEE, Jashore University of Science and Technology
Jashore-7408, Bangladesh
firozkhanshisir@gmail.com

Md. Shafin Shafayat
Department of EEE, Jashore University of Science and Technology
Jashore-7408, Bangladesh
shafinshafayat@yahoo.com

***Abstract*— A structure for $Al_xGa_{1-x}As$/InP multi-junction solar cell is proposed by using numerical simulation and the simulation is done with the help of Analysis of Microelectronic and Photonic Structures (AMPS-1D) simulator. The thickness of p-layer of $Al_xGa_{1-x}As$ top cell is varied from 20 nm to 150 nm keeping the n-layer thickness at a constant value of 800 nm. The thickness of n-layer of InP bottom cell is varied from 100 nm to 7000 nm and p-layer thickness is fixed at 200 nm. The band gap of $Al_xGa_{1-x}As$ top cell absorber layer is varied from 1.42 eV to 1.79 eV. The highest power conversion efficiency (PCE) obtained is 35.643% (V_{oc}=2.385 V, J_{sc} = 15.789 mA/cm^2, FF = 0.88). The operating temperature is also varied from 0 °C to 70 °C. We observed that with the increase in operating temperature, the normalized open circuit voltage decreased almost linearly which shows better stability of this proposed multi-junction solar cell.**

Keywords*— *AlGaAs/InP, AMPS-1D simulation, fill factor (FF), power conversion efficiency (PCE), thermal stability.

I. INTRODUCTION

Solar energy may be used as an alternative source for the rapidly growing demand for energy [1]. As non-renewable sources are becoming scarce day by day, we have to use renewable energy. Fossil fuels like petrol, gasoline, coal etc. emit carbon-di-oxide (CO_2) when they are used. Carbon-di-oxide (CO_2) is one of the major factors of global warming. On the other hand, solar energy creates no pollution to our environment. To make the solar energy useable, a solar cell is needed. Single junction solar cell cannot utilize the photons that have wide-range of solar spectrum ranging from 300 nm to 2500 nm. For example, photons with wavelengths wider than 1100 nm cannot be absorbed by Si solar cell [2]. Addressing the aforementioned drawback of single junction solar cell, researchers have focused on multi-junction solar cell both theoretically and experimentally in the past few decades [3].

A thin-film solar cell is a device fabricated to transform light energy into electrical energy by using the photovoltaic effect. Thin film solar cells which are consisted of polycrystalline materials have better efficiency [4]. Now a days, this type of technology continues to gain popularity around the world. A multi-junction solar cell is consisted of two or more p-n junctions with dissimilar band gaps. A different part of the solar spectrum is absorbed by every junction in the solar cell. When photons hit the top of a single junction solar cell, they are absorbed if their energy is greater or equal to the band gap, otherwise, absorption will not happen. If the photon's energy is more than the band gap energy, the extra energy from the photon will be abandoned as heat. The concept of the multi-junction solar cell is based on the idea of using semiconductor materials with dissimilar band gaps to absorb different parts of the light spectrum. The thickness of the layers of the materials that make up the solar cell should be as skinny as possible so that the cost can be reduced. A well-designed multi-junction solar cell should have that ability to transform maximum light spectrum into current so that the external loads can be driven by it.

The solar cells consisted of silicon based materials are used all over the world which is up to 25% efficient [5]. The silicon source is ample in this world and good fabrication techniques have increased the availability of silicon based solar cells. However, silicon based solar cell efficiency is lower compared to the other types of solar cell such as III-V based solar cells[6]. The III–V zinc blende and wurtzite compound semiconductors materials are renowned for their tremendous performance as highly efficient solar cells. The III-V solar cells such as GaAs, InP, AlGaAs and GaInP are eminent for generating a higher efficiency more than 25% [7, 8].

$Al_xGa_{1-x}As$ is selected as the top cell absorber layer. It is a compound semiconductor which has a face-centered cubic crystal structure and a wide range of absorption coefficient ($\alpha \sim 10^5 cm^{-1}$) [9]. Indium phosphide (InP) is a binary semiconductor consisted of Indium and Phosphorus. It is a III-V semiconductor having a face-centered cubic crystal structure. $Al_xGa_{1-x}As$ and InP lattice constants are 5.6533+0.0078x Å and 5.868 Å respectively [10]. Based on the above criteria, it may be expected that a multi-junction solar cell made of $Al_xGa_{1-x}As$ and InP should be very efficient. In this paper, we propose a two-junction solar cell having $Al_xGa_{1-x}As$ as top cell and InP as bottom cell. Back to back connection is used to connect the two subcells.

II. MODELING AND SIMULATION

To understand the performance of $Al_xGa_{1-x}As$/InP multi-junction solar cell, the numerical simulation work is done by using Analysis of Microelectronic and Photonic Structures (AMPS-1D). It is a one dimensional device physics code that can be applied to any two terminal devices. The AMPS-1D software is capable of simulating the electrical features of diode, sensor, photo-diode, and photovoltaic device analysis. AMPS-1D can be used to generate I-V characteristics with or without illumination, photo diodes under voltage and light bias and so on. At present, AMPS-1D is used by more than 200 groups of researcher for solar cell design [11].

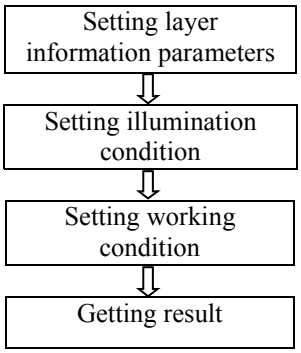


Fig. 1. General block diagram of simulation process of solar cell in AMPS-1D

The density of state (DOS) mode or the lifetime mode can be implemented for the simulation process by AMPS-1D [12]. Both modes are described in AMPS-1D manual [13]. Lifetime model is used in our simulation. The Poisson equation, the electron and hole continuity equations are three semiconductor device equations which are solved numerically by AMPS-1D [14]. The Poission equation, hole and electron continuity equations are given below respectively [15].

$$\frac{d}{dx}(-\in(x)\frac{d\psi}{dx}) = q[p(x)-n(x)+N_d^- - N_A^+ + p_t(x)-n_t(x)] \quad (1)$$

$$\frac{dp_n}{dt} = G_p - \frac{p_n - p_{n0}}{\tau_p} - p_n\mu_p\frac{d\xi}{dx} - \mu_p\xi\frac{dp_n}{dx} + D_p\frac{d^2p_n}{dx^2} \quad (2)$$

$$\frac{dn_p}{dt} = G_n - \frac{n_p - n_{p0}}{\tau_p} + n_p\mu_n\frac{d\xi}{dx} + \mu_n\xi\frac{dn_p}{dx} + D_n\frac{d^2n_p}{dx^2} \quad (3)$$

Here, Ψ represents electrostatic potential, n is denoted for free electron, p indicates free hole, n_t means trapped electron, p_t symbolizes trapped hole, N_D^- denotes ionized donor-like doping concentration, N_A^+ represents ionized acceptor-like doping concentration and ξ symbolizes electric field. These parameters are represented as a function of coordinate position x. G represents generation rate, D indicates diffusion coefficient, ε denotes permittivity and q symbolizes the charge of an electron. To calculate recombination current, Shockley–Read–Hall (SRH) model is used.

Fig. 1 explains the general block diagram of the simulation process of a solar cell in AMPS-1D. The basic structure of $Al_xGa_{1-x}As$/InP multi-junction solar cell is shown in Fig. 2. Standard Test Condition (STC) is considered in our simulation process. According to STC, cell temperature is 25 ^{0}C, irradiance is 1000 Wm^{-2} and air mass is 1.5.

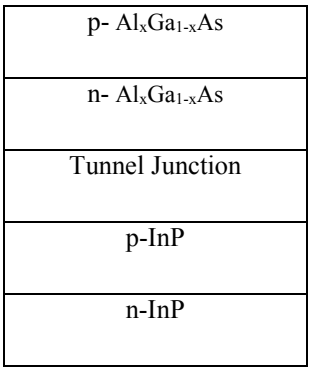


Fig. 2. Basic structure of $Al_xGa_{1-x}As$ /InP solar cell

More than fifty parameters can be varied in a specific solar cell model [16]. To work with fifty parameters is very abstruse and tiresome. So, the number of parameters should be reduced by keeping most of them as constant. It is very hard to select the accurate parameters that can be applied to simulate the various layer of a multi-junction solar cell model. For this simulation, the most important parameters are selected and summarized in Table I [17].

TABLE I. PARAMETERS USED IN SOLAR CELL SIMULATON

Layer properties	**p-AlGaAs**	**n-AlGaA**	**p-InP**	**n-InP**
Dielectric constant, ε	12.05	12.05	12.5	12.5
Electron mobility, μ_e [cm^2/Vs]	1100	1100	4000	4000
Hole mobility, μ_p [cm^2/Vs]	120	120	150	150
Carrier density, n or p [cm^{-2}]	p:1e17	n:1e17	p:5e16	n:5e16
Optical band gap, Eg [eV]	1.79	1.79	1.35	1.35
Effective density, N_c [cm^{-3}]	4.94e18	4.94e18	5.7e17	5.7e17
Effective density N_v [cm^{-3}]	1.74e19	1.74e19	1.1e19	1.1e19
Electron affinity [eV]	3.74	3.74	4.38	4.38

III. RESULTS AND DISCUSSION

Lattice constant has an effect on the band gap energy of semiconductor (cubic semiconductor). The band gap (E_g) is inversely proportional to the square of lattice constant [18]. The bandgap of $Al_xGa_{1-x}As$ absorber layer is varied from 1.42 eV to 1.79 eV according to the following Vegard's law [19].

$$E_{g,\,dir}(x) = 1.42\text{ eV} + x\,1.2475\text{ eV} \quad (4)$$

where $E_{g,dir}$ is the direct energy band gap of $Al_xGa_{1-x}As$ for $x < 0.45$ and x is Aluminum mole fraction in the $Al_xGa_{1-x}As$ alloy.

With the increase of x, the band gap is also increased. Effect of bandgap change on cell efficiency is depicted in Fig. 3. For x=0.3, E_g becomes 1.79 eV and the corresponding efficiency becomes 22.243%. It is the highest value of efficiency. So the band gap of the top cell is chosen as 1.79 eV.

A thin-film multi-junction solar cell is a second generation solar cell. It is created by depositing two or more thin layers. Currently, the thickness is varied from a few nanometers to micrometers which is much thinner than first generation solar cell. Thin-film solar cells are commercially used in recent times.

The thickness of a solar cell is a major parameter and choosing the optimum thickness is often an important factor [20]. One of the primary objectives of the present day's PV research work is making the cells thinner because diminishing the thickness reduces the creation time and fabrication cost. Lessening the semiconductor thickness of multi-junction solar cell decreases the semiconductor material expense as well as enhances the efficiency if the cell made with optimum thickness has a back-surface field [21]. When a solar cell absorbs light, an electron-hole pair is generated which must then be separated so that the charges can be put to use in a circuit. A large amount of energy will be reflected back if the cell is too thick [22]. Again, if the cell is too thin then it will not absorb most of the incident light as a thick one and the recombination rate will be increased. The increased recombination will result in decreased efficiency. Therefore, optimization in the thickness of the absorber layer is compulsory to gain the highest efficiency of solar cell.

The band gap of $Al_{0.3}Ga_{0.7}As$ is considered 1.79 eV as it will result in maximum efficiency. The p-layer thickness is varied from 20 nm to 150 nm and n-layer thickness is kept constant at 800 nm. The performance of the cell is simulated and outputs are represented in Fig. 4. From Fig. 4, it is observed that the thickness greater or less than 100 nm causes the efficiency to be reduced. So the p-layer thickness is fixed to be 100 nm.

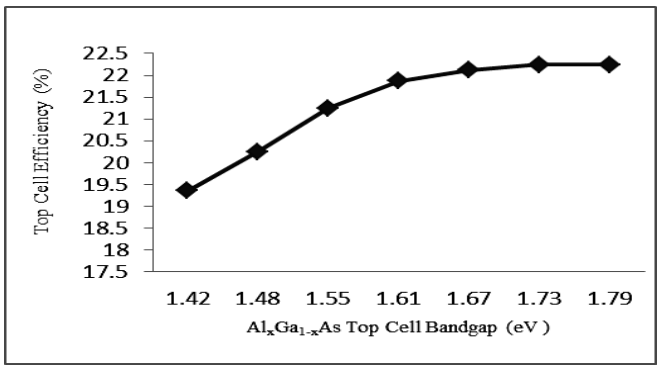


Fig. 3. Effect of top cell band gap change on efficiency

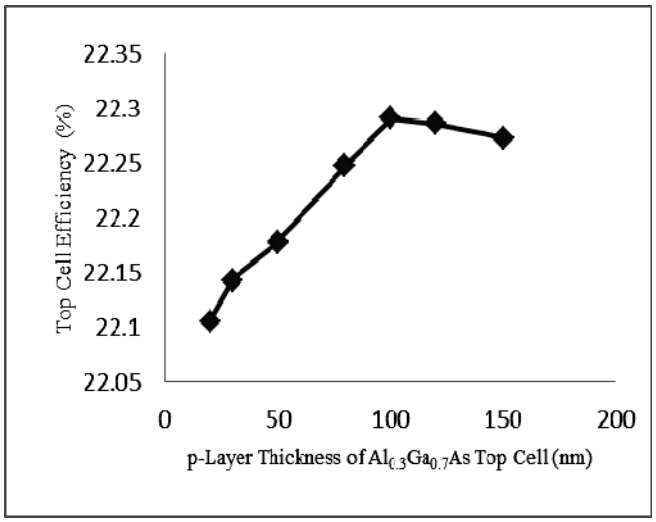


Fig. 4. Effect of p-layer thickness on efficiency for top cell

The band gap of InP is 1.35 eV. The thickness of n-layer is varied from 100 nm to 7000 nm and p-layer thickness is fixed at 200 nm. The performance of the cell is simulated and the results are shown in Fig. 5. If the absorber layer is thinner than 5000 nm it will cause a reduction in efficiency. On the other hand, the absorber layer thicker than 5000 nm causes no significant change in efficiency. So, the n-layer thickness is chosen as 5000 nm.

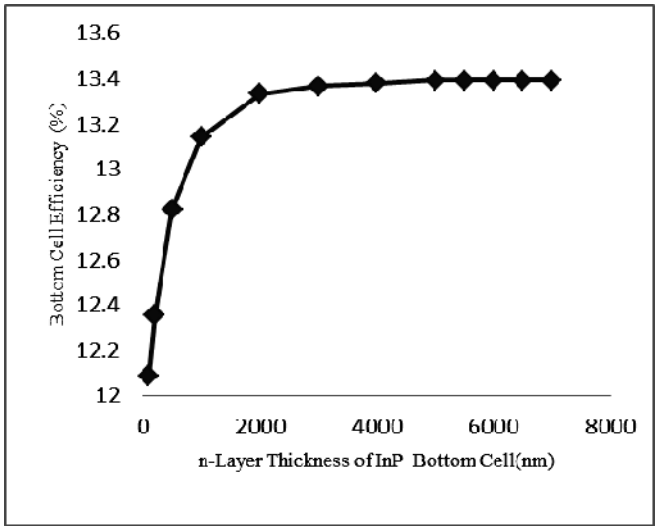


Fig. 5. Effect of n-layer thickness on efficiency for bottom cell

TABLE II. SIMULATED SUB CELLS RESULTS

Cells	Open Circuit Voltage (V)	Short Circuit Current (mA)	Fill Factor	Efficiency (%)
Top	1.539	15.783	0.916	22.247
Bottom	0.846	18.552	0.858	13.460

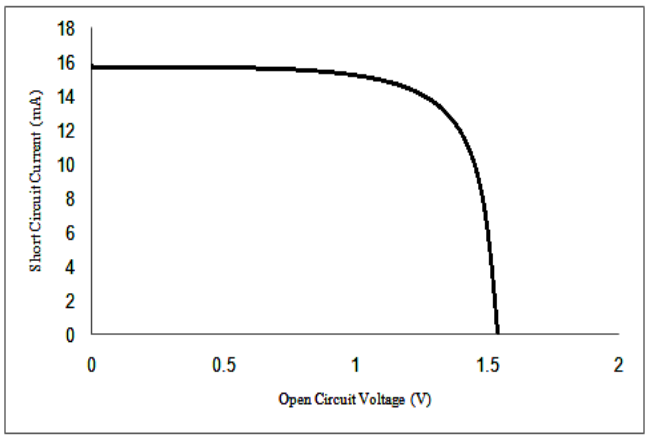


Fig. 6. I-V curve of $Al_{0.3}Ga_{0.7}As$ top cell

A perfect choice of the materials can give extended spectral coverage and thickness, bandgap optimization leads to an increase in overall power conversion efficiency [23]. In the baseline case from [17], the subcell efficiency of $Al_{0.3}Ga_{0.7}As$ and InP was 21.034% and 13.329% respectively and the overall efficiency of $Al_{0.3}Ga_{0.7}As$/ InP multijunction solar cell was 34.363%. Simulated subcell results are shown in Table II. The $Al_{0.3}Ga_{0.7}As$ top cell efficiency is 22.247% and the InP bottom cell efficiency is 13.460%. The I-V curve of $Al_{0.3}Ga_{0.7}As$ subcell is shown in Fig. 6. The I-V curve of InP subcell is depicted in Fig. 7. The I-V characteristic curve of $Al_{0.3}Ga_{0.7}As$/ InP multi-junction solar cell is represented in Fig. 8.

Open circuit voltage and fill factor decrease with the increase of operating temperature because bandgap slightly decreases and short circuit current increases with a decrease in the bandgap. As a result, with an increase in operating temperature efficiency decreases. If the normalized efficiency of a solar cell decreases linearly when the operating temperature is increased then the solar cell has better thermal stability [24]. The temperature is varied from 0 ^{0}C to 70 ^{0}C and the result is shown in Fig. 9. We observe that the normalized open circuit voltage decreases almost linearly that indicates the thermal stability of this proposed multi-junction solar cell.

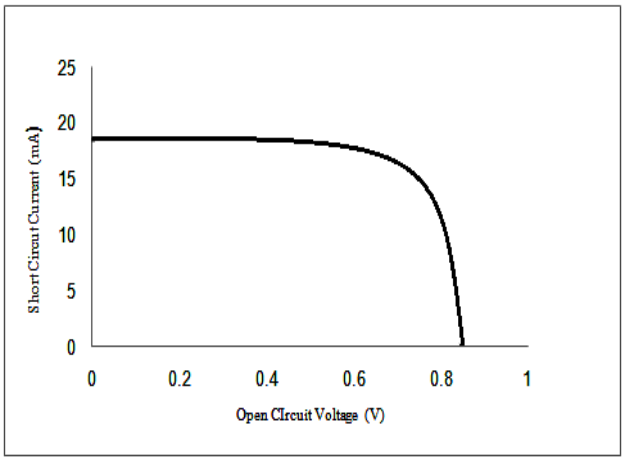


Fig. 7. I-V curve of InP bottom cell

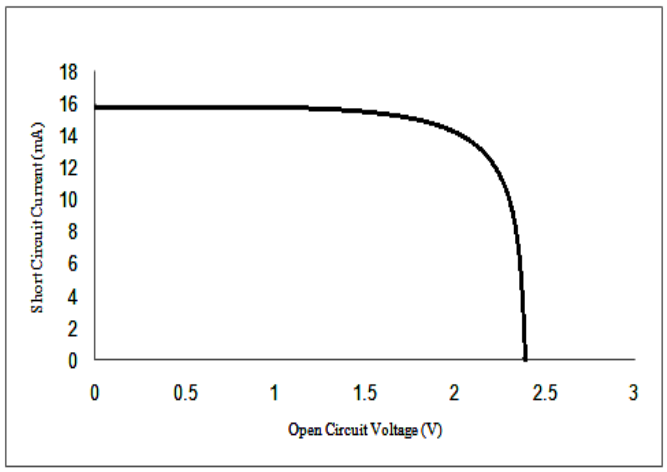


Fig. 8. I-V curve of $Al_{0.3}Ga_{0.7}As$/ InP multi-junction solar cell

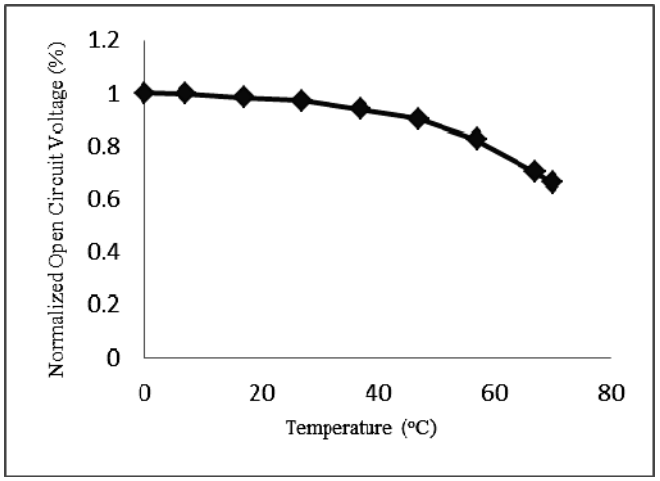


Fig. 9. Temperature effect on normalized open circuit voltage

IV. Conclusion

In this research work, we have designed and optimized a relatively thin $Al_{0.3}Ga_{0.7}As$/InP multi-junction solar cell. The simulated performance parameters of $Al_{0.3}Ga_{0.7}As$/InP multi- junction solar cell are V_{oc}=2.385 Volts, J_{sc}=15.789 mA/cm^2 and FF=0.88. The power conversion efficiency of the multi-junction solar cell at 298 K is 35.643%. The thermal stability of the proposed solar cell is also better. The proposed multi-junction solar cell is depicted in Fig. 10.

p-$Al_{0.3}Ga_{0.7}As$(Thickness= 100nm)
n-$Al_{0.3}Ga_{0.7}As$ (Thickness=800nm)
Tunnel Junction
p-InP(Thickness= 200nm)
n-InP(Thickness= 5000nm)

Fig. 10. Structure of optimized $Al_{0.3}Ga_{0.7}As$/ InP multi-junction solar cell

ACKNOWLEDGMENT

We would like to thank Dr. Fonash's group of Pennsylvania State University (PSU) who developed AMPS-1D program.

REFERENCES

[1] Khan, N., Rahman, K., Haque, F., Dhar, N., Islam, M., Akhtaruzzaman, M., Sopian, K. and Amin, N. (2014). "Design Optimization CdTe Thin Film Solar Cells from Numerical Analysis", *8th International Conference on Electrical and Computer Engineering*. Dhaka: IEEE, pp.508-511.

[2] J. Duffie and W. Beckman, *Solar Engineering of Thermal Processes*, 3rd ed. New Jersey: John Wiley & Sons, 2006.

[3] J. M. Olson, D. J. Friedman, S. R. Kurtz, *Handbook of Photovoltaic Science and Engineering* ,Chichester, West Sussex, U.K. ,Wiley, 2011.

[4] M. Hossain, N. Amin, M. Matin and M. Aliyu, "A Numerical Study on the prospects of high efficiency ultrathin ZnxCd1-xS/CdTe Solar Cell", *Chalcogenide Letters*, vol. 8, no. 3, pp. 263 – 272, 2011.

[5] Masuko, K., Shigematsu, M., Hashiguchi, T., Fujishima, D., Kai, M., Yoshimura, N., Yamaguchi, T., Ichihashi, Y., Mishima, T., Matsubara, N., Yamanishi, T., Takahama, T., Taguchi, M., Maruyama, E. and Okamoto, S. (2014). Achievement of More Than 25% Conversion Efficiency With Crystalline Silicon Heterojunction Solar Cell. *IEEE Journal of Photovoltaics*, 4(6), pp.1433-1435.

[6] Sumaryada, T., Rohaeni, S., Syafutra, H., Irzaman, Maddu, A., Hardhienata, H., Alatas, H. and Damayanti, N. (2016). Modeling the Thermal Response of Al0.3Ga0.7As/GaAs/Ge Multijunction Solar Cells. In: *International Conference on Energy Sciences (ICES 2016)*. Bandung, Indonesia: Journal of Physics, pp.1-5.

[7] Friedman, D., Kurtz, S., Kibbler, A., Bertness, K., Kramer, C., Matson, R., Arent, D. and Olson, J. (1992). Anisotropic Electrical Transport and Surface Microstructure in GaInP2. *MRS Proceedings*, 280, pp.9-12.

[8] King, R., Law, D., Edmondson, K., Fetzer, C., Kinsey, G., Yoon, H., Sherif, R. and Karam, N. (2007). 40% efficient metamorphic GaInP⁄GaInAs⁄Ge multijunction solar cells. *Applied Physics Letters*, 90(18), p.183516.

[9] W. Callister, *Materials Science and Engineering*, 7th ed. New Delhi, India: Willey India (P) Ltd., 2007, p. 747.

[10] S. Adachi, "GaAs, AlAs, and A_xGa_{1-x} As: Material parameters for use in research and device applications", *Journal of Applied Physics*, vol. 58, no. 3, pp. 1-29, 1985.

[11] H. Zhu, A. Kalkan, J. Hou and S. Fonash, in *AIP Conf. Proc*, Osaka, Japan, 1999.

[12] J. Ray, C. Panchal, M. Desai and U. Trivedi, "Simulation of CIGS Thin Flim Solar Cells Using AMPS-1D", *Journal of Nano- and Electronic Physics*, vol. 3, no. 1, pp. 747-754, 2011.

[13] S.J. Fonash, *A manual for One-Dimensional Device Simulation Program for the Analysis of Microelectronic and Photonic Structures (AMPS-1D)*, (The Center for Nanotechnology Education and Utilization, The Pennsylvania State University,University Park, PA 16802).

[14] M. Burgelman, J. Verschraegen, S. Degrave and P. Nollet, "Modeling thin-film PV devices", *Progress in Photovoltaics: Research and Applications*, vol. 12, no. 23, pp. 143-153, 2004.

[15] S. Sze, *Physics of Semiconductor Devices*, 2nd ed. New York, USA: John Wiley & Sons, 1981, p. 51.

[16] Burgelman, M., Verschraegen, J., Degrave, S. and Nollet, P. (2004). Modeling thin-film PV devices. *Progress in Photovoltaics: Research and Applications*, 12(23), pp.143-153.

[17] E. Reineri and T. Yu, "*AlGaAs/InP High-Efficiency Solar Cell*." M.S. Thesis,San José State University, California, USA, 2010.

[18] R. Dalven, "Empirical Relation between Energy Gap and Lattice Constant in Cubic Semiconductors", *Physical Review B*, vol. 8, no. 12, pp. 6033-6034, 1973.

[19] S. Adachi, *GaAs and Related Materials Bulk Semiconducting and Superlattice Properties*, 1st ed. Singapore: World Scientific, 1994, p. 223.

[20] H. Firoozi and M. Imanieh, "Improvement the Efficiency CIGS Thin Film Solar Cells by Changing the Thickness Layers", *International Journal Of Engineering And Computer Science*, vol. 6, pp. 22055-22061, 2017.

[21] C. Sah, "Thickness dependence of solar cell performance", *Elsevier*, vol. 25, no. 9, pp. 960-962, 1982.

[22] M. Green, *Solar Cells - Operating Principles Technology and Systems Applications*, 1st ed. New Jersey, USA: Prentice-Hall, Inc., 1982, p. 211.

[23] D. Benmoussa, B. Hassane and H. Abderrachid, "Optimization of a tandem solar cell GaAs / Ge using AMPS-1D", *Journal of Energy Technologies and Policy*, vol. 3, no. 3, pp. 9-13, 2013.

[24] Amin, N., Matin, M., Aliyu, M., Alghoul, M., Karim, M. and Sopian, K. (2010). Prospects of Back Surface Field Effect in Ultra-Thin High-Efficiency CdS/CdTe Solar Cells from Numerical Modeling. *International Journal of Photoenergy*, 2010, pp.1-8.

Cows Urine is a prominent source of electrolyte for Microbial Fuel Cell

Md. Abdul Halim
Department of Chemical Engineering
(Assistant Professor)
Jashore University of Science and Technology
(Faculty of Engineering and Technology)
Jashore, Bangladesh
halim.che09@gmail.com

Md Ibrahim
Department of Chemical Engineering
(Student)
Jashore University of Science and Technology
(Faculty of Engineering and Technology)
Jashore, Bangladesh
ibrahim250395@gmail.com

Md. Rasel Molla
Department of Chemical Engineering
(Student)
Jashore University of Science and Technology
(Faculty of Engineering and Technology)
Jashore, Bangladesh
raselchejust@gmail.com

Rituparna Kundu
Department of Chemical Engineering
(Student)
Jashore University of Science and Technology
(Faculty of Engineering and Technology)
Jashore, Bangladesh
kritu.bristi@gmail.com

Md. Bokhari Hasan
Department of Chemical Engineering
(Student)
Jashore University of Science and Technology
(Faculty of Engineering and Technology)
Jashore, Bangladesh
bokhari8355@gmail.com

Md. Aminul Islam
Department of Chemical Engineering
(Student)
Jashore University of Science and Technology
(Faculty of Engineering and Technology)
Jashore, Bangladesh
omy.jstu@gmail.com

***Abstract*— Demand of energy is increasing in the whole part of the world, it is estimated that in near future a significant energy crisis will be occurred. Furthermore, fossil fuels play a noticeable role for environmental pollutions. So, an alternative clean and environmental amiable renewable energy source is the most significant demand for the modern scientists. Bio-fuel cell especially microbial fuel cell can partially eliminate this crisis by producing electricity from biodegradable waste. Considering this aspect we produced double chamber MFCs where Zinc and copper plates are used as anode and cathode materials. In this study urine and fish waste mixed waste water are used as a source of substrate, aerobic bacteria helps to decompose it, in anode chamber where oxygen amplifies the reduction reaction in cathode chamber. To find out the effects of urine and fish waste on the performance of voltage, current and power we used digital multimeter and Hantek 365A data logger. Urine mixed waste water gives the maximum voltage 1.02 V, current 1.67 mA and power 1.7034 mW, whereas, for fish mixed waste water these values were 0.968 V, 1.365 mA and 1.3213 mW respectively. The whole operation we done batch wise, initial pH 8 and cell volume was 5 liter.**

Keywords— MFC, Urine, Fish, Electrode, Renewable.

I. INTRODUCTION

Energy is used to do everything from walking to sending astronauts into space. Energy is essential to all living organisms. The sun is the source of all the energy available on Earth. Energy choices and decisions impact Earth's natural systems in ways we may not be aware of it, so it is essential that we choose our energy sources cautiously. The authentic value of energy is more than just dollars and cents; there are significant economic, political and social factors and consequences to consider as well [1]. The importance of energy sources is increasing day by day and we need to conserve energy that will continue over the next few decades. Nonrenewable energy sources thrust into fossil fuels that come from beneath the ground and take thousands of years to form. Renewable energy sources are developing quickly and can supply a region with its long-term energy needs far into the future. Fossil fuels contain energy held millions of years ago from sunlight by living organisms. The energy in fossil fuels such as oil, natural gas, and coal comes from energy that producers captured from sunlight long ago. Energy stored in these fuels is liberated by burning them, which also releases carbon dioxide into the ambience [2].

Renewable and nonrenewable energy sources are mainly used as an elementary source of energy further it used to produce electricity. Nuclear energy is produced from uranium, a nonrenewable energy source whose atoms are split through a process called nuclear fission to create heat and, eventually, electricity. Microbial fuel cell (MFC) is a bio-electrochemical system [3] that drives an electric current by using bacteria and mimicking bacterial interactions found in nature. The first MFCs, exhibited in the early 20th century, used a mediator: a chemical that transfers electrons from the bacteria in the cell to the anode. Around the year 1970s unmediated MFCs emerged in the history of MFC, where bacteria electrochemically active redox proteins can transfer electrons directly to the anode [4]. In the 21st century MFCs inaugurated to find a mercantile use in wastewater treatment [5].

Microbial fuel cells have manufactured a pregnant posterior for human beings and the planet we live on. The research done so far and with the infliction that we have currently will help the reduction of our need to use fossil fuels and allow us to reach energy from wastes products, bacteria ridden water, enable us to use prior salted water to drink, farm with, this will greatly benefit 3rd world countries and areas plagued with drought where there is no electricity. Microbial Fuel Cell (MFC) may be such a renewable energy source for the coming decades. Microbial Fuel Cell (MFC) may be such a renewable energy source for the next decades.

In this research we compared the effect of additives (urine and fish waste) on the yield (current, voltage and power) of microbial fuel cell. For both cases the experiment was done in batch wise and length of time was nine days.

II. MATERIALS AND METHOD

A. Cell Construction

Two chambered MFC was constructed with batch operation. Both anodic and cathodic chambers were made up of glass (volume) with two openings, one for addition or removal of sample and other joining for the other chamber. Each chamber had 5000 ml working volume. The two chambers were connected using an agar salt bridge with a length and diameter of 5 cm and 0.5 cm respectively. The joints were covered by M-seal to prevent the leakage. Two similar sized zinc (Zn) and copper (Cu) electrodes were used as anode and cathode. Anode and cathode were joined to the copper wires, which in turn were commenced. The anodic chamber contained municipal wastewater which had following composition (per liter): 0.02g NH_4Cl, 0.015g $CaCl_2.2H_2O$, 0.33g KCl, 0.33g NaCl, 0.015g $MgSO_4.7H_2O$, 0.001g $CuSO_4.5H_2O$, 0.001g $ZnSO_4.7H_2O$, 0.001g $MnSO_4$, 0.1N Phosphate Buffer (pH 7), 5% Dextrose. The cathodic chamber contained 0.1M potassium sulphate prepared in 5000mL volume. Here also used microbe which was collected from micro-organism sources. This operation is shown in is fig. 1.



Fig.-1. Schematic diagram of two-chamber microbial fuel cells

B. Substrate Collection

We used municipal waste water and fish waste as substrate. The municipal waste water was collected from the municipal waste storage and the fish waste was collected from fish market.



Fig-2. Source of substrates

C. Salt bridge preparation

For the preparation of salt bridge we had to pass two steps. Firstly we had made 0.1M salt solution with 5gm agar. Finely ground salt such as table salt dissolves much faster than coarsely ground salt (rock salt). If the solution is warms the agar will begin to melt, warm the agar until it becomes liquid. We heated much to mix well to dissolve the salt in the agar. After mixing well we removed it from the heat. Then some pieces of gauge cloths were kept into the solution and then kept it for 2-3 hours. Then that we had prepared the agar/salt electrolyte solution we filled the salt bridge with it. A short PVC pipe about 4-5cm long was filled with the gauge which absorbed the agar-salt solutions.

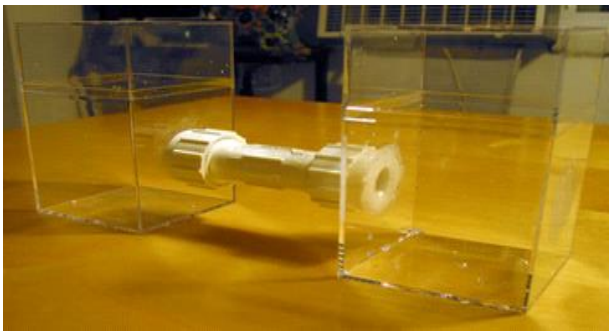


Fig-3. Salt bridge preparation

D. Microbe

We obtained our algae culture from natural sources. We went to a local lake; a swamp is good to have one. We were looking for bottom feeders which indicate a rich organic base. We mixed the mud with a portion of cow waste and also mixed a little portion of sugar and then mixed them very well. After that we kept the mixer in an anaerobic condition.

E. Constructing the Electrodes

As electrode (anode and cathode) we used Copper (Cu), zinc (Zn) and carbon felt. We used Zn as anode and Cu as cathode in both electrochemical cells. Then we used Cu wire for connecting the both electrode. Wire layers should be placed on the outside to contain the carbon paper and maintain integrity of the electrodes in the mud or sand.

F. Data Collection ad Data Analysis

Data was collected through manual and automatic in both systems. We used a digital multi-meter and measured voltage and current at fixed interval. For automatic data collection, we used Bluetooth/USB Data Logger named Hantek 365A which was connected to a computer with the B-Type Plug of USB cable to Hantek365's USB port provided by the manufacturer. Based on the recorded voltage, the current and powers were calculated using the following equations:

Current

Current was calculated using Ohms law:

$V=IR$

Current density:

$I_{anode} = I/A = V/(A*R)$

Where, V=Voltage; I=Current; R=External resistance; I_{anode}=Current density; A=Area of the anode electrode

III. RESULT AND DISCUSSION

A. Result

For this experiment we constructed rectangular shape double chamber microbial fuel cell. The operating condition for whole process was pH-8, electrode distance 5cm and operating temperature ranges from 25 to 30 degree centigrade. Here we mixed 10% (w/v) both urine and fish waste with municipal waste water and electrode pair was Zinc and Cupper.

Fig-4 compares the current obtained from urine and fish waste mixed waste water. From here it is clear that the obtained value from urine mixed waste water is higher than fish waste mixed waste water. Maximum current obtained from urine mixed waste water was 1.67 mA whereas, for Fish waste result was 1.365 mA, both results was obtained in the first day of operation. Then the current decrease sharply, after 3 days this rate became slow. After nine day operation the value was0.386 mA and 0.247 mA respectively for urine and fish waste mixed waste water.

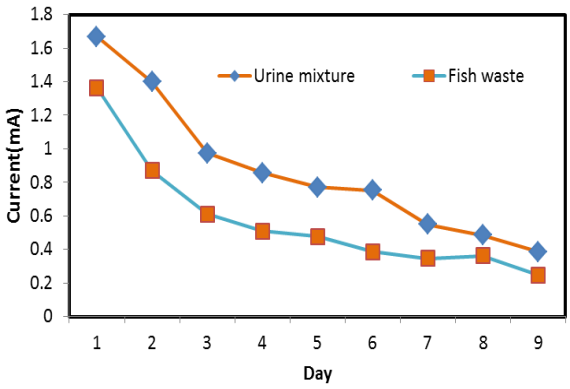


Fig-4 Change of current with time

Our findings is higher than Kamau et al [6], where he used cow dung and got a maximum current 1.156 mA. In fig-5 total current in nine days operation was 7.84 mA and 5.178 mA for waste water and fish waste respectively.

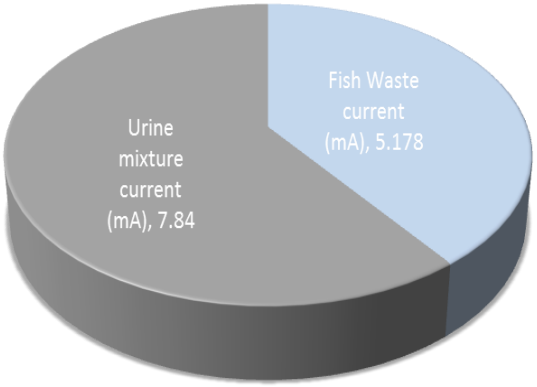


Fig-5 Total current of waste water and fish waste

Figure-6 shows the variation of the voltage for two additives urine and fish waste with municipal waste water. In first day of operation maximum voltage for urine mixture was 1.02 V, which was 5.1% higher than fish waste mixed waste water (0.968 V). Voltage decreased from 1.02 V to 0. 545 V and 0.968 V to 0.602 V for waste water with urine and fish waste respectively from first day to five day then declined slowly for both two. Finally after nine days operation the voltage (0.49 V, and 0.388 V) were reached for both sample, this was done because for the formation of biofilm on the anode surface. Our finding is much higher than Simeon and Raji 2016 [7], where they got 731 mV by using urine in soil based MFC. Fig-7 shows that in 9 days operation total voltage produced 6.871 V and 5.851 V for urine mixed and fish waste mixed waste water respectively.

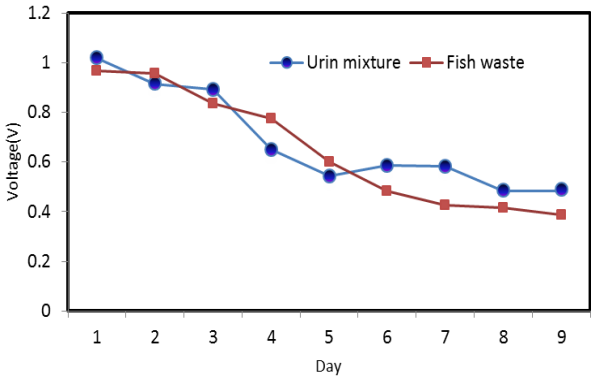


Fig-6 Change of voltage with time

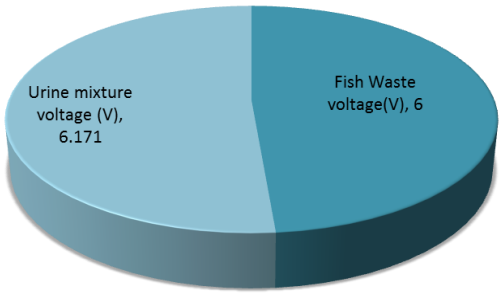


Fig-7 Total voltage of waste water and fish waste

The figure-8 below compares the power obtained for waste water mixed with (10% w/v) urine and (10% w/v) fish waste. At first day the obtained value was 1.7034 mW for urine mixture and 1.32132 mW for fish waste mixture. Then the power decreased for both connections, after three days operation the yield for both cases decreased slowly. These findings are higher than Kamau et al [6], where he got 0.0045 mW using cow dung. In fig-9 the total power were 6.017362 mW for urine mixed waste water and 3.929824 mW for fish waste mixed waste water.

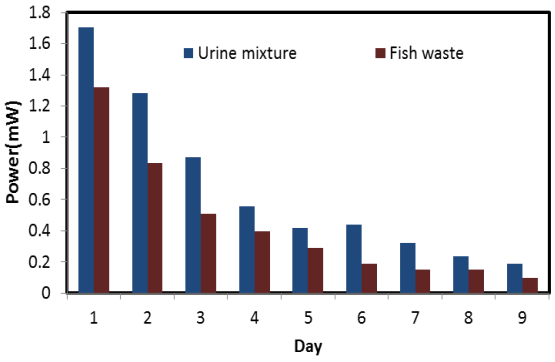


Fig-8 Change of power for waste water & fish waste



Fig-9 Total power for nine days

B. *Discussion*

In our experiment we find that cow urine is better than fish waste for generating power from microbial fuel cell. This experiment can be used for society when we store the current in a capacitor. Our result is higher than Ogugbue et.al 2015 [8], where they got maximum power 0.355 mW by using copper and carbon electrodes but we used zinc and copper electrodes with same electrode distance, again it was higher than both Kamau et al [6] and Simeon and Raji 2016[7].

IV. CONCLUSION

Urine as well as fish waste are prominent activator for waste water treatment in microbial fuel cell. The process we observed for 9 days, urine mixed waste water gives 5.12% higher voltage, current 34.1% and power 34.68% than fish mixed waste water. All these cases urine mixed waste water gives higher yield than fish mixed waste water, because some vitamins (citric, succinic acid, carbolic acid), minerals (lactose, calcium salts) ions such as Na, N, Mg, Fe, Si, Cl, Mg, P, K, S and enzymes are present in cow's urine [9], again fish waste contain average 2.83% nitrogen (N), 2.54% phosphorus (P), 0.10% potassium (K), 6.99% calcium (Ca), and 0.53% magnesium (Mg) on a dry-weight basis [10]. So that urine is a better electrolyte than fish waste.

ACKNOWLEDGMENT *(Heading 5)*

We would be very much grateful to the authority department of Chemical Engineering, Jashore University of Science and Technology (JUST). We are also grateful to the officers and staff of the central laboratory of JUST, who helped us in whole experiment.

REFERENCES

[1] https://climateinterpreter.org/content/why-energy-important

[2] https://climateinterpreter.org/content/human-energy-use-and-consumption

[3] A. Ebrahimi, D. G. Najafpour, D. Y. Kebria, Performance of microbial desalination cell for salt removal and energy generation using different catholyte solutions, Desalination, 432, pp. 1–9, 2018, doi:10.1016/j.desal.2018.01.002.

[4] S. P. S. Badwal, S. S. Giddey, C. Munnings, A. I. Bhatt and A. F. Hollenkamp, Emerging electrochemical energy conversion and storage technologies, Frontiers in Chemistry, 2, 79, 2014, doi:10.3389/fchem.2014.00079. PMC 4174133. PMID 25309898.

[5] "MFC Pilot plant at the Fosters Brewery". Archived from the original on 2013-04-15. Retrieved 2013-03-09.

[6] J. M. Kamau, D. N. Mbui, J. M. Mwaniki, F. B. Mwaura and G. N. Kamau, Influence of External Resistors on Power, Current and Power Density, J Thermodyn Catal, 8, pp.1, 2017, DOI: 10.4172/2157-7544.1000182.

[7] M. I. Simeon, and A. O. Raji, Experimental utilization of urine to recharge soil microbial fuel cell for constant power generation, Journal of Engineering and Environmental Sciences, 1(1), pp. 129-135, 2016.

[8] C. J. Ogugbue, E. E. Ebode and S. Leera, Electricity Generation From Swine Wastewater Using Microbial Fuel Cell, Journal of Ecological Engineering, 16 (5), pp. 26-33, 2015, DOI: 10.12911/22998993/60450.

[9] N. K. Jain, V. B. Gupta, R. Garg and N. Silawat, Efficacy of cow urine therapy on various cancer patients in Mandsaur District, IndiaeA survey. Int J Green Pharm, 4, pp. 29-35, 2010.

[10] S. J. Naylor , R. D. Moccia and G. M. Durant, The Chemical Composition of Settleable Solid Fish Waste (Manure) from Commercial Rainbow Trout Farms in Ontario, Canada, N AM J AQUACULT, 61:1, pp. 21-26,1999.

An Intelligent Street Light System Based on Piezoelectricity Generator with Noticeable Zebra Crossing Point and Bus Stand

A. Z. M. Tahmidul Kabir
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
tahmidulkabir@gmail.com

Nirmol Deb Nath
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
debnathnirmol@gmail.com

Fukrul Hasan
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
dipta.hasan555@gmail.com

Mirza Tanvir Sadik
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
mirzasadik4032@gmail.com

Sakhwat Hossian Lisun
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka-1229, Bangladesh
sakhwatlisun306@gmail.com

Abstract— **The world's total population and the total number of vehicles are increasing day by day. The consumption rate of total power is also rising. Sometimes it becomes a big problem to provide the nation with enough power in underdeveloped or a developing country only because of the high-power consumption by the consumers. It is also observed that the road line continuously consumes a huge amount of electricity, and this amount of electricity can be saved if a controlled process can be operated on the road. In order to save the waste of electricity from the road and to operate street light in a controlled manner, an intelligent lighting system is introduced with a Piezo electricity generator. With the appearance of vehicles or people on the road, the light intensity can be changed in this system. This system also introduces the zebra crossing light to warn a driver at night when a person passes a road using the zebra crossing. This warning from the zebra crossing light can save a person from an accident. Another special feature of this system is the Piezoelectric generator, which provides extra electricity. In this system, several components are used to operate the system properly. Light Dependent Resistor and Ultrasonic sensor are used to control street lights by sensing night and day or by detecting vehicles or pedestrians at night. The lights at the zebra crossing point are controlled by a Passive Infrared Resistor. All these features of this system will provide an uninterrupted supply of electricity as this system can make a major contribution in saving a country's most needed energy.**

Keywords— Ultrasonic sensor, Light Dependent Resistor, Piezoelectric generator, PIR sensor.

I. INTRODUCTION

With the rise of civilization, every day counts for new progress in different aspects of life. At the beginning of civilization, there were fewer streets for lesser population than now. However, with the population growth, the demands and facilities used by consumers have accelerated. As a result, to fulfill the demand of this increased population, better transportation facilities need to be provided along with the assurance of safer road. The economic growth of a country is highly dependent on power supply, communication and transportation system.

However, as Bangladesh is a developing country, the street facilities are not properly managed and executed. Among many other reasons, poor lighting system is one of the major concerns of the government. It has been also observed that unwanted situation like robbery, snatching, hijacking, sexual harassments etc. are also occurring daily where the street-light facilities are not up to the mark [1].

In Dhaka city (Capital of Bangladesh) about 420MW power is wasted every month for not turning on/OFF the street light on the right time. It is observed that street lights are turned ON usually one or two hours before sunset and are turned OFF
one or two hours after sunrise. This is happening because of a very poor monitoring system and some other issues. As a result, 34 lakh (BDT) is wasted every month. These wasted powers could be used to fulfill the demand of 4,000 non-urban households [2].
In the conventional system, street light is controlled from the control station and they keep turned ON the street light every day from 6 PM to 7 AM [3].

In Bangladesh, people face snatching, robbery, hijacking, sexual harassment every day. These types of incidents

happen at a place where the street light is insufficient. To overcome the insufficiency of light we have designed a system which will provide maximum light when a single person is available on the street or at a bus station. This system will work automatically to turn on or OFF the light.

Bangladesh is a developing country and its development depends on power supply but there is huge lacking of power supply in this country according to the total demand. Total production of power is not sufficient in Bangladesh but a huge amount of power is wasting every day for street light. This wastage can be reduced, if the street light can be controlled through an efficient system. Our smart street light system with Piezoelectricity generator will control those lights by turning it on or OFF at an exact time. This system can operate the lights in several modes. The street lights can be kept turned ON or OFF and also its intensity can be controlled according to the appearance of the people or the vehicles with self-light or no light.

This system will also work for the pedestrians using zebra crossing. When a person will walk through the zebra crossing line, the lights beside the line will turn on automatically and will warn the driver at night. This system works in such a manner by which a huge amount of electricity can be saved from the street light consumption and the Piezoelectricity generator produces electricity by itself with the help of the running vehicles and stores it in a battery. Through this system another source of electricity is added for the street light and if there is a load shedding or other issue of absence of electricity from the national grid, then the street light can operate itself by the stored energy. This system will also provide a ceaseless supply of electricity and the roads will be safer than the existing ones.

This system will detect the day-night situation and will turn on/OFF the light automatically. This system will detect the darkness even in daytime. This system will work even in the rainy days when sufficient sunlight is unavailable. Another problem is found that street lights are not operational. It is a common scenario where several numbers of street lights are found damaged or malfunctioned in a particular place. Our system will detect the damaged light immediately and this information can be monitored from a control station so that the damaged light can be replaced as soon as possible. Combining all these features our system will not only work for reducing the power consumption of street light but also ensure a safer life and efficient monitoring system of the street-light.

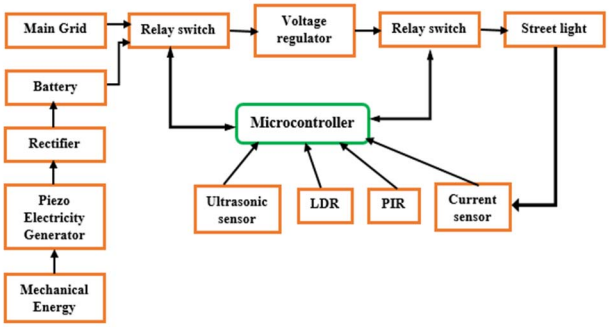
II. INTEGRATED EQUIPMENT

Various types of components are embedded in this system to make the project effective. The particular combination of all these equipment makes this system comparatively accurate and user-friendly. The main characteristics of these components are described below.

No. of component	No. of unit	Features of these components for this project
Arduino Mega	1	• Main processing unit. • Both analog and digital pins are used. • 5V is used to operate all of the components. • To power up the Arduino 9V is used in this project.
Piezo electric generator	5	• To convert mechanical energy to electrical energy.
Light-dependent resistor	6	• 1 LDR sensor is used to detect day or night period. • Rest of the LDR sensor detects the light of the vehicle.
Ultrasonic sensor	5	• To detect objects. • Person or vehicle both can be detected.
Current sensor	5	• To check malfunction in every individual pole.
Relay switch	1	• For switching purpose.
Voltage regulator	1	• To operate the device in fixed voltage.
Rectifier	1	• Convert AC signal to DC signal.
Led light	15	• Used as a street light.
PIR sensor	2	• To detect human.

III. SYSTEM DESCRIPTION

A. Block diagram



The above figure shows the block diagram of the system. In this diagram, all the components used in this system are shown and this diagram mainly describes the overall system in a single figure. In this system, several components are used as inputs and outputs. Main grid and a battery are the two-power sources of this system. The battery reserves its power from the Piezoelectricity generator. The microcontroller is the main control unit of this system.

Several components are connected with the microcontroller. Relay switches, ultrasonic sensor, Light Dependent Resistor, Current sensor, street lights, all are connected with the microcontroller. The details of the system are given below:

The system is operated by two power sources. The main power source of the system is the main grid line and another power source is the battery. The battery reserves its power from the Piezoelectricity generator. Piezoelectricity generator will produce electricity by converting the mechanical energy into electrical energy. Then the converted electricity is rectified with the rectifier. To store the power into battery, AC needs to be converted into DC. Both of the power sources are connected to a relay switch and that switch is connected to a microcontroller. The relay switch and the microcontroller are connected to each other to create a feedback system. At first, the microcontroller checks the power of the battery. If the power is sufficient to operate the system, then it allows the battery to give the power otherwise it allows the main grid to operate the system. Then the supplied voltage either from the grid or from the battery is passed through a voltage regulator to maintain a stable voltage in the system. LDR and Ultrasonic sensor are connected with the microcontroller and all of them work together to detect the vehicles or human in the street. The ultrasonic sensor actually detects any kind of object in the street and then LDR checks whether the object has its own light or not. Depending on the signal from the LDR, microcontroller takes the decision to operate the street lights through relay switches. PIR sensor is used in the zebra crossing line and in the bus station to detect the human. When a person walks on the zebra crossing line to cross the road, the corresponding lights for the zebra crossing line will turn ON. It will give a warning to the driver and the zebra crossing line will be focused brightly.
The same procedure is used for the bus station. When a person appears at the bus station, the PIR sensor detects the presence of the person and the microcontroller allows the bus station's lights to be turned ON. The current sensor is used in this system with the microcontroller to detect the malfunction of the system. If any kind of malfunction is detected in the system, the microcontroller will send a report to the control room to take a step immediately. This is the system description of the system.

B. simulation

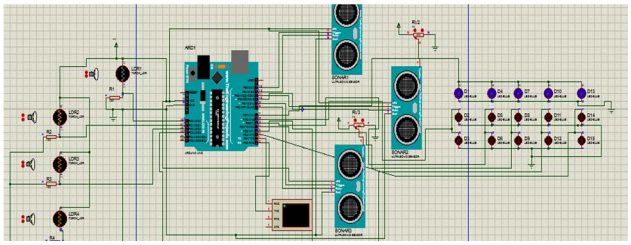


Fig 2: General night condition (minimum intensity) without any object.

In this figure, it is seen that the main LDR (LDR-1) senses lower light from the surroundings. The system is detecting the environment as a general night condition. In this situation, the LED's D1, D4, D7, D10, D13 are turned ON and the rest of them are turned OFF. This situation is treated as a normal night condition without any objects so the four LED's are turned ON and this is the minimum light intensity of the system.

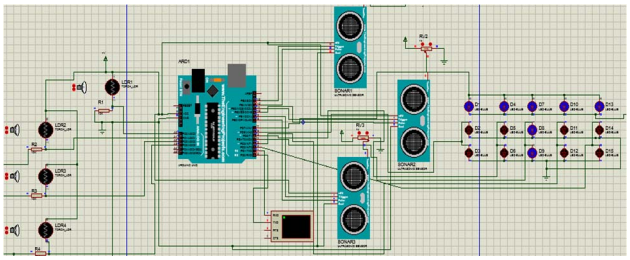


Fig.3: Maximum intensity of a particular pole with detection of human.

In this figure, a condition with the detection of a person in the system is shown. A person is under the pole no.3 and this triggers SONAR3 turning all the lights (D7, D8, D9) of pole 3 ON, whereas in other poles, only one light is turned ON for each. This picture actually shows the condition when a person will cross a certain pole on the road and the corresponding pole will give the maximum intensity.

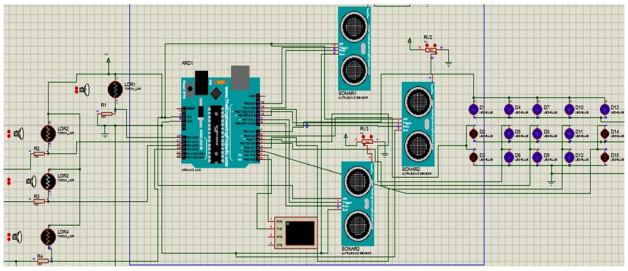


Fig 4: Maximum intensity of three poles at a time with detection of vehicle.

Fig.3 shows a condition that describes a situation with a vehicle running on the road. When a vehicle will pass the road, all lights of the three poles will be turned ON. In this figure, it is shown that when a vehicle with self-light is approaching towards the poles, LDR2 and SONAR2 get triggered and all lights of pole2, pole3, pole4 are turned ON. This figure actually describes such a condition that when a vehicle will be passing the road, corresponding three poles will give the maximum intensity.

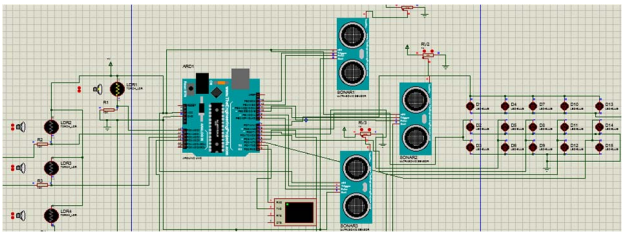


Fig 5: Normal daylight condition.

Figure 5 shows the normal daylight condition when the LDR1 senses the daylight from the surroundings. It gives a signal to the microcontroller and all lights of the poles will remain completely turned OFF. In this situation, even if there is any person or vehicle on the road, all the lights will be turned OFF in any situation.

IV. Overview of the Prototype

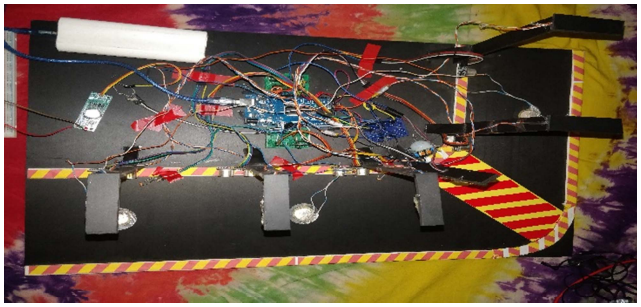


Fig6: Prototype of The System.

A prototype was made for describing the system properly. In this prototype in each street light pole, there is a combination of ultrasonic sensor and LDR. Another LDR is situated at a different place which is far from the street light pole. This LDR works as the main LDR of the system. It detects the day/night condition by measuring the natural light intensity and gives different signals to the microcontroller. A PIR sensor is used in the zebra crossing line and at the bus station to detect the human appearance. In this portion the details of the prototype are given below:



Fig7: Low light detection of the surrounding

Fig.7 shows an overview of the prototype during night. In this figure, it is seen that the street lights are turned on with a minimum intensity. When surrounding lights are lower, at that time the street lights are turned on with its inbuilt automation system.



Fig 8: Detection of objects with self-light

Fig.8 shows the condition of the system when an object having self-light approaches towards it. when a vehicle or any object will run on the road with its self-light turned on, the street light will be operated in such a way that the vehicle gets the proper light on the road. When the vehicle keeps running on the road, three lights will be turned on at a time with high intensity and after passing each street lamp, it will be turned OFF automatically. This system actually shows that there will be a number of three lights which will give the light for the passing vehicle.

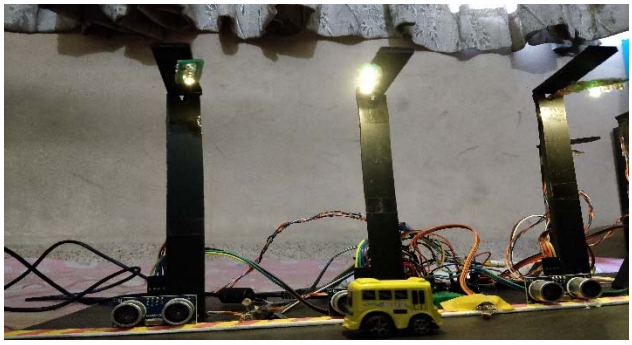


Fig9: Detection of objects without self-light.

Fig.9 shows the condition of the system when an object without self-light approaches towards it. When a person will walk on the footpath, a street light will be turned on immediately. When the person will reach under the next street light, that light will be turned on and the previous one will be turned OFF automatically.

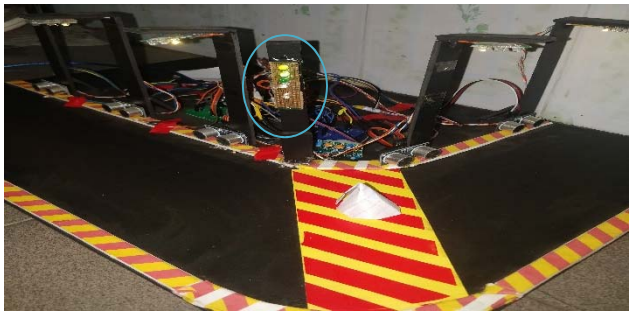


Fig10: zebra crossing point.

Fig.10 shows the zebra crossing point of the system. When a person will walk on the zebra crossing line to cross the road, the PIR sensor will detect the human and give a signal to the microcontroller to turn on the corresponding lights. Those lights will focus the zebra crossing line brightly and will give a warning to the drivers. The bus station's lights were operated in this prototype in the same manner as the zebra crossing lights.

V. Prototype Cost

SL NO.	Component List	Price Per Unit (US Dollar)	Quantity	Total Cost (US Dollar)
01.	Arduino Mega	$10	1	$10
02.	Battery	$5.95	1	$5.95
03.	Piezo electricity generator	$0.35	15	$5.25
04.	Rectifier	$0.85	1	$0.85
05.	Voltage regulator	$2.35	1	$2.35
06.	Relay module	$9.5	6	$57
07.	Led	$0.02	15	$0.3
08.	PIR sensor	$0.95	2	$1.9
09.	Ultrasonic sensor	$0.95	5	$4.75
10.	LDR sensor	$0.12	6	$0.72

11.	Current sensor	$2.35	5	$11.75
12.	Total Cost			$100.82

VI. Features And Comparative Advantages

Features available in the device:

- Smart lighting system to ensure safe road.
- Reduces the waste of electricity.
- Minimize the power consumption using led light.
- Reduce the dependency on the main grid connection.
- Easy to implement as it is cost-efficient, eco-friendly, robust and relatively accurate.
- Smart lighting system on the zebra crossing point.

Comparative advantages of the device:

There are lots of smart street light systems proposed earlier. One of those paper [4] is about IOT based smart street light energy saving system. In that system LDR sensor used to ON/OFF the street light according to the availability of surrounding light intensity level. Ultrasonic sensor used to detect vehicle or person when comes to the range then automatically increase the intensity level up to 100% of road light. In that system, there is no zebra crossing light system, absence of detecting a bus station having people or not, absence of Piezoelectricity generator and especially the damaged light monitoring system, while in our designed system all of these features are added parallelly.

In another [5] paper it was proposed that street light is controlled based on TRIAC. TRIAC controls the voltage level for varying the intensity of street light according to the traffic which is measured by sensor and based on sunset or sunlight which information comes from Internet source as power electronic circuits are combined with IOT. In this paper TRIAC based control system was proposed and IOT was also proposed. In that paper, there was also absence of Piezoelectricity generator, monitoring and detecting of people at a bus station, automatic light of zebra crossing line as well as the monitoring system of the damaged light.

Another [6] paper described street light monitoring and control based on WSN. Integrated doppler sensor was used for detecting vehicle and helping to increase the efficiency of power consumption. Current sensor was used to identify any possible malfunction. This paper came with the feature of detecting malfunction but they could introduce Piezoelectricity generator to produce more power and to save the power consumption from main grid. They could introduce the zebra crossing line system and monitoring system of a bus station. Our system will work with all these features as well as malfunction detection.

In [7] this paper, design was proposed about intelligent control of road light and reducing the power consumption by LDR, IR sensor and all of the components are connected to microcontroller. ARDUINO Uno used as microcontroller. This system also missed some features that could be included in it. They also did not introduce with another power source which could be used when the electricity supply is OFF. There is no malfunction monitoring system, zebra crossing light system as well as people and vehicles detection system. All this combination could be used on that system to make the system much better.

In [8] this paper Piezoelectric sensor was used to observe vehicle movement and save power by controlling the intensity of road light. Piezoelectric sensor was used to make the project cost-effective. ZIGBEE technology used to connect between circuit and IOT for storing information about vehicle and people. This paper appeared with an extra feature of ZIGBEE technology but there was no malfunction detection system. People or vehicles detection system was also missing in this paper. These features could be integrated to the paper to make the system more efficient and robust.

The author of [9] this paper used Piezoelectric material as the source of energy of street light. It is cost efficient and renewable energy source so the whole project easy to implement in a large scale. Full wave rectifier circuit was used to develop power extraction capability from Piezoelectric harvests. This paper proposed that there was no another electricity supply except Piezoelectricity generator but this system sometimes may face some trouble to supply enough amount of power. In Dhaka city about 420 MW is consumed by the street light consumes in every month [2]. This huge amount of power may not be produced by the only Piezoelectric material and in a rainy day or in a holiday this production may be reduced very badly. This paper also did not propose any system of automatic light controlling, zebra crossing line and especially malfunction detection. These features could be added on that system.

VII. Novelty Of The Project

This project is feasible for two reason. Firstly, a new smart lighting system is established at the zebra crossing which will allow the drivers on the road notice a person crossing the road letting a secured crossing of the pedestrian. This can greatly reduce the accidents that happens on the road due to lack of visibility or awareness by the drivers.

Secondly, this system will be established even on the bus stand light which will only turn ON when any passenger or person is near the stand leading to lower and efficient power consumption.

The above given points prove the feasibility of the project and similar type of advantage can be established to every other lighting system based on the requirement.

VIII. Limitations

Our design system is prototype design as here a few amounts of piezoelectric material are used as a result it is difficult to charge a battery with less time. It will be beneficial If it is used on a large scale and the number of piezoelectrics is increased then it can possible to charge a battery with less time.

IX. Future Aspects

This project can be improved by adding image processing technology to detect a car or person more accurately. By

using this technology if someone breaks the traffic signal in the night period then the details information of that vehicle can be identified easily.

X. Conclusion

From this project, we have not depended only on the grid connection as sometimes it is not available due to lacking electricity so piezoelectric material plays an important role by providing a renewable source of energy. These unique features of this system will provide an uninterrupted supply of electricity as this system can make a significant contribution to saving the most needed energy in a country. Power consumption is reduced as light intensity in this system can be changed with the appearance of vehicles or people on the road. Led light is used here which is eco-friendly and cost-efficient. The zebra crossing light warns the driver at night when a person crosses the road through the zebra crossing.

References

[1] Dhaka Tribune, "dhaka tribune web site," dhaka tribune, 28 january 2018. [Online]. Available: https://www.dhakatribune.com/opinion/special/2018/01/28/snatchers-wheels-terrorise-dhaka. [Accessed 9 february 2019].

[2] Dhaka Tribune, "Dhaka Tribune Web site," Dhaka Tribune News, 4 august 2014. [Online]. Available: https://www.dhakatribune.com/uncategorized/2014/08/23/much-power-gets-wasted-in-streetlamps. [Accessed 9 February 2019].

[3] H. B. Khalil, N. Abas and S. Rauf, "Intelligent Street Light System in Context of Smart Grid," in *ICCCNT 2017*, Delhi, India, 2017.

[4] R. Kodali and S. Yerroju, "Energy efficient smart street light," *2017 3rd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, Tumkur, 2017, pp. 190-193. doi: 10.1109/ICATCCT.2017.8389131

[5] O. Rudrawar, S. Daga, J. R. Chadha and P. S. Kulkami, "Smart street lighting system with light intensity control using power electronics," *2018 Technologies for Smart-City Energy Security and Power (ICSESP)*, Bhubaneswar, 2018, pp. 1-5. doi: 10.1109/ICSESP.2018.8376692

[6] A. Lavric, V. Popa and S. Sfichi, "Street lighting control system based on large-scale WSN: A step towards a smart city," *2014 International Conference and Exposition on Electrical and Power Engineering (EPE)*, Iasi, 2014, pp. 673-676. doi: 10.1109/ICEPE.2014.6969994

[7] A. Abdullah, S. H. YusOFF, S. A. Zaini, N. S. Midi and S. Y. Mohamad, "Smart Street Light Using Intensity Controller," *2018 7th International Conference on Computer and Communication Engineering (ICCCE)*, Kuala Lumpur, 2018, pp. 1-5. doi: 10.1109/ICCCE.2018.8539321

[8] R. Abinaya, V. Varsha and K. Hariharan, "An intelligent street light system based on piezoelectric sensor networks," *2017 4th International Conference on Electronics and Communication Systems (ICECS)*, Coimbatore, 2017, pp. 138-142. doi: 10.1109/ECS.2017.8067855

[9] K. Haque, M. R. Khan, N. Nowrin and H. U. Zaman, "Smart street lights using piezoelectric materials," *2016 International Conference on Microelectronics, Computing and Communications (MicroCom)*, Durgapur, 2016, pp. 1-4. doi: 10.1109/MicroCom.2016.7522506

Electrical and Morphological Characterization of Zinc-Doped α- Fe_2O_3 Thin Films at Different Annealing Temperature

Arman Ibn Shahjahann[1], Abu Bakar Md. Ismail,
Department of Applied Physics and Electronic Engineering, University of Rajshahi,
Rajshahi 6205, Bangladesh
swaran.apeeru@gmail.com

Abstract— **α Fe_2O_3 is naturally n-type material. Zn was used to dope α-Fe_2O_3 (Zn: Fe_2O_3) to turn it into p-type. The thin film of Zn: Fe_2O_3 was obtained by spin coating a blend of Zinc acetate (Zn $(CH_3COO)_2$ $2H_2O$) and $FeCl_2$ after annealing the film in air. Zn-doped-Fe_2O_3 was found as a p-type semiconductor by Hall Measurement. Effect of annealing at 450°C and 550°C temperature on electrical properties were also investigated. The resistivity of those films was found to be between 1~10 Ω-m. It was observed that the resistivity of the films increases with increasing temperature for most of the sample. It was also observed that the resistivity of the film is inversely proportional to the increasing number of film layer and thickness within most of the sample. Atomic Force Microscopy (AFM) was used to observe the surface morphology of the annealed film on glass substrate. The variation of the average roughness of annealed film on the glass substrate was from 60nm to 98nm and the thickness was 160nm to 540nm.**

Index Terms—*Thin film, Zn-doped-Fe_2O_3, Resistivity*

I. INTRODUCTION

Non-conventional energy sources have a great significance for the world's future, given the environmental issues related to energy generation and energy's importance in our society. To overcome a predictable energy crisis, there are intense researches on alternative energy harvesting techniques [1-3] . Solar energy is a good solution for these energy crises [4]. Hematite (α-Fe_2O_3) is an attractive material for PEC(photoelectrochemical) studies due to its ban-gap (Eg = 2.2 eV), lying nearly in the optimum range for the solar splitting of water.Based on this band gap, solar energy has absorbed were sun emits maximum energy. Besides, α-Fe_2O_3 is naturally abundant in the earth's crust and is, therefore a low-cost material. It is also corrosion-resistant in acidic and alkaline medium [5-10]. One major drawback of Fe_2O_3 is its unfavourable photo-response that arises mainly due to its low conductivity and consequent recombination of photo-generated carriers. Moreover to make a homo-junction of Fe_2O_3, which is naturally, n-type, p-type conduction in Fe_2O_3 is needed that is long-term stability.

After the first report of Fujishima on photo-assisted electrochemical water oxidation on TiO_2 in 1972[11] researchers have been investigating the wide bandgap metal-oxide semiconductors such as titanium dioxide (TiO_2), zinc oxide (ZnO), iron oxide like hematite (α-Fe_2O_3) and tin dioxide (SnO_2) for their application in solar cells, solar fuel, photocatalysis and energy storage devices [12-18]. In our recent past, Photoresponse of p-Type Zn:Fe_2O_3 Oxide Thin Films has been investigated by William B. Jr, John P and U. M. Khan. [19]

One of the fundamental aim of this research to make p-type Zn doped Iron(III) Oxide thin films, analyzing their electrical properties and to learn the fabrication which can be used as a solar cell or homojunction material for any kinds of semiconductor device in future. The present communication describes the PEC study on Zn-doped α-Fe2O3 thin films prepared at different doping concentrations (0.1M, 0.005M and 0.0025M). Resistivity was found to depend upon the doping concentration of Zn and thickness. For thin film deposition, there are various kinds of deposition methods like Chemical vapour deposition, Spin coating, Thermal evaporation, Dip coating. Spray pyrolysis etc[20-22]. Here we have applied spin coating methods in this research project, which is one of the easiest techniques for preparing thin films [23-24]. In this research Zn:Fe_2O_3 thin films with different doping concentrations were prepared by the spin coating method and doping effects of these films after annealing were also studied. After synthesis Zn:Fe_2O_3 solution is taken by micropipette and drop cast on glass surface then spin coated.

II. EXPERIMENTAL METHODOLOGY

A. SOLUTION PREPARATION FOR DOPING

The p and n-type semiconductor were separately prepared by a simple one-pot spin coating method on ultrasonically cleaned 1x1 inch sizes glass substrates. The substrates were also chemically cleaned with the piranha solution (5:1 H_2SO_4 and H_2O_2 mixing) for the increase of hydrophilic properties in which enhance water-compatibility. After that, the dried substrates with atmospheric air were available for the use.

0.11 M $FeCl_2$ Solution Preparation:

Ethanol	100 ml
$FeCl_2$	1.3943 g

0.01M $(CH_3COO)_2Zn.2H_2O$(Zinc acetate) Solution Preparation:

Ethanol	100 ml
$(CH_3COO)_2$ $Zn.2H_2O$	219.527 mg

when the two solutions were mixed and stirred for 10 minutes. After that Zn doped α-Fe_2O_3 solution had prepared. In this way, 0.005M and 0.0025M Zn doped α-Fe_2O_3 solution had prepared. After synthesizing Fe_2O_3, 5μL of the solution was taken by micropipette and dropped on loaded substrate surface then spin-coated with the following rotational speed and time:

SP1	T1	SP2	T2
500rpm	3s	7000 rpm	30s

B. ANNEALING

Two sets of the sample were annealed after deposition at two different temperature (450^0C & 550^0C) for 4 hours in a thermal annealing furnace (carbolite CWF-1200). Then they were left to be cooled naturally to the room temperature.

C. Studies with Atomic Force Microscope (AFM)

Atomic force microscope (AFM) [25] of model (XE 70) park system was used to study the roughness (nm) [26] of the all film surface.

III. Results & Discussion

A. Resistivity Measurement

The electrical resistivity of Zn:Fe_2O_3 semiconductor measured by using Van der Pauw's technique [27-28]. The resistivity variation of different layer Zn doped α-Fe_2O_3 thin film at different temperature and different molar concentration are shown in the following table:

Table 1: The resistivity variation of different layer at 450˚C temperature with different molar concentration

No. of Layer	0.01M Resistivity (Ω-m)	0.005M Resistivity (Ω-m)	0.0025M Resistivity (Ω-m)
Layer 1	10.8	2.2	1.7
Layer 3	8.0	1.8	1.1
Layer 5	4.0	1.1	0.9

Table 2: The resistivity variation of different layer at 550˚C temperature with different molar concentration

No. of Layer	0.01M Resistivity (Ω-m)	0.005M Resistivity (Ω-m)	0.0025M Resistivity (Ω-m)
Layer 1	8.2	6.0	0.7
Layer 3	7.0	2.9	0.6
Layer 5	7.8	3.1	1.9

It was observed that the resistivity of annealed films increases with temperature. The increase in resistivity was not linear. It was seen that the Resistivity of the films decreased with decreasing the molar concentration.

Resistivity variation of different layer Zn doped α-Fe_2O_3 thin film at a different temperature and different molar concentration are shown in Fig. 1, Fig. 2, Fig. 3

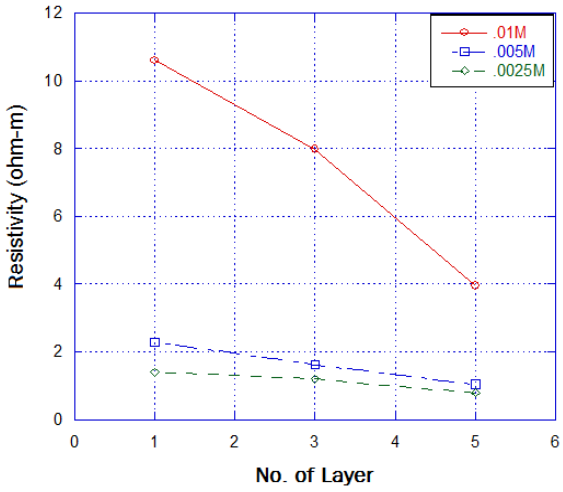


Fig. 1 Resistivity of different molar concentration 0.01M, 0.005M and 0.0025M and different layer at temperature 450˚C .

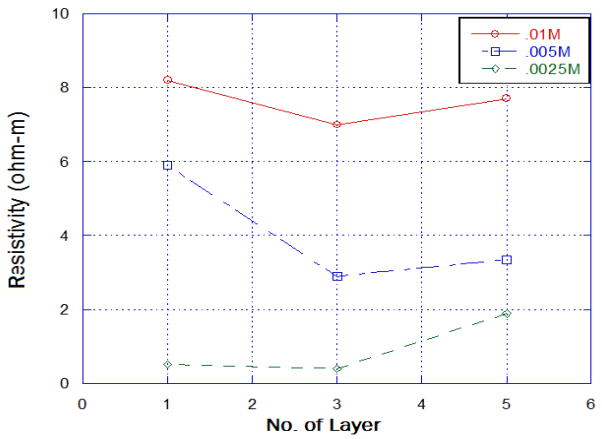


Fig. 2 Resistivity of different molar concentration 0.01M, 0.005M and 0.0025M and different layer at temperature 550˚C .

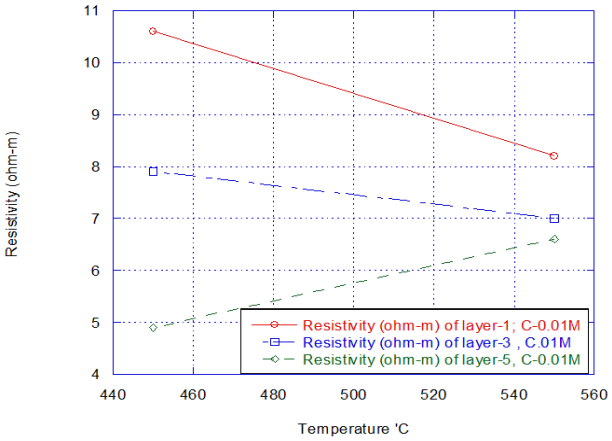


Fig. 3 Resistivity of different layer and different temperature at 0.01M molar concentration

Resistivity at different temperature and different molar concentration in layer 5 as shown in Fig. 4

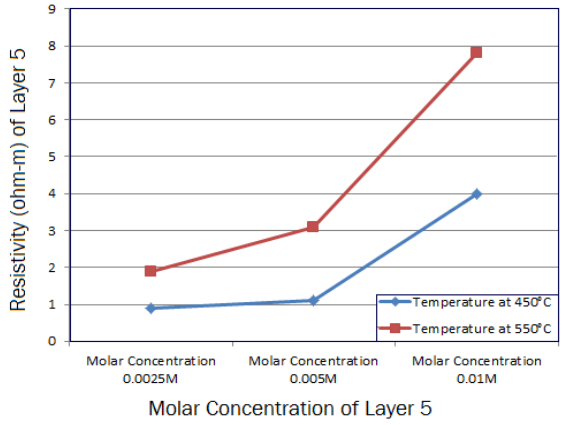
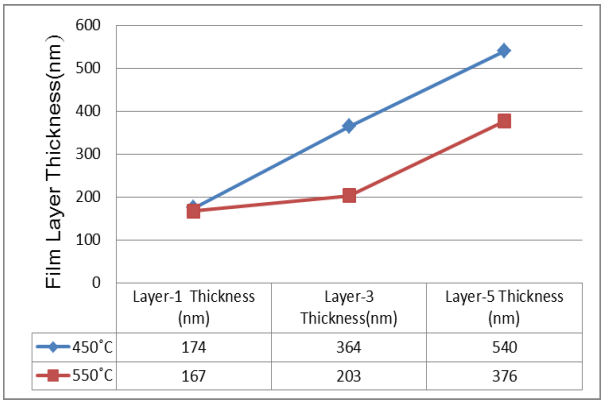


Fig. 4 Resistivity at different temperature and different molar concentration in layer 5.

B. Surface Morphology Study using AFM

The thickness of the Zn doped α-Fe_2O_3 thin film varied with annealing temperature and a number of the film layer. Following Fig-5 has shown the thickness (nm) variation of Zn doped α-Fe_2O_3 thin film with temperature:



	Layer-1 Thickness (nm)	Layer-3 Thickness(nm)	Layer-5 Thickness (nm)
450˚C	174	364	540
550˚C	167	203	376

Fig-5: Thickness variation of the thin film layer at different annealing.

The thickness of the film has been increased with increasing the number of the film layer and decreased with the increase of annealing temperature. This was due to the fact that increasing of annealing temperature reduces various types of defects like voids, crystal defects and dislocation density.

The relation between annealing temperature & the thin film surface roughness was inversely proportional. Variation of the average value of Roughness of Zn doped α-Fe_2O_3 at 450 ˚C, 550˚C and AFM images are the following:

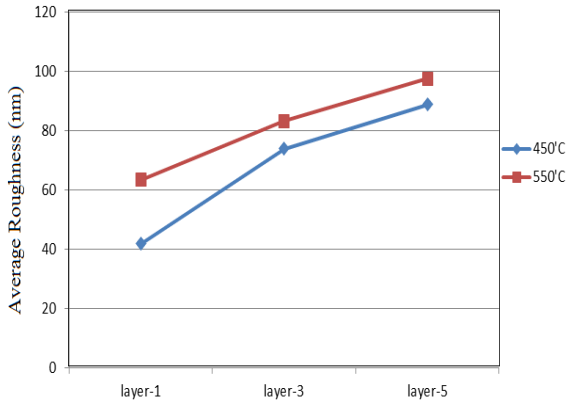
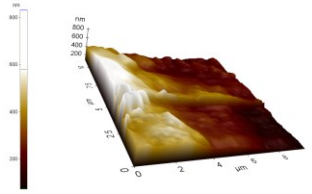
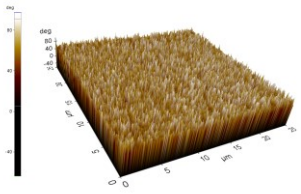


Fig.6: Average value of Roughness (nm) of the film surface of different layer at different temperature of the thin film

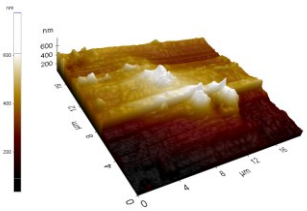


(a)

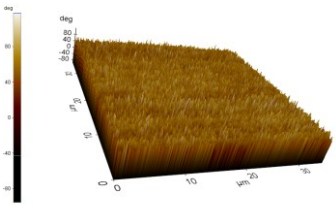


(b)

Fig. 7 Surface morphology of Zn doped α-Fe_2O_3 Layer 5 at 450˚C near to the Edge (a) and middle of the surface (b)



(a)



(b)

Fig.8: Surface morphology of Zn doped α-Fe_2O_3 Layer -5 at 550˚C near to the Edge (a) and middle 0f the surface (b)

IV. Conclusions

The effect of different parameters (annealing temperature, thickness and molar concentration) on the electrical property of the spin-coated thin film of Zn:Fe_2O_3 on the glass substrate has been investigated. It was observed that the resistivity of annealed films increases with temperature. The increase in resistivity was not linear. It was seen that the Resistivity of the films decreased with decreasing the molar concentration. The semiconductor type was not stable for very long time, their properties were converted from p-type to n-type as well as the carrier concentration were degraded with the time to leave due to their solid format of zinc. But the lightly doped Zn:Fe_2O_3 retained its p-type conduction for comparatively long time. Because lightly doped Zn:Fe_2O_3 can be able to form comparatively highly attraction bond.

It was also observed that the resistivity of the thin film decreases with increasing of film thickness (increasing number of film layer). Films with lower thickness may have defects like lattice defects, voids etc. These defects significantly contribute to resistivity. When the film thickness was increased more atoms are deposited on film surface which reduces the defects (voids, crystal defects [29] and dislocation density[30]) on film. As a result, resistivity decreases.

The surface of Zn:Fe_2O_3 was investigated with atomic force microscope (AFM) to see the effect of annealing on surface roughness. It was observed that the surface roughness and thickness was reduced after increasing the annealing temperature. This was due to the fact that increasing of annealing temperature reduces various types of defects like voids, crystal defects and dislocation density. So the film becomes more smooth which was expected for thin film.

References

[1] Chen, S. Shen, L. Guo, S. S. Mao, Chem Rev, 110 (2010) 6503.
[2] P. Zhao, C.X. Kronawitter, X. Yang, J. Fu, B.E. Koel, Phys. Chem. Chem. Phys. 16 (2014) 1327.
[3] Y.K. Hsu, Y.C. Chen, Y.G. Lin, ACS Appl. Mater. Interfaces 7 (2015) 14157
[4] Fizeau, H. (1851). *Philosophical Magazine* **2**: 568–573.
[5] J.S. Chen, et al, J. Am. Chem. Soc. 132 (2010) 6124.
[6] Y. Izumi, Coord. Chem. Rev. 257 (2013) 171.

[7] L. Mai, X. Tian, X. Xu, L. Chang, L. Xu, Chem. Rev. 114 (2014) 11828.
[8] J. Schneider, et al, Chem. Rev. 114 (2014) 9919.
[9] L. Jing, W. Zhou, G. Tian, H. Fu, Chem. Soc. Rev. 42 (2013) 9509.
[10] H. Zeng, G. Duan, Y. Li, S. Yang, X. Xu, W. Cai, Adv. Funct. Mater. 20 (2010) 561.
[11] A. Fujishima, K. Honda, Nature 238 (1972) 37. [1] E.J.W. Crossland, et al,
[12] J.S. Chen, et al, J. Am. Chem. Soc. 132 (2010) 6124.
[13] Y. Izumi, Coord. Chem. Rev. 257 (2013) 171.
[14] L. Mai, X. Tian, X. Xu, L. Chang, L. Xu, Chem. Rev. 114 (2014) 11828.
[15] J. Schneider, et al, Chem. Rev. 114 (2014) 9919.
[16] L. Jing, W. Zhou, G. Tian, H. Fu, Chem. Soc. Rev. 42 (2013) 9509.
[17] H. Zeng, G. Duan, Y. Li, S. Yang, X. Xu, W. Cai, Adv. Funct. Mater. 20
[18] Nature 495 (2013) 215
[19] J. Am. Chem. Soc., 2004, 126 (33), pp 10238–10239,Publication Date (Web): July 30, 2004, Copyright © 2004 American Chemical Society
[20] *Electrical Engineering Materials and Devices*, 2nd Ed. 2001
[21] *Solid State Electronic Devices*, Prentice Hall, 1995
[22] S. Tolansky in *"Multiple Beam Interferometry of surfaces and films"*, Oxfored University Press,1948.
[23] J.G. Gotlling and W.S. Nicol. J. Opt. Soc. Am., Vol. 56, 1966, p. 1227.
[24] *Electrical Engineering Materials and Devices*, 2nd Ed. 2001
[25] Effect of roughness as determined by atomic force microscopy on the wetting properties of PTFE thin films†Authors J. D. Miller,S. Veeramasuneni, J. Drelich,M. R. Yalamanchili, G. Yamauchi
[26] Calibration of atomic - force microscope tips Jeffrey L. Hutter and John BechhoeferView Affiliations, Department of Physics, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada
[27] Van –der-Pauw. Philips. Res. Rept., Vol. 13, 1958, p.l.
[28] *On the Van der Pauw method of resistivity measurements*, A.A.Ramadan[a]. R.D.Gould[b]. A.Ashour[a] Volume 239, Issue 2, 1 March 1994, Pages 272-275
[29] *Crystal Defects and Crystalline Interfaces*, By Walter Bollmann, 1961, Pages 37-142
[30] *Crystallography and Crystal Defects* By Anthony Kelly, Kevin M. Knowles , 2nd edition, Pages 241-332

Fabrication of Digital Sol-Gel DIP Coating System with Smart Beaker Carrier

Sarafat Hussain Abhi
Department of Electrical and Electronic Engineering, Rajshahi University of Engineering & Technology (RUET),
Rajshahi-6204, Bangladesh
email: sarafat.rueteee@gmail.com

Md. Shoaib Akther
Department of Electrical and Electronic Engineering, Rajshahi University of Engineering & Technology (RUET),
Rajshahi-6204, Bangladesh
email: shoaibkiron@gmail.com

Dr. Md. Faruk Hossain
Department of Electrical and Electronic Engineering, Rajshahi University of Engineering & Technology (RUET),
Rajshahi-6204, Bangladesh
email: faruk94_ruet@yahoo.com

***Abstract*— In this work, a digitalized sol-gel system is prepared for fabrication of efficient thin films. The proposed system of this paper offers dipping and withdrawing of a substrate with precise timing and has an extra feature of holding multiple beakers having multiple solutions for fabricating complex layered films. This system also provides rotation of the beakers containing required solution to be deposited on a substrate with the help of pressing only a button. Duration of dipping time of the substrate can also be set using this system. In short this digitalized system offers multiple features of thin film fabrication using sol-gel technique with high percentage of minimizing defects.**

***Keywords*— Sol-Gel technique, L293D motor driver, DC gear motor, 16*2 LCD.**

I. INTRODUCTION

Thin films have many applications in this modern age like fabrication of electronic components, solid-state devices, microelectronics IC's etc. Films are generally defined by layering a substrate (glass, FTO, plastic etc.) with a solution having the required properties for the purpose of fabricating that film after that the layer on top of the substrate is either heated on rapidly thermal annealed. If the film thickness is less than one micron then the film is called thin film. There are many ways of fabricating a film such as sol-gel dipping, spin-coating, sputtering, spray-pyrolysis, thermal evaporation, CVD, PVD etc. Among them sol-gel technique is not only easier but also cost efficient. By this technique both amorphous and crystalline films can be fabricated [1]. Conventional sol-gel technique was briefly discussed in the year 2002 [2], but that technique was not automatic and requires manual control which is not appropriate for preparing uniform thin films. M. F. Ali, U. Saha and M. F. Hossain designed an automatic length controlled sol- gel system in the year 2012 which can deposit thin film for 2 cm, 3 cm, 4 cm, 5 cm, 6 cm & 7 cm [3], but that system requires the use of ribbon which is very much subjected to vibration and the dipping and withdrawing speed cannot be controlled in that system. Later in the year 2015 Dr. M. F. Hossain and Ayesa Tuz Johra proposed a digital sol-gel system which was designed to solve the problem of using ribbon by using mechanical gear but that was not completely vibration less [4]. Our proposed system has the extra feature of having a beaker carrier and movable position of the horizontal circular surface attaching the beakers. This machine also has less vibration compared to previously designed systems and has fewer defects on the resultant prepared films. By using our proposed system multilayered films can be fabricated as the sol-gel dip coating system is a very useful technique for fabricating & characterizing self-assembled structures & multi-layered thin films [5].

II. DESIGN AND COMPLETE SETUP OF THE PROPOSED SYSTEM

A. Block Diagram

Fig. 1 shows the basic block diagram representation of our proposed digital sol-gel system. The main parts of the digital sol-gel system are dc power supply unit, 3*3matrix keypad, 16*2 LCD unit, motor driver, pulley, belt, lead screw, motion filter, substrate holder and a microcontroller unit which is the heart of the automatic arrangement. DC power supply unit rectifies 220V ac supply into 5V dc which is supplied to various components of the system.

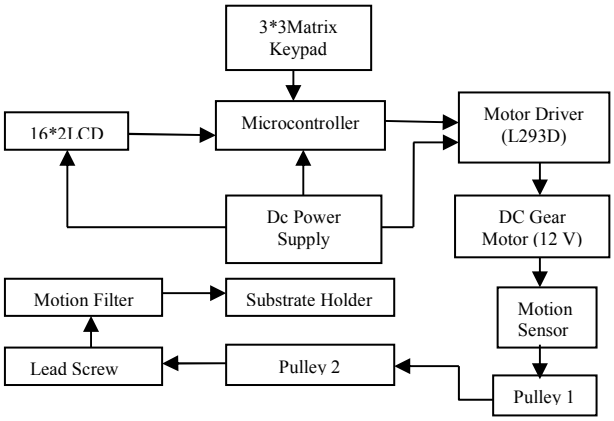


Fig. 1. Block diagram representation of digital sol-gel system.

The dc gear motor used in this work is of 12 V dc bipolar motor. A motor driver a 16 pin IC named L293D has been used. Pulley belt and lead screw mechanism has been used to convert the circular motion of DC gear motor into linear motion and a crocodile clip is connected with the rod which is used to hold the substrate. As a input device a 3*3 matrix keypad and as a display unit a 16*2 LCD character display has been used. After giving input value through matrix keypad microcontroller sends out control signal to the motor driver to drive the DC gear motor. According to the control signal, the DC gear motor starts rotation either in clockwise or anticlockwise direction.

B. Initial Setup of the Automatic Sol-gel System

Initial setup of our proposed system is given in Fig. 2. The total body is covered by steel sheet. The structure of the machine is made with aluminum sheet to increase the stability of the machine. The lower portion consists of the electronics part that is microcontroller, DC power supply unit, motor driver, LCD, matrix keypad, pulley 1, pulley 2, belt etc.

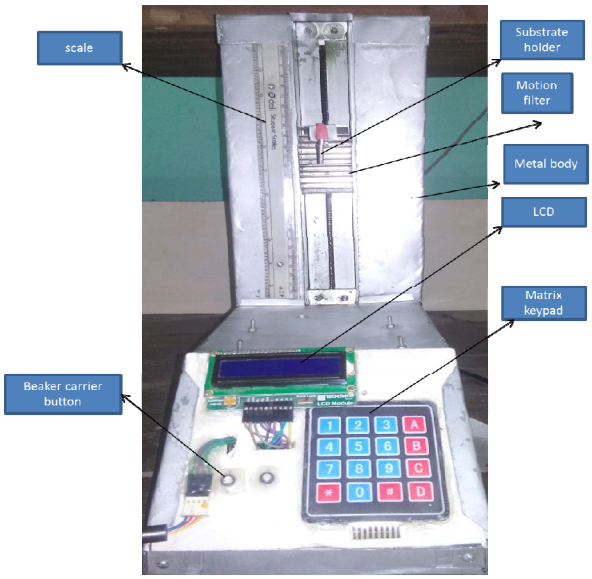


Fig. 2. Initial setup digital sol-gel system (without beaker carrier).

Fig. 3 shows the internal view of the machine. Pulley 1 is attached with the motor shaft and Pulley 2 is connected with pulley 1 by a belt. Then pulley 2 is connected to a lead screw which is placed vertically from lower portion to the upper portion of the machine. A rod (to which the substrate holder is attached) with a screw is attached to the lead screw. So when the lead screw rotates the rod moves upward or downward according to the rotation of the lead screw. Thus the circular motion of the motor shaft is converted to linear motion.

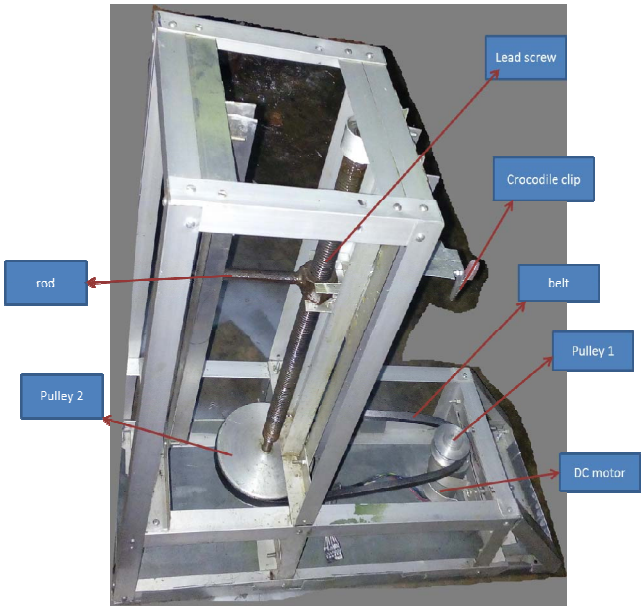


Fig. 3. Internal view of the designed machine.

A motion filter is used to reduce the mechanical vibration when the motor rotates to have a uniform speed of the substrate. The width of pulley 1 and pulley 2 are 1 inch and 3.5 inch respectively. This means that to rotate pulley 2 for one time, pulley 1 has to rotate for three and half times. Now the lead screw has 13 threads per inch. That means when the screw rotates for one time its vertical distance is 1/13 inch. So to have a vertical distance of one inch the lead screw has to rotate 13 times, pulley 2 has to rotate for 13 times, pulley 1 has to rotate for 45.5 (= 13*3.5) times. Thus we can have a very lower speed of the linear motion. The lowest speed that we could achieve is 1 mm/sec.

C. Complete Setup of the System with Beaker Carrier

Suppose we want to have layers of different solution on a substrate or layer of different solution on different substrates. So we will have beakers of different solutions. We can't place several beakers on the beaker base and every time we have to change the beaker manually to have layers of different kind of solutions on the substrate. To remove this problem an additional feature I have made which is the smart beaker carrier. We can see it in Fig. 4 (a), it contains four sections to place four beakers on it. And with a press on the button associated for the beaker this beaker carrier can rotate clockwise. It rotates 90 degree at a time and consequently we can place our required beaker underneath the substrate holder among the four beakers. A stepper motor is used to have the circular motion of the beaker carrier.

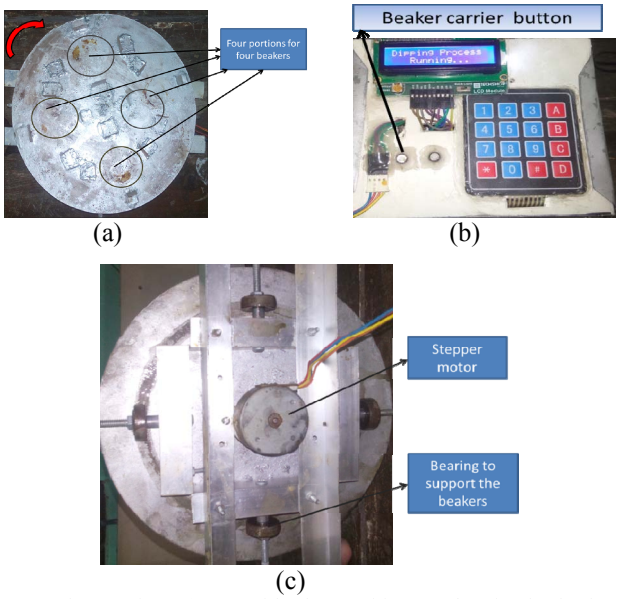


Fig. 4. Beaker carrier: (a) No. of beaker position on the circular beaker carrier, (b) Beaker carrier button & (c) Bearing and motors of the beaker carrier.



Fig. 5. Final setup of the digital sol-gel system (with beaker carrier).

The stepper motor of the beaker carrier takes power from the switch placed in front of sol-gel machine. Now if we wish we can use the beaker carrier or we can do the dipping process without it. There are four bearings under the base of the beaker to support the four beakers placed on it which is shown in Fig. 4 (c). The base is mounted with the shaft of the stepper motor. When the motor rotates the beaker carrier rotates clockwise for 90 degree at a time when we press the button. The final setup of our proposed system is shown in Fig. 5.

III. Operating Principle of Digital Sol-gel System

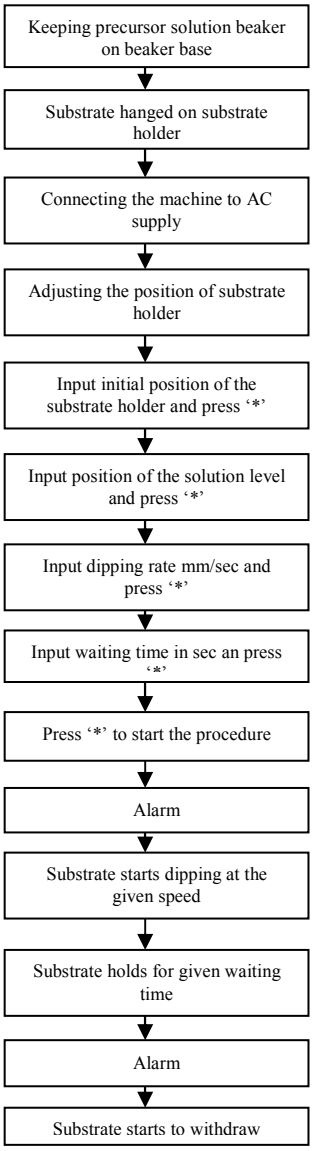


Fig. 6. Operating principle of the proposed system.

The operating principle of digital sol-gel system is shown in Fig. 6 which is basically a block diagram. At first prepared precursor solution is kept on the beaker base and a clean substrate is hanged on the substrate holder. Then the machine plug is connected to the Ac line. After making the primary setup, power switch is turned on and a message "Welcome to Nano Lab of Ruet….Automatic sol-gel machine" comes out on display. After 2 sec a message "press '*' to Start, A or B to adjust" comes to display and we can adjust the position of the substrate holder. Then a massage "Input Initial Position" comes out on display and we have to input the initial position of the substrate holder. Then a massage "Input Solution Level" comes out on display and we have to input the solution level. After that a massage "Input DIP Speed" comes out on display and we have to input the required dipping speed of the substrate. Then a massage "Input Waiting Time" comes out on display and we have to input the desired waiting time before withdrawing the substrate. There must not to be made hurry while pressing two keys one after another. Because intentionally there have been made a delay of 0.5s between two keys pressing for programming purpose.

After correctly giving all of the input values from keyboard the motor starts to rotate in anticlockwise direction to move the substrate in downward direction at a rate of given rpm. And a message "dipping process running" comes on display. After completing the dipping length the substrate wait in the solution for a given waiting time. Then it starts to withdraw at a given rpm and when it reaches to a predetermined position it closes an alarm switch and alarm sound is heard and a message "Thank you" comes on display and the current operation breaks down. Then the digital sol-gel system goes to the start position and is prepared for another operation.

IV. Results and Discussion

To examine the results of the proposed digital sol-gel system we prepared ZnO thin film using this machine. Images of different lengths of substrates covered by ZnO thin films that have been deposited by proposed digital sol-gel process is shown. All the prepared films have been characterized by optical properties with UV visible spectrometer in the wavelength range of 300-900 nm. Fig. 7 show the images of ZnO films prepared by length variation by proposed digital sol-gel process. It can be easily figured out that the proposed digital sol-gel system is successful in depositing ZnO thin films on the required length of substrate without any problem. It is also clearly shown that the surface of prepared films has been grown uniformly.

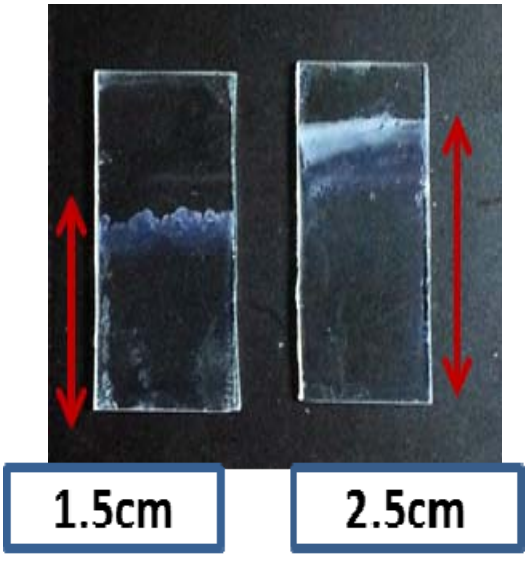


Fig. 7. ZnO thin film by length variation.

Fig. 8 shows the optical transmittance spectra of the prepared film of length 2.5cm. The film has a very sharp cut-off at around 375 nm wavelength and reaches to peak at around 400 nm.

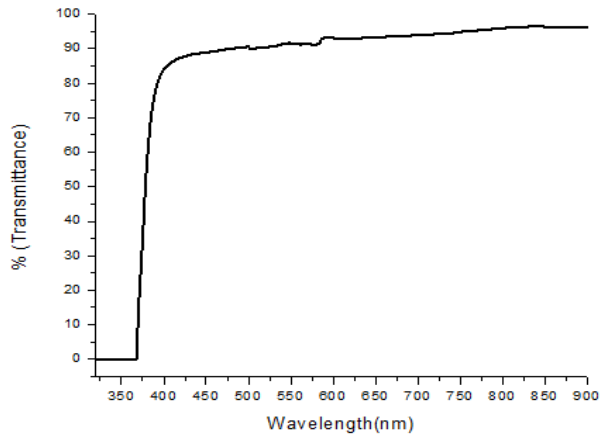


Fig. 8. Transmittance spectra of ZnO thin film (2.5cm) made by proposed machine.

According to Tauc's law relation between absorption co-efficient (α) on photon energy (hν) can be written as [6]:

$$\alpha h\nu = A\left(h\nu - Eg\right)^{m} \quad (1)$$

where, $m = \frac{1}{2}$ and 2 for direct and indirect transitions respectively. α is known as absorption coefficient, A is defined as the edge width parameter and hν is called the photon energy. Fig. 9 shows the curves of $[\alpha h\nu]^2$ vs. photon energy hν for direct band gap transitions of ZnO thin film.

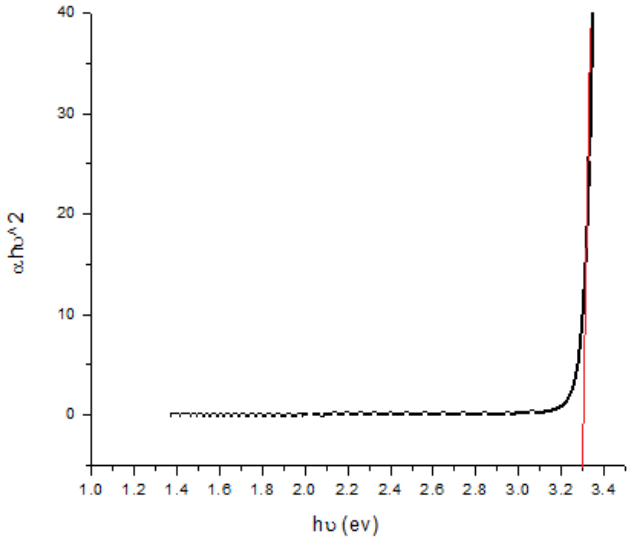


Fig. 9. The curve of $[\alpha h\nu]^2$ vs. photon energy hν for ZnO thin film (2.5cm).

V. Conclusion and Future Work

The work reported in this paper is mainly focused on development of a new user friendly automatic sol- gel system and after that preparation of ZnO thin film for result verification of the system. Starting with aim of the work, first the limitations of existing digital sol-gel system were identified. In spite of the higher cost the new system overcomes the limitations of existing systems to a great extent. Lastly ZnO thin film was prepared successfully by length variation. And prepared films were characterized with respect to optical property (transmittance, absorbance, direct bandgap) by using UV visible spectrophotometer for 300-900 nm. This paper represents the modeling of a new automatic sol-gel system. The prepared ZnO can be used for various applications like transparent conductive oxide, UV detector, optical window, DSSC solar cells, waste water purification, air purification etc. Future directions related to the work in this thesis are improving the modeling of substrate holder more substrate can be deposited at a time, to achieve more lower speed of the substrate up to 1cm/min.

Acknowledgement

For this work the authors of this paper is very much thankful to Rajshahi University of Engineering and Technology (RUET) for giving them the opportunity to use Nanotechnology and Photovoltaic's lab for successful outcome of this experiment.

References

[1] J.L. Vossen and W. Kern, Thin Film Process II, Elsevier, p. 502, ISBN 9780080524214

[2] S. M. Attia, J. Wang, G. Wu, J. Shen and Jianhua " Review on sol-gel drived coatings: process, technique and optical applications", journal of materials science and technology, vol. 18, no. 3, pp. 211-217, 2002.

[3] M. F. Ali, U. Saha and M. F. Hossain "Design an Automatic Dipping and Withdrawing System for the Deposition of Thin Films", ICECTE 2012.

[4] Dr. M. F. Hossain and A. T. Johra, "Optimization of Microcontroller Based Digitalized Dipping and Withdrawing System for Sol-Gel Method", ICEEICT 2015.

[5] C. J. Brinker , G. C. Frye, A. J. Hurd And C. S. Ashley, "Fundamentals of sol-gel dip coating", Thin Solid Films, vol. 201, no. 1, pp. 97-108, 1991.

[6] N. Shakti and P. S. Gupta, " Structural and Optical Properties of Sol-gel Prepared ZnO Thin Film", Applied Physics Research, vol. 2, no. 1, pp. 19-28, 2010.

Enhancement of Sensitivity for Surface Plasmon Resonance Biosensor with Higher Detection Accuracy and Quality Factor

Mohammad Hasibul Hasan Hasib
Electrical and Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
hasib.eee.ruet@gmail.com

Jannati Nabiha Nur
Electrical and Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
aubonti88@gmail.com

Kamrun Nahar Shushama
Electrical and Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
shushamanahar@gmail.com

Imteaz Rahaman
Electrical and Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
imtiaz.arup48@gmail.com

Md. Masud Rana
Electrical and Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
md.masud.rana.ruet@gmail.com

Mohammad Al Mahfuz
Electrical and Telecommunication Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
almahfuz.ete.ruet@gmail.com

***Abstract*—In this paper, we demonstrate a highly sensitive Kretschmann configuration based surface plasmon resonance (SPR) biosensor with high detection accuracy (DA) and quality factor (QF). Five layers of biosensor model is proposed using numerical simulation and graphical analysis. In this model, graphene and heterostructures of black phosphorus improve the sensitivity of biosensor. Silver (Ag) is coupled with prism CaF_2 for better reflectivity, since CaF_2 has the less refractive index (RI). The mentioned model gives the highest sensitivity for 633 nm wavelength p-polarized light is 263.51 °/RIU with higher detection accuracy (DA) 2.14 and quality factor (QF) 57.915 RIU^{-1}.**

Keywords—Surface plasmon resonance, Sensitivity, Biosensor, Detection accuracy, Quality factor.

I. INTRODUCTION

In this contemporary age, optical sensors become a preferable phenomenon due to its enriched accuracy, better sensitivity, low cost and inferior bulk [1]. These sensors are mainly classified into two types based on fiber optics and optical free space. Hence surface plasmon resonance (SPR) is considered as optical free space sensors [2]. Due to the accelerated progression of SPR optical technology, it has become a promising platform for photonics, optoelectronics and biosensors [3]. SPR is a technology which appears at the metal-dielectric surface at the time of light reflects at a SPR angle and due to metal (noble metal nanoparticles, Al nanoparticles, conventional semiconductors and 2D materials) surface conditions, there occurs an intense change of reflected light [4]. SPR based on optical sensor, which is one of the most impressible methods for the biomolecules detections by resonance and it is a physical process under reflection conditions [5]. In practical life, SPR sensors are convenient with various applications such as organic chemical sensing, gas detection [6], food quality maintaining, monitoring glucose, molecule detection, [7] environment monitoring and digital devices like as smartphone platforms or clinical diagnosis [8]. Especially, biosensors technology in the biomedical area has drawn a wide sight in research field such as haemoglobin concentration, toxicology and in microbiology [9]. In conventional SPR biosensors that had less competency of sensing small biomolecules like virus, bacteria and DNA for the low coalition onto the metal surface [10]. Therefore, this limitation can be exhibited by several gateway where mixing a core-shell of nanoparticle metal acting as the effective material over the conventional biosensors [11]. For enhancing the sensitivity of biosensor, hetero-structures and transition metal di-chalcogenides (TDMC) are more preferable with two dimensional (2D) material such as graphene layers. Those TDMC (MoS_2, $MoSe_2$, WS_2, WSe_2) have distinctive electrical and optical properties and graphene has high carrier mobility, better optical transparency, strength and flexibility [12]. Graphene aids biomolecule adsorption for their wide surface area and due to high π conjugation, preferring compatible dielectric top layers in SPR bio-sensing [13]. Wu et al. proposed a model where they had used graphene coated biosensor for enhancing sensitivity in SPR [14]. Since 2014, black phosphorus has become new promising 2D materials for its direct band gap, better carrier mobility, enrich electrical and optical properties [15]. Qiao et al. showed a few layer black phosphorus based model with polarized light for attaining high mobility [15]. Other materials such as Silver (Ag), Gold (Au), Cobalt (Co), Aluminum (Al), Copper (Cu) are also well known in SPR sensing model and have a better oxidization with good corrosion behavior in biosensors [14]. For better sensing model, BK7, CaF_2, SF10 prism glasses are very popular but among them, CaF_2 has the minimum refractive index which has a tremendous effect on sensitivity [16, 17]. Herminjard et al. circumstantiated the sensitivity of CaF_2 as a prism glass [17] and Shushama et al. presented graphene coated BK7 glass based SPR sensor [16].

Finally, in this paper we have proposed a modified and standard Kretschmann arrangement for a SPR based biosensor utilizing Silver (Ag) layers, black phosphorus (BP) and graphene layer with CaF_2 prism glass. Using MATLAB simulation with transfer matrix, we have been carried out highest sensitivity for graphene coated CaF_2 glass.

II. NUMERICAL MODELING AND DESIGN CONSIDERATION

For SPR technology, two configurations Otto and Kretschmann are preferred in most of the research but Kretschmann is more suitable for coupling prism with metal film layers[18]. The proposed final SPR biosensor is analyzed using five-layers structure with 633 nm wavelength and represented in Fig.1. This operating p-polarized 633 nm red laser is taken for developing the optical nonlinearity for specific frequency. Here we have applied the transfer matrix method and Fresnel equations founded on N layers structure which calculate the reflectivity of the reflected light [16] and due to having no approximation this method is the most efficient. Along with the Z-axis the thickness of the layers dk are considered and the refractive index (RI) of kth layer and dielectric constant are considered as ek and nk. The fields of tangential at first boundary $Z=Z_1=0$ are related with term of the final tangential boundary field at $Z=Z_{N-1}$ [19].

$$\begin{bmatrix} p_1 \\ Q_1 \end{bmatrix} = M \begin{bmatrix} P_{N-1} \\ Q_{N-1} \end{bmatrix} \tag{1}$$

Here P_1 and Q_1 are the element of electric and magnetic fields at the tangential boundary of the first layer and P_{N-1} and Q_{N-1} are the boundary field at N^{th} layer where characteristic matrix is M with Mij [14].

where,

$$M_{ij} = \left(\prod_{K=2}^{N-1} M_K\right) = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{2}$$

$$M_K = \begin{bmatrix} \cos\beta_K & (-i\sin\beta_K / q_K) \\ -iq_K \sin\beta_K & \cos\beta_K \end{bmatrix} \tag{3}$$

$$q_K = \sqrt{\frac{\mu_K}{e_K}}\cos\theta_K = \frac{\sqrt{(e_K - n_1^2 \sin^2\theta_1)}}{e_K} \tag{4}$$

$$\beta_K = \frac{2\pi d_K}{\lambda}\sqrt{(e_K - n_1^2 \sin^2\theta_1)} \tag{5}$$

The reflection intensity for p-polarized light with reflection coefficient and reflectivity are expressed as,

$$r_p = \frac{(M_{11} + M_{12qN})q_1 - (M_{21} + M_{22qN})}{(M_{11} + M_{12qN})q_1 + (M_{21} + M_{22qN})} \tag{6}$$

$$R_p = \left|r_p\right|^2 \tag{7}$$

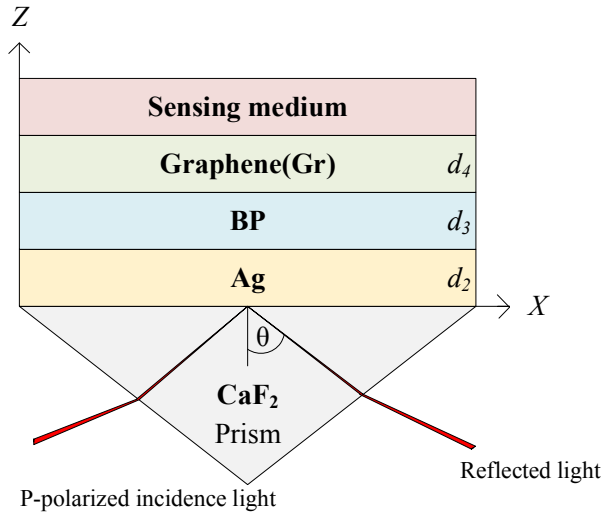


Fig. 1. Schematic structure of the proposed SPR biosensor.

In Fig. 1, proposed five layers SPR sensor has one prism glass of CaF_2 act as first layer and thickness of the metals are described as d_2, d_3 and d_4: silver (Ag) as the second layer, black phosphorus (BP) as the third layer, graphene (Gr) as fourth layer and at last sensing medium layer.

III. REFRACTIVE INDEX FOR VARIOUS LAYERS

In this proposed structure, CaF_2 glass has been applied instead of other glasses (BK7 or SF10) for enhancing the sensitivity with the combination of other consecutive layers because of holding a comparatively lower refractive index. The RI of the prism glass can be calculated by the according equation with λ = 633 nm [23(16)].

$$n_x = \sqrt{\left(\frac{1.03961212\lambda^2}{\lambda^2 - 0.00600069867} + \frac{0.231792344\lambda^2}{\lambda^2 - 0.0200179144} + \frac{1.03961212\lambda^2}{\lambda^2 - 103.560653} + 1\right)} \tag{8}$$

According to Drude-Lorentz model, refractive index of proposed metals can be calculated as follow [23(16)],

$$n_m = \sqrt{\left[1 - \frac{\lambda^2 \lambda_C}{\lambda_P^2(\lambda_C + i\lambda)}\right]} \tag{9}$$

Where λ_C and λ_P are the collision and plasma wavelengths. After the calculations, the refractive indexes are Ag = 0.8-4.2346i and BP = 3.5-0.01i respectively. The RI of graphene in complex form is developed as $n_4 = 3 + \frac{C}{3} i\lambda$ (λ = wavelength of vacuum and $C \approx 5.446$ μm^{-1}) [23]. Here the RI of sensing medium is $n_5 = 1.33$.

IV. SENSITIVITY AND RESONANCE ANGLE OF BIOSENSOR

Sensitivity (S) is the most vital element for evaluating the performance of a biosensor. It is determined by the ratio of change in the resonance angle ($\delta\theta$) and RI chance in sensing medium (Δn_s) [16].

$$S = \frac{\delta\theta}{\Delta n_s} \tag{10}$$

The resonance condition can be expressed as below. [16(11)],

$$\frac{\omega_o}{C} n_x \sin\theta_{SPR} = \frac{\omega_o}{C}\sqrt{\frac{\varepsilon_m n^2 d}{\varepsilon_m + n^2 d}} \tag{11}$$

Where $\sin\theta_{SPR}$ is the resonance angle in prism with RI n_x, Real part of metal permittivity is ε_m and n_d is RI of sensing medium or dielectric metal. and an incident light beam for propagation constant is ratio of ω_o and C (ω_o = frequency of light and C = speed of light).

The direction accuracy is figure out by the ratio of change in resonance angle ($\delta\theta$) and the full width at half maximum (FWHM) of reflectance curve follow. The FWHM should be low for high accuracy [20].

$$DA = \frac{\delta\theta}{FWHM} \tag{12}$$

Finally, quality factor (QF) is the sensitivity of biosensor to FWHM ratio [12].

$$QF = \frac{S}{FWHM} \tag{13}$$

V. Results and Analysis

In this paper, we have simulated various kind of response curve for the proposed structural SPR biosensor. For simulation, the transfer matrix procedure is used and the design is modified by the Kretschmann configuration. The highest sensitivity is obtained by considering first performance parameter then considering the relationship of RI with resonance angle and sensitivity and finally, optimized the different layers of the thickness of various metals we used. The five-layers model contains CaF_2 (prism), Ag, BP and graphene. The basic performance of SPR biosensor can be obtained by reflectivity, with respect to angle of incidence at changeable RI. When p-polarized light incidence on prism then the surface is excited of the SPR. Then a sharp dip is analyzed at an appointed angle. This phenomenon is happened due to a change in the RI for an accordance of molecule interaction in the sensing surface of biosensor [26].

A. Determination of performance parameter

In performance determination, the values of RI are used for simulation taken from the aforementioned section and at first considering the prism with Ag (28 nm) and got 222.97 °/RIU sensitivity in absence of BP and graphene layers. For this configuration, the quality factor (QF) and detection accuracy (DA) are obtained 68.61 RIU^{-1} and 2.538 respectfully. Now, considering Ag (28nm) and BP (2×0.53 nm) layers with the prism in absence of graphene layers and got the highest sensitivity of 258.10 °/RIU with QF = 52.14 RIU^{-1} and DA = 1.931. Here the thickness of BP layers is obtained by the equation L×0.53 nm. Where number of the layers is L . When only BP is absent and graphene layers are taken by L×0.34nm, then the sensitivity is decreased for the effect of graphene and obtained a sensitivity 237.63 °/RIU, QF = 65.16 RIU^{-1} and DA = 2.412.

Finally, the proposed model with Ag (28 nm), BP (2× 0.53 nm), graphene (1×0.34 nm) shows the highest sensitivity 263.51 °/RIU among all the model and QF = 65.16 RIU^{-1} and DA= 2.412. The highest sensitivity curve for the proposed model is figured out in Fig .2.

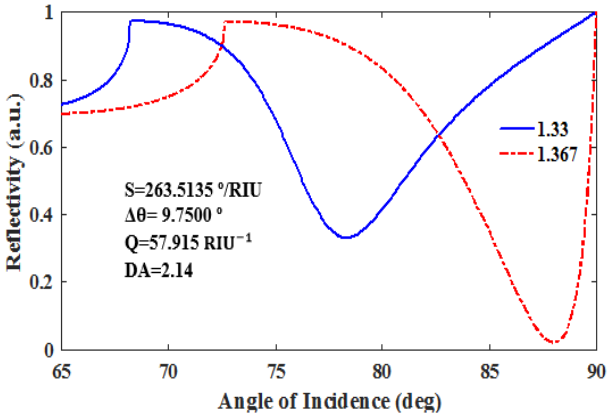


Fig. 2. Reflectivity versus angle of incident at highest sensitivity.

The circumstance above, points out the implication of change in RI for attaining highest sensitivity in the SPR biosensors. Table.1 reported the above discussion with a suitable comparison.

TABLE I. SENSITIVITY, QF AND DA COMPARISON FOR DIFFERENT CONDITIONAL MODEL.

Ag (1) nm	BP (0.53) nm	Gr (0.34) nm	Sensitivity (°/RIU)	QF= (S/FWHM)	DA=(δθ/FWHM)
28	0	0	222.973	68.614	2.538
28	2	0	258.10	52.143	1.931
28	0	1	237.63	65.162	2.412
28	2	1	263.51	57.915	2.144

B. RI with resonance angle and sensitivity

For the minimum change in RI, the resonance can be drastically changed. For analyzing better sensitivity, mostly high resonance angle is preferred. So the higher resonance angle, the better it responses in shortest change of RI. Fig. 3 shows the relation curve between resonance angle and different RI with metal configurations.

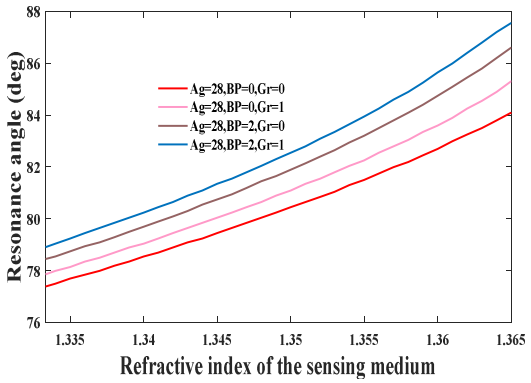


Fig. 3. Curve plotting for resonance angle versus RI of medium.

Finally, in Fig. 4 the relation curve for the sensitivity versus refractive index is plotted and the highest sensitivity is gained at 1.3670 for the proposed configuration. In simulation, the maximum RI change is considered in the range of 1.3300 to 1.3700.

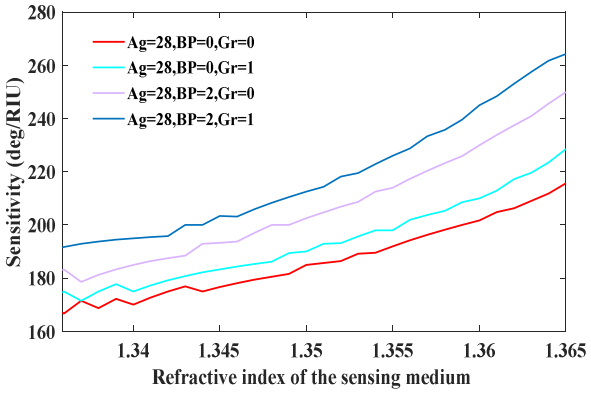
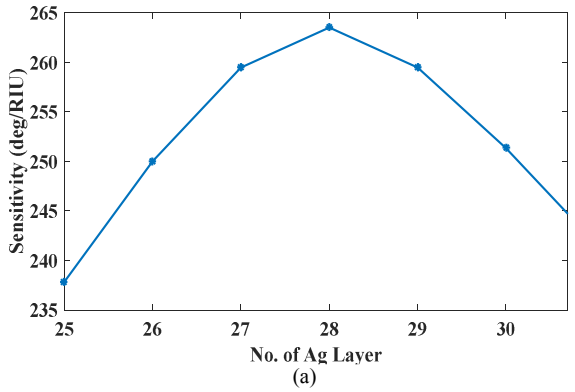


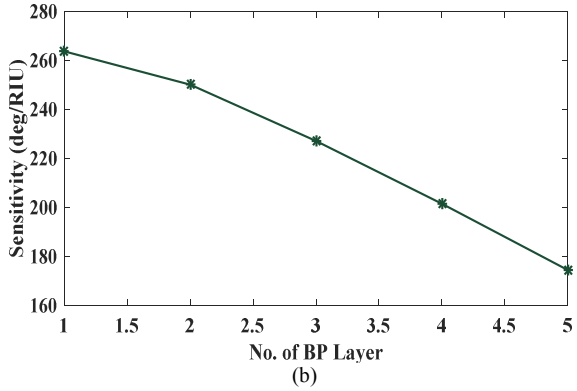
Fig. 4. Sensitivity versus RI of the medium.

C. *Optimization of different layers thickness*

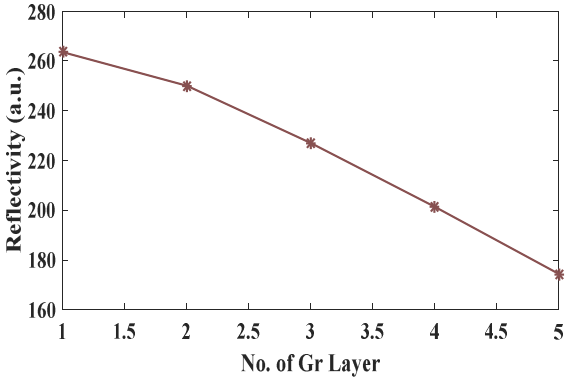
Obtaining the highest sensitivity with Ag layers, Fig. 5(a) is analyzed. Which represents the condition between sensitivity and thickness of Ag layers and for the thickness 28 nm the highest sensitivity 263.51 °/RIU is attained. In the Fig. 5(b), the curve for BP layers is plotted and for L = 2 layers highest sensitivity is achieved and by decreasing or increasing from L =2, the value of sensitivity is decreased. Finally, in Fig. 5(c) the layers optimization of graphene is detailed and for the increasing of graphene layers, sensitivity is decreased. So, the highest sensitivity 263.51 °/RIU is attained for the L =1 layer of graphene.



(a)



(b)



(c)

Fig. 5. Layers optimization for sensitivity versus (a) thickness of Ag layers (keeping BP and graphene constant) (b) thickness of BP layers (keeping Ag and graphene constant) (c) thickness of graphene layers (keeping Ag and BP constant).

TABLE II. COMPARISON OF SENSITIVITY FOR DIFFERENT SOUND PRISM BASED SPR BIOSENSORS.

References	Configuration	Wavelength (nm)	Sensitivity (°/RIU)
Ouyang (2016)[21]	Prism/Au/si/MoS_2	633	131.70
Lin (2016) [26]	Prism/Au/MoS_2/Au	633	182
Pal (2017) [20]	Prism/Au/Bp	633	180
Shushama (2017)[16]	Prism/Au/Si/MoS_2/Au/ Gr	633	210
Hamid (2019)[22]	Prism/Air/MoS_2/Nono composite/Mos_2/Gr	633	200
Proposed	Prism/Ag/BP/Gr	633	263.51

A polished comparison among the sensitivity is formed based on previous analogous works of biosensors for prism based SPR model which is precised in Table.2 From the table it is accomplished that the proposed biosensor ensures the highest sensitivity than the prior reported works.

VI. CONCLUSION

The proposed biosensor based on graphene coated layer with Ag and BP offers the highest sensitivity among the previous models. The highest sensitivity 263.51 °/RIU is obtained by optimizing prism, metals, the thickness of layers and RI. This highest sensitivity has also the highest DA about 2.14 and QF about 57.915 RIU^{-1}, which are the parameter for accomplishing the performance of the biosensor. So, a good SPR biosensor must have high sensitivity and better detection accuracy. In the conventional model, graphene is normally coupled with gold (Au) but in this model we have used BP to enhance the sensitivity response. A long range of RI is used for ensuring the perfect highest sensitivity of the biosensor and this biosensor can be used in multiple purpose of biomedical and optics such as DNA hybridization by proper orientation between Adenine (A) and Thymine (T) or between Guanine (G) and Cytosine (C) with match and miss-match monitoring.

REFERENCES

[1] R. Narayanaswamy and O.S. Wolfbeis, "Optical sensors: industrial environmental and diagnostic applications", *Springer Science & Business media*, vol. 1, 2013.

[2] W.M. Mukhtar, S. Shaari and P. S. Menon, "Propagation of surface plasmon waves at metal thin flim/air interface using modified optical waveguiding assembly", *Optoelectron Adv. Mat*. vol. 7, pp. 9-13, 2013.

[3] S. A. Maier, "Plasmonics: The promise of highly integrated optical devices," *IEEE Journal of selected topics in Quantum Electronics*, vol. 12, no. 6, pp. 1671-1677, 2006

[4] L. Song, L. Zhang, Y. Huang, L. Chen, G. Zhang, Z. et al., "Amplifying the signal of localized surface plasmon resonance sensing for the sensitive detection of Escherichia coli O157:H7". *Scientific Reports*, v.7, no. 1, pp. 3288, 2017.

[5] N. E. Agbor, J.P. Cresswell, M. C. Petty and A. P. Monkman, "An optical gas sensor based on polyaniline Langmuir-Blodgett films", *Sensors and Actuators B: Chemical*,1997, vol. 41, no. 1-3, pp. 137–141, 1997.

[6] S. Chakma, M.A. Khalek, B.K. Paul, K. Ahmed, M.R. Hasan, A.N. Bahar, "Gold-coated photonic crystal fiber biosensor based on surface plasmon resonance: design and analysis", *Sensing and bio-sensing research* . Res. 18, pp. 7–12, 2018.

[7] E.K. Akowuah, T. Gorman, H. Ademgil, S. Haxha, G.K. Robinson, J.V. Oliver, "Numerical analysis of a photonic crystal fiber for biosensing applications", *IEEE Journal of Quantum Electronics* vol. 48, no. 11, pp. 1403-1410, 2012.

[8] Y. Liu, Q. Liu, S. Chen, F. Cheng, H. Wang, and W. Peng, "Surface Plasmon Resonance Biosensor Based on Smart Phone Platforms", *Scientific Reports*, vol. 5, no. 1, pp. 12864, 2015.

[9] J. Ju, W. Liu, C. M. Perlaki, K. Chen, C. Feng, and Q. Liu, "Sustained and Cost Effective Silver Substrate for Surface Enhanced Raman Spectroscopy Based Biosensing", *Scientific Reports*, vol. 7, no. 1, pp. 6917, 2017.

[10] R. P. Liang, G. H. Yao, L. X. Fan and J. D. Qiu, "Magnetic Fe3O4@Au composite-enhanced surface plasmon resonance for ultrasensitive detection of magnetic nanoparticle-enriched α-fetoprotein. *Analytica chimica acta* 737 (2012), pp. 22-28.

[11] Widayanti, K. Abraha, and A. B. S. Utomo, "Computational Study of Sensitivity Enhancement in Surface Plasmon Resonance (SPR) Biosensors by Using the Inclusion of the Core-Shell for Biomaterial Sample Detection", *Biosensors*, vol. 8, no. 3, pp. 75, 2018.

[12] M. S. Rahman, M. R. Hasan, K. A. Rikta, and M. S. Anower, "A novel graphene coated surface plasmon resonance biosensor with tungsten disulfide (WS 2) for sensing DNA hybridization", *Optical Materials*, vol. 75,pp. 567–573, 2018.

[13] N. Maalouli, S. Szunerits, J. Pierre, E. Wijaya, Vilcot, R. Boukherroub, "Recent advances in the development of graphene-based surface plasmon resonance (SPR) interfaces", *Analytical and bioanalytical chemistry*, vol. 405, no. 5, pp. 1435-1443, 2013.

[14] L. Wu, H. S. Chu, W. S. Koh, E. P. Li, "Highly sensitive graphene biosensors based on surface plasmon resonance". *Optics Express*, vol. 18, no. 14, pp. 14395–14400, 2010.

[15] J. Qiao, X. Kong, Z.X. Hu, F. Yang, W. Ji, "High-mobility transport anisotropy and linear dichroism in few-layer black phosphorus", *Nature communications*, vol. 5, pp. 4475, 2014.

[16] K.N. Shushama, M. M. Rana, R. Ium, M. B. Hossain, "Sensitivity enhancement of graphene coated surface plasmon resonance biosensor," *Optical and Quantum Electronics*, vol. 49, no. 11, pp. 381, 2017.

[17] S. Herminjard, L. Sirigu, H. P. Herzig, E. Studemann, A. Crottini, J. P. Pellaux et al., "Surface Plasmon Resonance sensor showing enhanced sensitivity for CO_2 detection in the mid-infrared range", *Optics Express*, vol. 17, no. 1, pp. 293, 2009.

[18] Z. Lin, L. Jiang, L. Wu, J. Gou, X. Dai,Y. Xiang and D. Fan., "Tuning and sensitivity enhancement of surface plasmon resonance biosensor with graphene covered Au-MoS 2-Au films," *IEEE Photonics Journal*, vol. 8, no. 6, pp. 1-8, 2016.

[19] S. Zeng, S. Hu, J. Xia, T. Anderson, X.Q. Dinh et al., "Graphene–MoS2 hybrid nanostructures enhanced surface plasmon resonance biosensors", *Sensors and Actuators B: Chemical* 207 (2015), pp. 801-810.

[20] S. Pal, A. Verma, Y. K. Prajapati, and J. P. Saini, "Influence of black phosphorous on performance of surface plasmon resonance biosensor", *Optical and Quantum Electronics* vol. 49, no. 12, pp. 403, 2017.

[21] Q. Ouyang, S. Xeng, L. Jiang, L. Hong, G. Xu, X. Q. Dinh et al., "Sensitivity enhancement of transition metal dichalcogenides/silicon nanostructure-based surface plasmon resonance biosensor," *Scientific reports*, vol. 6, p. 28190, 2016.

[22] H.Vahed, and C.Nadri, "Sensitivity enhancement of SPR optical biosensor based on Graphene–MoS2 structure with nanocomposite layer". *Optical Materials*, vol. *88*, pp.161-166, 2019.

Design and Implementation of Solar Power Wireless Battery Charger

[1]Abu Shufian, [1]Md. Mominur Rahman, [2]Koushik Ahmed, [1]Riadul Islam, [1]Mahmodul Hasan, and [1]Toufiqul Islam
Department of Electrical and Electronic Engineering
[1]American International University-Bangladesh, Dhaka-1229, Bangladesh
[2]Ahsanullah University of Science and Technology, Dhaka-1208, Bangladesh
shufian2aiub@gmail.com, mominur18aiub@gmail.com, koushik5699@gmail.com, riadul999@gmail.com, mahmodul033@gmail.com, and toufiq.sajib@gmail.com

***Abstract*—This paper deals with wireless power transmission technology. A battery of an electronic device will be charged wirelessly. The solar panel converts the sun light into electrical energy. Power from a solar panel is sent through a transmitter circuit and received by a receiver circuit wirelessly based on Faraday's law of induction. As magnetic resonance coupling is more efficient and useful than magnetic inductive coupling, in this paper, magnetic resonant coupling is used. There are a transmitter and receiver coil. The transmitter coil is driven at the same frequency of the resonant frequency of the receiver coil. For those reasons, the LC tuned circuit has been built to maintain the same frequency. As power is transmitted wirelessly, some power is wasted. As a result, efficiency is decreased to only 11 %. After all, it will be a future charging solution in the modern world. With the help of wireless battery charger technology, power can be supplied to the electric cars, drones, hospital's equipment, and smartphones. This device also helps to reduce the necessity of using wires.**

Keywords— copper coil, inductive coupling, receiver coil, resonant, solar power, transmitter coil, wireless battery charger.

I. Introduction

Solar energy is a very environment-friendly power source and one type of renewable energy. Besides that, the charge controller makes it more efficient for use as it gives the necessary maximum power to the battery or appliances. Solar panel uses photovoltaic cells that convert light energy from the source of light into electricity. This renewable energy can be used for various purposes, i.e. to run the home appliances, to operate machines and motors. In this study, solar power is used to charge batteries wirelessly. Wireless power transmission system follows the rules of inductive coupling by Faraday's laws. Where transmitting power is converted into high frequencies AC current that is sent to a transmitter coil for which magnetic field is created. The same magnetic field produces current within the receiver coil. This current is converted into usable DC current. Normally wired charger charges batteries but some devices such as, electric toothbrush which is not suitable for a wired charger. Mobile phones can also be charged by wireless, charger station can offer this type of feature.

The solar or photovoltaic system is a very effective energy resource that has established a great impact on the society and industries. In remote areas, it is quite difficult for people to charge their devices because of unavailability of electricity. With this proposed system, it is possible to charge any device without electricity. In this study, an inductive coupling has been used which creates a magnetic field between two objects and transfers energy from source to receiver. With this energy, the battery can be charged without any physical connection.

II. Literature Review with in-Depth Investigation

Transmitting power through a wireless system was invented by Faraday. His research was about current conduction in the wire. If current flows through one wire of a conductor then another wire near to it will gain some current as well. For this, the second wire has to be placed very near to the current carrying wire. This principle was developed by Tesla who built towers to transmit power through the entire world. But the project was unsuccessful because his sponsor discontinued providing financial support for the transmission. The reason for this was, the experiment results were not satisfactory [1].

Many research work has been done for wireless power transmission. A wireless sensor network along with a charging car can increase the lifespan of a network with the help of wireless power conversion technology. The wireless energy can be radiated in every direction as well as a specific direction. So, the charging model can be divided into two model, point-to-point and point-to-multipoint [2]. An ideal case was first mentioned by assuming zero traveling time for the charger. A provably near-optimal solution was obtained based on the discretization and logic point representation techniques. Finding the shortest Hamiltonian cycle helps to discover the moving path of the original problem.A feasible solution is derived with this moving path [3]. Research had been done in the same field and the proposed device's efficiency was very high [4]. In a project named "Automatic wireless mobile charger", the researchers tried to develop a better device to deliver power wirelessly [5]. One more study states that mobile will be charged automatically and wirelessly through radio frequency harvesting [6]. The demand for fossil fuel is increasing day by day. An electric vehicle, developed by wireless technology can lessen the use of fossil fuel [7]. A wireless remote control sysem is applied for protection of the linear generator where the electrical power is supplied from a solar panel connected with a battery and a charger [8]. Another automatic portable grass cutter machine is proposed in which a solar panel, a charger and a battery are used for power supply [9].

In recent years, it has been proved that electricity can be transferred at a distance of about two meters. The efficiency of this technology is 40% where 60% power was wasted. Later, the efficiency has been improved [5].

III. Methodology and Modeling

Wireless power transmission means transferring power without physical wire. Transmission is possible for inductive coupling. The system implementation is done in two parts, one is transmitting side and another one is receiving side.

Firstly, solar array takes energy from the sun. Then, this DC power goes to maximum power point tracking (MPPT). Secondly, MPPT looks at the output of the panels and compress it to the battery voltage. Actually, it figures out what is the best power that the panel can put out to charge the battery and converts received voltage to the best voltage to get maximum ampere into the battery. Thirdly, maximum DC power goes to the wireless charger circuit. In there, wireless battery charging uses an inductive or magnetic field between two objects. At the receiver end, a typical battery is connected. Finally, the battery will be charged. Part by part implementation has been described below.

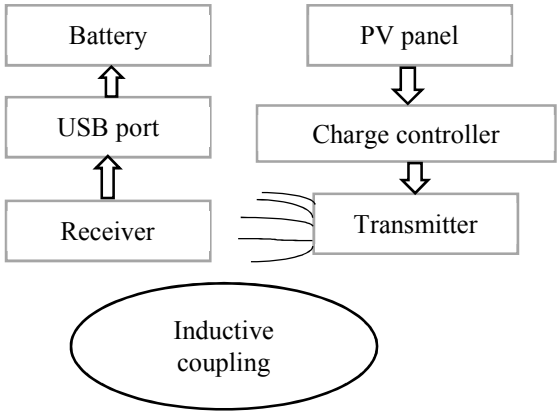


Fig. 1. Block diagram for wireless battery charger.

A. *Transmitter Side*

Transmitter side is comprised of IC 555 timer, general purpose transistor BC547, n-channel MOSFET IRF540N, some capacitors, some resistors, and a tuned LC circuit. This tuned LC circuit is also called the tank circuit.

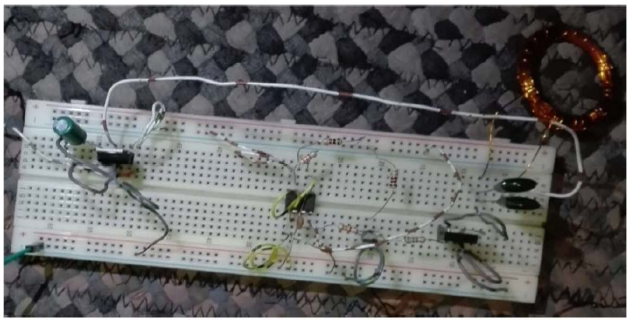


Fig. 2. Design of a transmitter circuit for this system.

B. *Solar Setup for Power Supply*

During the hardware implementation portion, a storage device is required to store the produced energy by photovoltaic cells.



(a) (b)

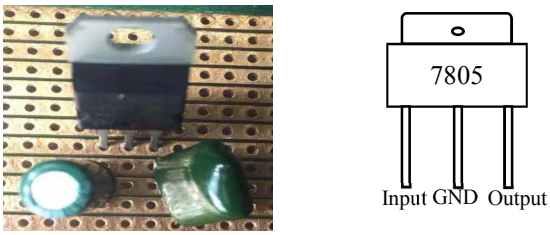
Fig. 3. (a) Measurement of open circuit voltage, (b) measurement of short circuit current.

A battery is a common storage device which is used to store electrical energy produced by photovoltaic cells. To charge a battery by using a solar PV, a charge controller is an important device. Therefore, it will be necessary to connect a charge controller in series between the solar panels and the battery.

C. *Setup of Voltage Regulator*

The voltage regulator IC 7805 maintains a constant voltage output for fluctuating of an input voltage and the output voltage is +5 volt. It has 3 pins. Pin number 1 is for input, pin number 3 is for output and pin number 2 is for ground. In this system, pin number 1 is connected with the source as well as a capacitor C5. It is also connected with the tuned LC circuit.

Pin number 2 is connected with capacitor C5 and C6 and then connected with the ground. Pin number 3 is connected with capacitor C6 and also connected with IC 555 timer through its pin number 4.



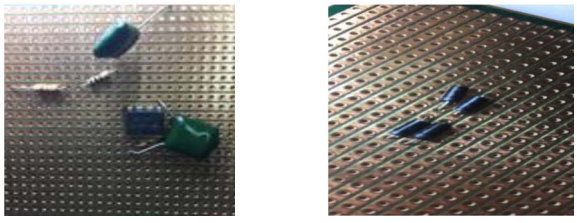
(a) (b)

Fig. 4. (a) Voltage regulator setup, (b) voltage regulator IC.

D. *Setup of IC 555 Timer*

This is an integrated circuit which is used for pulse generation purpose. There are three types of moods of a timer IC, i.e. Astable, Monostable, Bistable. In this system, Astable mood is used. It puts out a continuous spectrum of rectangular pulses. Each pulse has a specified frequency.

In this system, pin 1 is grounded. Triggering pin 2 is shorted with pin 6 and also connected with a capacitor C6. A transistor is connected with resistor R3 and pin 3. Pin 4 is connected with pin 3 of voltage regulator IC 7805. A grounded capacitor C4 is connected with controlling voltage pin 5. The threshold pin 6 and discharging pin 7 is connected through a single register R2. Pin 8 is shorted with 4.



(a) (b)

Fig. 5. (a) Setup of IC 555 timer, (b) rectifier setup of the system.

E. *Tuned LC Circuit* (*transmitting end*)

The tuned LC circuit is made with capacitors C1 and C2. An inductor is parallelly connected with C1 and C2. This inductor is made with 45 turns of copper wire. The diameter of copper coil is 7.3 cm. This circuit acts as a load circuit. It determines the frequency of oscillation. For this tuned circuit, both the transmitter and receiver circuit is inductively coupled. A voltage is induced at both transmitter and receiver circuit.

F. *Receiver End*

Transmitted oscillating induced voltage from the magnetic field is received here. The induced voltage is used to charge the battery. In the receiving part, some equipment are used such as LC tuned circuit, a current regulator IC MC34063, Schottky diode 1N5819, and some passive equipment.

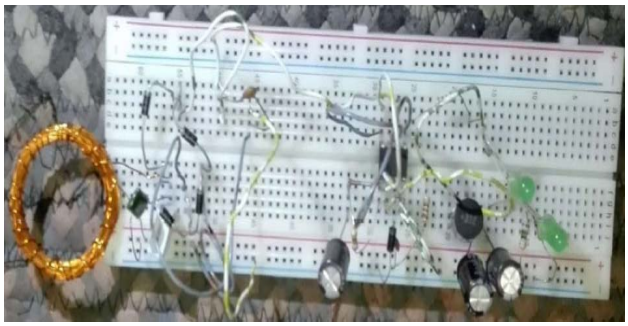


Fig. 6. Receiver side of the proposed system.

G. *Tuned LC Circuit* (*receiving end*)

An inductor made of 50 turns of copper wire is parallelly connected with two capacitors C7 and C8. Rotating magnetic flux comes over this copper coil. This coil is assembled with 7.3 cm in diameter.



(a) (b)

Fig. 7. (a) Copper coil for transmitting end, (b) Copper coil for receiving end.

H. *Setup of Bridge Rectifier*

A bridge rectifier is made with 4 diodes. Tuned LC circuit is connected with the rectifier. This bridge rectifier extracts DC current from the rotating magnetic field. The rectifier is connected with C9 capacitor. Here, DC voltage is ripple free for that capacitor. Rectifier setup has been shown in Fig. 5(b).

I. *Setup of IC MC34063*

This IC is a buck/boost converter. In this system, it works as a buck converter.

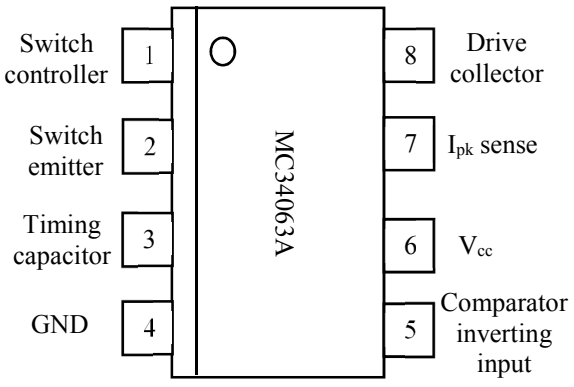


Fig. 8. Buck-boost converter IC.

Buck converter helps to maintain the adjustable voltage and current. Switch collector pin 1 is shorted with driver collector pin 8 and I_{pk} pin 7, then the common point of them is connected with resistor R8. Switch emitter pin 2 is connected with a diode and an inductor. Then, the other parts of the inductor are connected with two capacitor C12 and C13. A load is connected with this common point of two capacitors.

IV. Software and Hardware Interface

Transmitter side is assembled with a solar panel, two IC, transistors, and some other passive devices. The receiver side is assembled with a rectifier, one IC and some passive components. The range of source voltage should be 06 volt to 20 volts. And current should be maintained up to 1 ampere.

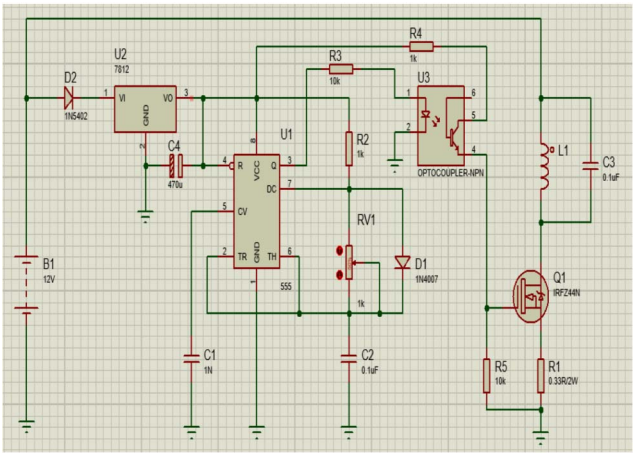


Fig. 9. Circuit diagram for transmitter side.

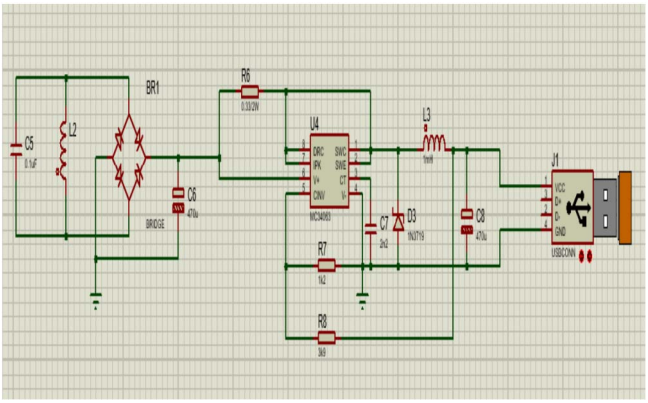


Fig. 10. Circuit diagram for receiver side.

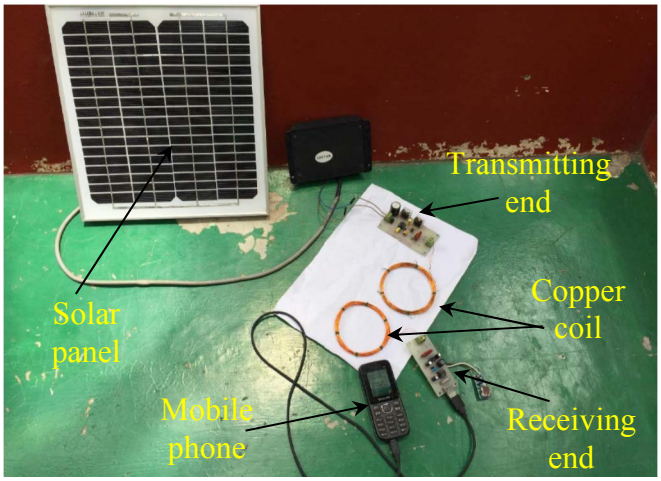


Fig. 11. Final setup for solar power wireless battery charger.

Table I. COMPONENTS OF SOLAR POWER WIRELESS BATTERY CHARGER.

Components	Amount
ICMC34063	1 pcs
Heat sink	2 pcs
3.3 resistor	10 pcs
Copper wire	25 Gauge
Solar panel	1 pcs
1n 4007 diode	4 pcs
IC 555 timer	1 pcs
Capacitor	8 pcs
Inductor	3 pcs
PCB	2 pcs
Schottky diode	1 pcs
BC547 (General purpose NPN transistor)	1 pcs
LM7805 (Series voltage regulator, 5V)	1 pcs
IRF540N (N-channel MOSFET)	1 pcs

V. RESULTS ANALYSIS

This proposed model is related to wireless power transmission. Power transmission happens because of the electromagnetic field. Transmitter side receives the input solar power and converts it into an oscillating magnetic field with the help of copper coil. At the receiver end, this oscillating magnetic field creates an induced voltage and finally, the battery of a device is charged.

The charging operation of a mobile phone is shown in Fig. 12.



Fig. 12. Charging a mobile phone.

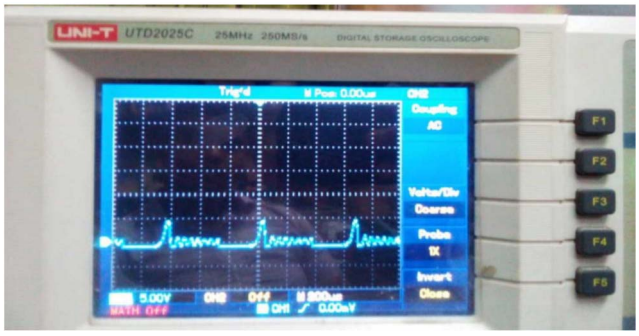


Fig. 13. Wave shape for the input voltage.

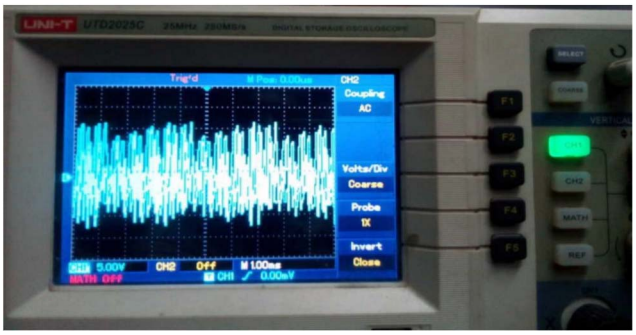


Fig. 14. Wave shape for the output voltage.

VI. CALCULATION

As the IC 555 timer is wired in "astable multivibrator" so, pulse oscillation has to calculate. It is known that there are three modes for IC 555 timer, i.e. bistable, monostable, astable. Astable multivibrator puts out a continuous stream of rectangular pulses.

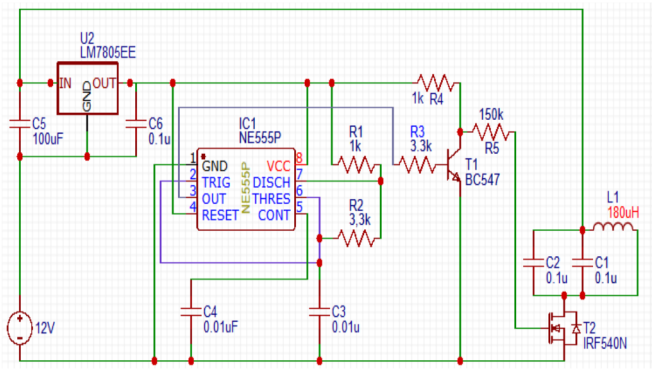


Fig. 15. Transmitter Circuit of the system.

Now, high time or charging time for each pulse

$$= \ln(2) \times C3 \times (R1 + R2) \quad (1)$$

$$= 0.693 \times 0.01 \times 10^{-6} \times (1000+3300)$$

$$= 29.79\mu \text{ sec.}$$

The low time or discharging time from each pulse

$$= \ln(2) \times C3 \times (R1 + R2) \quad (2)$$

$$= 0.693 \times 0.01 \times 10^{-6} \times (3300)$$

$$= 22.87\mu \text{ sec.}$$

The oscillating frequency of 555 timer IC or frequency of pulse stream,

$$f = \frac{1}{\ln(2) \times C3 \times (R1 + R2)} \quad (3)$$

$$= 18.94 \text{ kHz.}$$

L-C tuned frequency for the transmitter,

$$Ft = \frac{1}{2\pi\sqrt{LC}} \quad (4)$$

$$= \frac{1}{2\pi\sqrt{180 \times 10^{-6} \times 0.2 \times 10^{-6}}}$$

$$= 26 \text{ kHz.}$$

L-C tuned frequency for the receiver,

$$Fr = \frac{1}{2\pi\sqrt{LC}} \quad (5)$$

$$= \frac{1}{2\pi\sqrt{180\times10^{-6}\times0.2\times10^{-6}}}$$

= 26 kHz.

Output voltage = 3.8 Volt and Output current = 0.35 A

Output power =Voltage × Current

= 3.8 V × 0.35 A

= 1.3312 watt

Input Voltage = 12 V and Input Current = 1 A

Input power = Input Voltage × Input Current

= 12 V × 1 A

= 12 watt

Now, Efficiency = Output power / Input power

= 1.3312 / 12

= 0.11

= 11%

In this study, the efficieny is low (11%) because power loss occurs due to the power transferring process through some passive devices.

VII. Advantages Over the Traditional Model

The handset is usually charged these days by using electricity, generated and supplied from different power stations. But this proposed system does not need any connection with this type of conventional power sources to charge a handset. The handset will be charged through solar power. So, additional electricity is no longer needed to charge a phone's battery. In addition, this wireless system does not require any wire to connect with the device for charging. Handset just needs to drop onto the charging dock and things start working without having to plug anything into the phone. This is significantly easier than traditional wired charging method. So, plugging and unplugging a cable is no longer necessary in every time to charge a phone.

VIII. Limitations Of the System

As we know, nothing is perfect and long lasting in today's world. Same goes with wireless charging. One of the main reason behind unsuccessful of wireless charging on extended large distance. The charging transmitter still needs to be connected to a power source so we cannot be moved everywhere. Certain wireless chargers cannot reach the same level of efficiency as traditional charging, so the process tends to be slower. As our project is solar based so sometimes it is very difficult to produce sufficient power as needed. In addition, the heat generation of wireless charging technologies is generally higher with the conventional charging method.

IX. Future Enhancement

Basically, the wireless solar charger is designed to charge the mobile battery. But in the future, by making some modifications, it can be used for different portable devices, for example, Laptop, iPad, digital camera, electric vehicles. With the development of the extended distance wireless charging, it is possible to charge the devices by walking outside on the street. It will work like a hotspot area similar to a Wi-Fi hotspot, the device gets charging while users are walking or speaking. Now, wireless charging would allow the vehicles to power up. This wireless charging vehicle would not use plugs or charging cords. Drivers would park their electric vehicles over a wireless energy source. Then the system will automatically transfer power to the battery charger on the vehicle. These electrochemical high rate batteries can be used for wireless charging vehicles. So within very short times vehicles will be charged.

X. Conclusion

Solar power wireless charger can charge the battery very efficiently without any wire. This wireless charging system is ideal to use for many portable devices such as mobile phones and other wireless applications. This study is based on solar power and wireless technology. So during load shedding and traveling, electronic devices can be easily charged without any wire and charger. This system is actually very efficient and environment friendly.

References

[1] Sourabh Pawade, Tushar Nimje, and Dipti Diwase, "Goodbye Wires: Approach to Wireless Power Transmission", IJETAE 2012; Volume 2, Issue 4.

[2] L. Xie, Y. Shi, Y. Hou and H. Sherali, "Making Sensor Networks Immortal: An Energy-Renewal Approach With Wireless Power Transfer", IEEE/ACM Transactions on Networking, vol. 20, no. 6, pp. 1748-1761, 2012. Available: 10.1109/tnet.2012.2185831

[3] L. Xie, Y. Shi, Y. Hou, W. Lou, and H. Sherali, "On traveling path and related problems for a mobile station in a rechargeable sensor network", Proceedings of the fourteenth ACM international symposium on Mobile ad hoc networking and computing - MobiHoc '13, 2013. Available: 10.1145/2491288.2491291

[4] Shilpa S, Manasa. H.M, Sadaf shahana Mariam and Mr. Keerthi Kumar M "Wireless Battery Charger", IJETAE 2016; Volume 35, Number 3.

[5] Abhishek Gupta, Ankush Tandon, Akhilesh Jeriya, Naveen Tanwar, Ramesh Kumar, Dinesh Kumar, "Automatic Wireless Mobile Charger", IJEER (ISSN 2348-6988) , Vol. 4, Issue 1, pp: (113-119), Month: January - March 2016, Available at: www.researchpublish.com

[6] Daniel W. Harrist, University of Pittsburgh,2004, "Wireless Battery Charging System using Radio Frequebcy Energy Harvesting".

[7] Chan Kam and Tony Tsang, "A Experiment Method of Wireless Power Transfer for Charging Devices", American Journal of Engineering Research (AJER), e-ISSN: 2320-0847 p-ISSN : 2320-0936, Volume-5, Issue-11, pp-110-123, www.ajer.org.

[8] O. Farrok, M. R. Habib, M. R. Afefin and A. Habib, "A new protection system of linear generators used for converting oceanic wave energy into electricity," IEEE Region 10 Humanitarian Technol. Conf., Dhaka, Bangladesh, Dec. 21–23, 2017, pp. 755–758.

[9] M. R. Habib, N. Khan, K. Ahmed, M. R. Kiran, M. A. Habib, M. T. Hasan, O. Farrok, "Two degree-of-freedom PID controller based an automatic solar power-driven grass cutting machine," Int. Conf. Comput., Commun., Chemical, Mater. Electron. Eng., Jul. 11-12, 2019 (Submitted).

Offshore Wind Energy Estimation in the Bay of Bengal with Satellite Wind Measurement

Navila Rahman Nadi
Lecturer,Dept. of EEE
Independent University, Bangladesh.
Dhaka, Bangladesh.
navila@iub.edu.

Ferhat Bingöl
Assist. Prof. ,Energy Systems Engineering
Izmir Institute of Technology
Izmir, Turkey
ferhatbingol@iyte.edu.tr

Merete Badger
Senior Scientist, Dept. of Wind Energy,
Technical University of Denmark(DTU)
Roskilde, Denmark
mebc@dtu.dk

***Abstract*— The objective of this paper is to obtain appropriate offshore location in the Bay of Bengal, Bangladesh for further development of wind energy. Through analyzing the previous published works, no offshore wind energy estimation has been found related to the Bay of Bengal. Therefore, this study can be claimed as the first footstep towards offshore wind energy analysis for this region. Generally, it is difficult to find offshore wind data relative to the wind turbine hub heights, thus a starting point is necessary to identify the possible wind power density of the region. In such scenario, Synthetic Aperture radars (SAR) have proven useful in previous studies. In this study, SAR based dataset- ENVISAT ASAR has been used for Wind Atlas generation of the Bay of Bengal. Furthermore, a comparative study has been performed with Global Wind Atlas (GWA) to determine a potential offshore wind farm production in a reasonable location at the bay. The annual energy production of that offshore windfarm has been analyzed by combining SAR, GWA and ASCAT datasets. Through ASAR based Wind Atlas and GWA comparison, some differences have been found where there are less samples from the ASAR datasets. Thus, Weibull statistical analysis are performed to have a better Weibull fitting and accurate estimation of Annual Energy production (AEP). The study summarizes that, satellite datasets can be a very useful method to detect potential zone if compared with any long time statistical result and bathymetry data together.**

Keywords— Wind Atlas, Satellite wind measurements, Weibull distribution, ASCAT ,ASAR, Offshore

I. INTRODUCTION

Increasing demand of offshore wind parks in various countries has promoted wind resource mapping over the sea. The original method to estimate the wind energy resource is by installing a meteorological mast at one or more specific sites for minimum one year. This mostly requires more time as well as becomes costly and this drawback leads to search of another alternative method which gives indications of wind resources and help in proper decision making of feasibility studies for the prospective wind parks. In such scenario, satellite remote sensing is useful. Execution of these satellite based images in order to derive wind speed in the early stages of wind farm planning is becoming praiseworthy because of its spatial and time coverage along with cost effectiveness [1]. This useful method can predict perfect area of higher wind resources which leads to the planning of feasibility studies and provide a faithful supplement to the traditional assessments [3].

The main study area is chosen as the Bay of Bengal, which is considered as the largest Bay of the world which comprises around 2090 km of length and 1610 km of width. The total surface area of Bay of Bengal is around 2 million km^2 and it is bounded by Sri Lanka and India in the west side, Myanmar and Andaman and Nicobar islands in the east side and Bangladesh in the North side. It has a high depth of around 5 km and average depth of 2600 m which can be an impeccable study area for offshore floating wind turbines [2]. Broad wind analysis has not been performed yet in this bay to find out the wind energy resources, only few studies have been done in the Indian coasts which just covers a specific western zone of the Bay of Bengal. Also, no wind atlas generation for larger domain and wind energy estimation has been performed in the Bay of Bengal which is the main concentration of this paper.

In our study, the obtained datasets were compared with the Global Wind Atlas (GWA) datasets. GWA provides long-term high resolution wind climatology at 50, 100, 200m hub heights for the whole world including onshore and 30 km from offshore as a free access online [12]. This wind atlas provides wind resource data at 1 km resolution and contains data at a higher spatial resolution of 250 m [12]. The information provided in the Atlas along with further specific regional studies, supports to identify the prospective areas for additional preliminary assessment of technical possibilities.

II. SATELLITE TECHNOLOGIES FOR OFFSHORE WIND

In this paper, two types of satellites were analyzed which are: Synthetic Aperture Radar (SAR) and Scatterometer. The studied datasets were attained from SAR methodology based Advanced Synthetic Aperture Radar (ENVISAT ASAR) and from Advanced Scatterometer (ASCAT) satellites.

A. Synthetic Aperture Radar (SAR) technology

SAR is a remote sensing technique which transmits microwave pulses of energy to illuminate an area of the Earth surface. The pulses gets scattered because of the surface properties and a portion returns to the instrument with a change in amplitude, polarization, travel time and phase [3]. All these backscattered signals for each foot print area stores in the satellite to process a two dimensional image of the surface. SAR has higher spatial resolution (around 1km) and preferred more for coastal wind energy measurements [5].The data acquisition from the SAR pulses can be performed in day or night even if it is cloudy or light rainy (Christiansen, 2006). SAR signals are electromagnetic waves with a polarization determined by the orientation of the electric field [3]. Some upgraded SAR systems have the capability of transmitting and receiving both polarization simultaneously, and deliver quad observable SAR images [4].

ENVISAT ASAR (Environmental satellite based Advanced Synthetic Aperture Radar) satellite is based on SAR technology and it is actually an Environmental satellite capable of scanning the surface in one of the four polarization combinations. [4]. The mission was later followed up by the

Sentinel-1 A/B mission. It is being established by European Space Agency (ESA) in 1 March,2002 until 9 May, 2012, when ESA lost the contact with the satellite [5]. The satellite take images and send to the base station down in earth to preserve. The spatial resolution of the images is around 150 m and swath mode around 400 km. The total incidence angle coverage is 15-45 degree [3] . In the scenario of ocean winds, the average of around 1km resolution is processed for SAR images in a purpose to lessen the natural speckle noise and modulation error [9]. In WSM (wide swath mode) of SAR data acquisition, it can map a point at mid-latitude (55°N) around 15 times per month [9]. The ENVISAT ASAR dataset is freely accessible from the European Space Agency (ESA) archive from where the required data can be downloaded and used for further scientific research studies.

B. Scattterometer Technology

Scatterometer is a real aperture radar that also uses microwave technology are sun – synchronous polar orbiting satellites similar to SAR. These kind of satellites have a fixed local time of observations for ascending and descending mode. Scatterometers operate in the radar scale of C-band (around 5 GHz) or in Ku-band (around 13 GHz)[5]. One of the prime differences in between Scatterometer and SAR is that, Scatterometters take multiple (3 to 4) views of the surface from the antennae while SAR takes only one directional view. Due to this higher area coverage capability Scatterometer's observation rate is higher than SAR [5].

In October 2006 the Advanced Scatterometer (ASCAT) was launched on the EUMETSAT (European Organization for the Exploitation of Meteorological Satellites) MetOp-A which is Meteorological Operational satellite program (MetOp) and became completely operational from May, 2007 until today[6]. From September 2012, MetOp-B became operative, whereas both of these satellites has same ASCAT instruments.[.Being a C-Band scatterometer, ASCAT operates two times a day (day and night) through its 3 vertically polarized antennas transmitting pulses at 5.255 GHz [7]. ASCAT has a double swath as the fan-beam antennae is slanted towards 45, 90, and 135° with satellite track where the extended side of antennae possesses two swaths each consisting of 550 Km wide separated by a gap of about 670 Km nadir gap. [14] Having two swaths ASCAT can provide mapping of all locations at a midpoint (55°N) around 40 times per month [9]. Comparing with ENVISAT ASAR it has more frequent imaging as ASAR maps 15 times per month for the same midpoint.

III. DATA ANALYSIS OF THE BAY OF BENGAL

The fundamental of this paper is to create a wind atlas for the Bay of Bengal by using ENVISAT ASAR datasets and through a comparative analysis with Global Wind Atlas (GWA), finding a suitable location for wind farm estimation. In this paper, the wind atlas methodology has been followed to create a wind atlas. This step by step process has been displayed in a flow chart below. The interest area of our study locates in UTM 45N zone, at 82° to 95° Eastern longitude and 16.5° to 23° Northern latitude.

A. ASAR datasets

The Bay of Bengal has been mapped through ASAR on-board by European Space Agency satellite ENVISAT from 2006 until 2012. A total number of 466 Wide Swath Mode (WSM) scenes from 2002 until 2012 are collected from the ESA's ordering system EOLI-SA (Earth Observation Link) with special request by DTU. Each of those WSM scenes cover 400 x 400 km.

After receiving the WSM scenes, the wind retrieval offshore wind maps were produced. This whole process of the wind retrieval is done with a patented software called by ANSWRS (Applied Physics Laboratory National Oceanic and Atmospheric Administration (NOAA) SAR Wind Retrieval System)[3]. ANSWRS is a well effective software for wind retrieval from the ASAR WSM scenes which has been used in all previous SAR based offshore wind atlas generation [3] [8][9] [10] [11]. The SAR wind directional data was attained at 1 degree latitude and longitude resolution from the US Navy Operational Global Atmospheric Prediction System (NOGAPS)[5]. NOGAPS is modeled as 6-hourly wind direction provider which were interpolated spatially to be suitable for usage with the SAR images as previously it has been mentioned that, SAR only takes one directional images. So it cannot detect the wind direction by itself. For geophysical model function CMOD-5.n is being used. CMOD-5.n is an updated version of the CMOD5 which provides neutral condition of wind speed at 10m [5]. In the below Figure 2, it represents two ENVISAT ASAR WSM wind scenes on the date of 26 June, 2010 and 9 March 2009 respectively in ascending and descending formats.

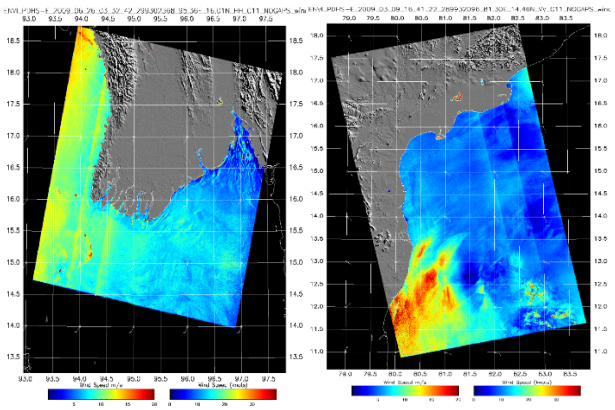


Figure 1: ENVISAT ASAR retrieved wind scenes of 26 June 2010 and 9 March 2009 at 16.41:22 and at 03.32:42 UTC in the Bay of Bengal accordingly.

The satellite wind retrieved fields received from the DTU were merged and combined for wind resource mapping using the DTU Wind tool S-WAsP. In S-WAsP, a second moment fitting for the Weibull scale (A) and shape (k) parameters are used to fit a distribution to all data and estimate the two Weibull parameters.

B. ASCAT datasets

In this study, the similar way, ASCAT datasets has been received and further retrieved to achieve wind scenes for that specific location. National Oceanic and Atmospheric Administration (NOAA)'s weather forecasting and warning mission operationally supports the production and utilization of ASCAT ocean surface vector wind products[15]. Royal

Netherlands Meteorological Institute (KINMI) has developed the ASCAT wind data processor for the standard ASCAT wind by using GMF CMOD5.N [15]. KNMI produces the gridded version (cell size of 12.5 km) product of ASCAT wind speed and direction within EUMETSAT OSI SAF. (SAF, 2013). In our study, ASCAT observations are extracted from the ASCAT Level 3 Coastal Ocean Surface Wind Vector product (MyOcean). L3 daily gridded wind product is available from the MyOcean catalogue after registration at their site for the defined location.

C. GWA datasets

For this paper, the global wind dataset were collected in modern Era-Retrospective Analysis for Research and Applications (MERRA) format from the free access internet as .lib files for definite co-ordinates and then matched those nodal points with the obtained SAR data locations. In the concerned area, the sample nodal locations of SAR .lib files were matched with the downloaded GWA location defined .lib files for 10m height.

IV. STATISTICAL ANALYSIS

For analyzing the wind speed data, the two parameter Weibull Probability distribution function (PDF) is most appropriate and acknowledged. This PDF provides an enhanced fitting to predict high accurate monthly distribution of wind speed. The Weibull Probability function is denoted in [16] as:

$$f(v) = (\frac{k}{A})(\frac{v_i}{A})^{k-1}\exp[-(\frac{v_i}{A})]^k \quad (1)$$

where $f(v)$ is denoted as the probability of observed wind speed, A is the Weibull scale parameter and k is the unit less Weibull shape parameter, v_i is wind speed. The two Weibull parameters and the average wind speed are related by [16]

$$\bar{v} = A.\Gamma\,(1+\frac{1}{k}) \quad (2)$$

By using the Weibull parameters, gamma function Γ and the air density ρ (~1.245 g m−3) the wind power density is predicted in [16] as:

$$E = \frac{1}{2}\rho A_w^3.\Gamma\left(1+\frac{3}{k_w}\right) \quad (3)$$

where $\bar{v}$ is the average wind speed, E is the power density and Γ is the gamma function. In this study, four methods of statistical analysis were compared to determine the shape and scale parameters of Weibull wind speed distribution for wind energy resource analysis, these are:

A. Empirical Method

Empirical method (EM) is a comparison set using the standard deviation and mean wind speed. It is considered as a distinctive situation of Method of moments (MOM). Through EM the parameters of Weibull distribution can be found from [16] as:

$$k = (\frac{\sigma}{\bar{v}})^{-1.086} \quad (4)$$

$$A = \frac{\bar{v}}{\Gamma\,(1+\frac{1}{k})} \quad (5)$$

where $\bar{v}$ is the average wind speed, σ is the standard deviation and Γ is the gamma function.

B. Method of Moments

In estimating Weibull distribution parameters Method of moments (MOM) is widely used. The equation format is dependent on the arithmetical iteration of mean wind speed ($\bar{v}$) and standard deviations (σ). The equations [16] are followed as:

$$k = (\frac{0.9874}{\frac{\sigma}{\bar{v}}})^{1.0983} \quad (6)$$

$$A = \frac{\bar{v}}{\Gamma\,(1+\frac{1}{k})} \quad (7)$$

where $\bar{v}$ is the average wind speed, σ is the standard deviation and Γ is the gamma function.

C. Method of Moment 3

By Moment 3 method of Weibull parameter estimation, first the k parameter is being determined from [17] through equation 4.12.

$$\frac{\bar{v}^3}{(\bar{v})^3} = \frac{\Gamma(1+\frac{3}{k})}{\Gamma(1+\frac{1}{k})^3} \quad (8)$$

Afterwards the A parameter is determined from

$$A = \frac{\bar{v}}{\Gamma\,(1+\frac{1}{k})} \quad (9)$$

where $\bar{v}$ is the mean wind speed and Γ is the gamma function

D. Power Density Method

The power density method (PDM) is a numerical analysis related to the mean wind speed and for this method initially the Energy Pattern Factor (E_{pf}) should be computed. E_{pf} can be defined as the ratio of mean of cubic wind speed to the cube of mean wind speed. The energy pattern factor (EPF) denoted in [16] as:

$$E_{pf} = \frac{\frac{1}{N}\sum_{i=1}^{N} {v_i}^3}{(\frac{1}{N}\sum_{i=1}^{N} v_i)^3} \quad (10)$$

Subsequently, the shape parameter (k) and scale parameter (A) can be calculated from the previous E_{pf} equation through below format described in [16] as:

$$k = 1 + \frac{3.69}{{E_{pf}}^2} \quad (11)$$

$$A = \frac{\bar{v}}{\Gamma\,(1+\frac{1}{k})} \quad (12)$$

where $\bar{v}$ is the average wind speed, v_i is the observed hourly wind speed, E_{pf} is the energy pattern factor. With the aim of evaluating the accuracy of the previously mentioned 4 Weibull methods, 3 types of error estimation has been calculated. Those are:

E. Root Mean Square Error (RMSE)

RMSE estimates the error by considering the deviation between the predicted and the experimental values. The lower the value of RMSE is, the more accurate the values are. RMSE should be as close by to zero as possible to have a successful forecasting of prediction. The equation formula of RMSE is described in [16] as:

$$\text{RMSE} = (\frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2)^{1/2} \quad (13)$$

where, y_i is the frequency of the observation , x_i frequency of Weibull (using those statistical methods) and N is the number of observations.

F. Chi-Square Test(χ^2)

This test is considered accurate when the value comes closer to zero. Basically it gives back the mean square of the deviations between the experimental and the calculated values for the distributions. Chi-squared test is calculated in [16] as following:

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - x_i)^2}{x_i} \quad (14)$$

where, y_i is the frequency of the observation x_i frequency of Weibull (using those statistical methods) and N is the number of observations.

G. *Coefficient of Determination (R^2)*

The coefficient of determination (R^2) represents a linear relationship between the calculated values and observed data. The higher value of it represents a better fit and the highest value it can get is 1. The function can be represented in [17] as:

$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} \quad (15)$$

$$where,\ SS_{RES} = \sum_{i=1}^{N}(y_i - f_i)^2 \quad (16)$$

$$SS_{TOT} = \sum_{i=1}^{N}(y_i - \bar{y})^2 \quad (17)$$

Here, y_i is the frequency of the observation, $\bar{y}$ is the mean of the observed data, N is the number of observation, f_i is the predicted value.

V. RESULTS

The purpose of this section is to combine all the satellite datasets to generate a Wind Atlas of the Bay of Bengal and have a comparative study among GWA and satellite datasets.

H. *Wind Atlas Generation*

After S-WAsP mapping of the achieved wind scenes, the time-series wind measurement dataset is being generated as .tab files which is a file format for observed wind climate. From then the forming of a generalized wind climate in .lib file begins which is format of the Atlas files. For this instance, a vector map of the whole zonal area is used for specific terrain analysis and a met-station for the site precise location is settled up in WAsP.

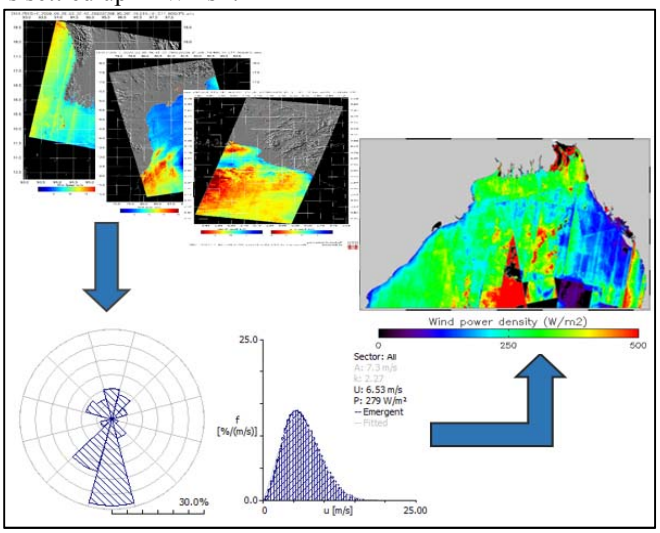


Figure 2: Wind resource mapping process from ASAR wind scenes

If the met station location is preferable and specific under that zonal area, the Generalized Wind Climate can be further calculated. For each .tab file, the same procedure is followed to create the respective .lib files. The generated .lib file contains the sectoral wise frequency of occurrence of the wind (the wind rose) as well as the wind speed frequency distributions in the same sectors (as Weibull A- and k-parameters). The wind climates are specified for a number of reference roughness classes and heights above ground level.

For specific roughness and height selection, it provides a combined 12 sectoral wind rose, sectoral frequencies of the wind direction, mean wind speed, mean power density along with Weibull shape (k) and scale (A) parameters

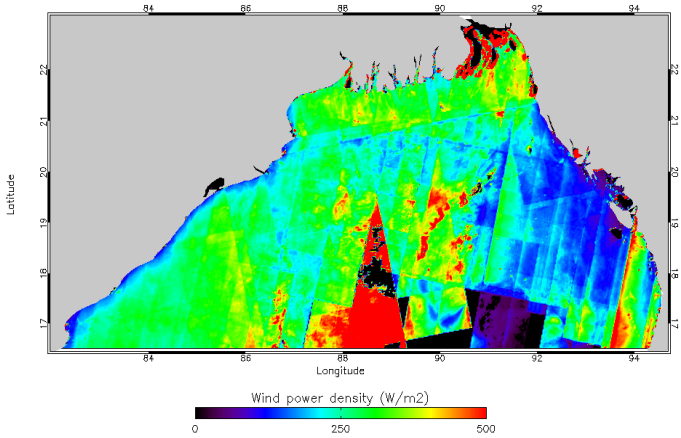


Figure 3: Wind Power density map of the Bay of Bengal through ENVISAT ASAR wind dataset

B. *Comparison and Outcomes*

The nodal comparison was done in an area basis (60 km by 60 km) keeping the GWA node points in the middle. So these area wise nodal comparisons were performed where 122 coordinated nodal location points are found in the defined zone. While comparing the Satellite based Atlas with GWA, we have found in further from shore areas, the number of samples are really scanty which leads to a lower data obtain scenario of that particular area. The higher the number of samples, the better the estimation can be provided. In addition, from previous studies it has been perceived that, wind resources can be predicted well enough through 70 samples considering an uncertainty of ±10% at a confidence level of 90% [13]. To gain more accuracy, those ENVISAT ASAR nodal locations were chosen which has more than 70 time series samples. This lead to 26 nodal sample comparison in between GWA and ENVISAT ASAR wind atlas.

In the figure 4, 5, 6 and 7, the scenes which has more than 70 time series samples has been selected. For this reason other nodal points of deep sea is showing as null values. From the figure, two particular interest areas can be found where the number of samples are moderate, which is in the Northern coastal area near Bangladesh and in the Western side of the Bay of Bengal near India. While comparing the mean wind speed verified ENVISAT ASAR wind Atlas map with GWA, ENVISAT ASAR wind Atlas map shows a moderate result, though the ASAR seems to be over predicting the wind speed and energy entirely, especially in the Northern coastal areas.

In the figure 4, it shows the calculated offshore mean wind speed from ENVISAT ASAR datasets. The color legend shows the variation of mean wind speed availability of that particular zone of the Bay of Bengal. From the legend it is clear that, the color orange to red is the high potential zone whereas the color below are indicating a lower wind speed scenario. Figure 4 represents a variation of around 6-7.5 ms^{-1} mean speed. In the Northern side of the Bay of Bengal, the mean wind speed is around 6.5-7.5 ms^{-1} where as in Eastern part it varies around 6-7.29 ms^{-1} On the other hand, figure 5 represents the GWA mean wind speed for the same location points where GWA shows a variation of wind speed from

5.05 to 7.82 ms- Alike SAR, It can be seen that in Western side of the Bay of Bengal, the mean wind speed is found to be around 6.5-7.2 ms^{-1}.

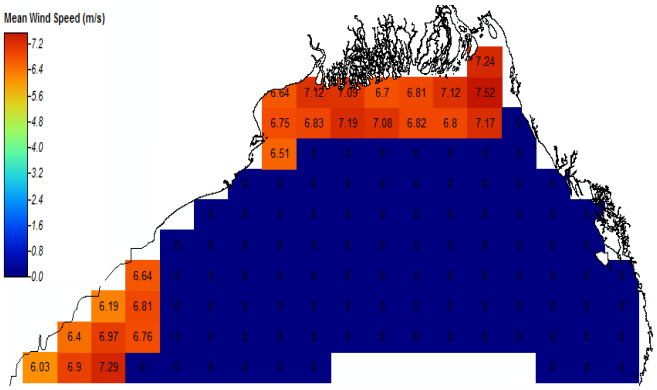


Figure 4 : Mean wind speed of the specified location from ENVISAT ASAR datasets

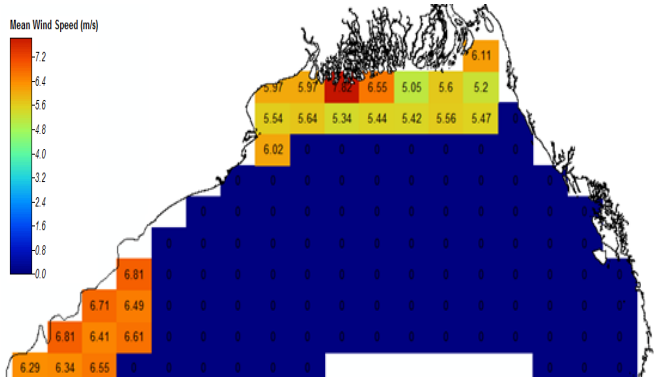


Figure 5 : Mean wind speed of the specified location from GWA datasets

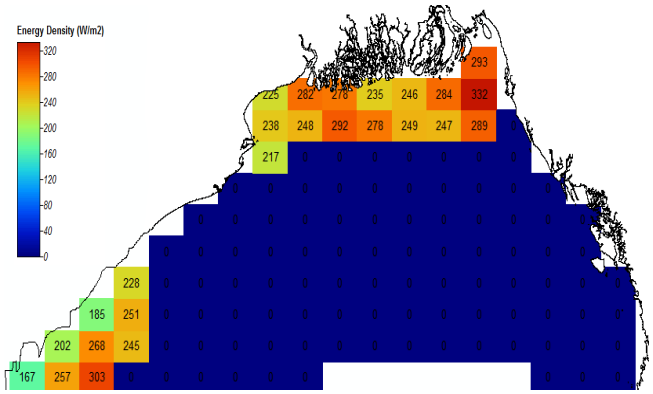


Figure 6. Energy density of the specified location from ENVISAT ASAR datasets.

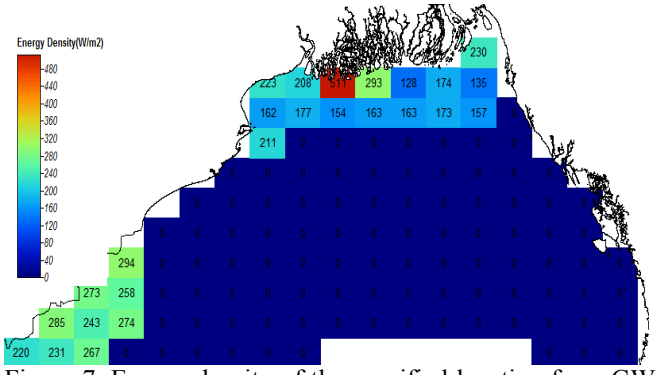


Figure 7. Energy density of the specified location from GWA datasets

Figure 6 represents the energy density from ENVISAT ASAR datasets for the specific location where it shows a good variation of 225-332 W/m2 in near coastal area of Northern side of the Bay of Bengal where as in Western coast it is a bit less comparatively (167-303 W/m2). On the other hand, from the wind energy density verified GWA map, the energy density is found to be 162-293 W/m^2 in the Northern nearshore area (Figure 7). Where as in a specified location, it is estimated to have around 511 W/m^2 of energy density. Another corner of the map which is situated in the Western coast, the variation of 231-294 W/m^2 energy density has been found.

VI. Test Case

A case study has been performed through assembling all the obtained datasets (ASAR, GWA and ASCAT). Through an engineering application this test case has been approached to estimate an offshore wind park production for an appropriate zone at the Bay of Bengal.

A. Selection of the suitable location

The main focus is to create a 400 MW wind farm at low depth sea locations but high in power density; which is 4-6 km close to the shore, so that it does not require any supplementary base station and additionally, it will reduce the cost of the offshore wind farm installation. The proceeding steps of this engineering application has been displayed in a flowchart below.

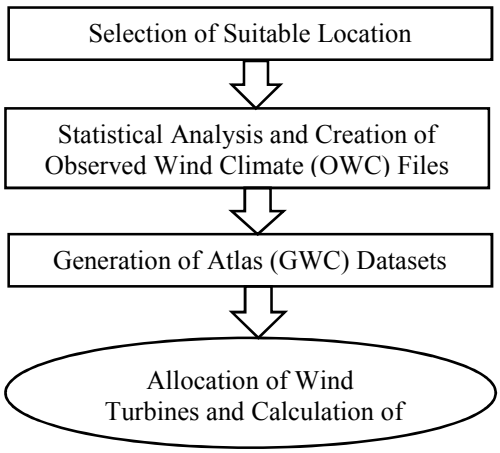


Figure 8. Flowchart of Case Study Phases

Our achieved SAR and GWA power density maps were analyzed to find a comparable higher energy density area comprising higher sample number of SAR dataset. When a point location has more number of SAR samples that represents more data acquisition in the time domain which can lead to a reliable estimation of wind energy density. Also, another preferences were to find a nearshore, lowest depth and highest potential area for case study purpose. To apply the newest technology of offshore wind turbine in the specified area, it leads to operational uncertainty and more expenses. That is why the conventional turbine installation techniques has been preferred in this section.

Therefore, based on four detailed criterion, the zone has been selected as (1) Sufficient number of satellite data sample to get more accurate results. (2) Highest wind energy density region as an aspect to highest energy production. (3) Nearest to shore to not to have a base station in offshore. (4) Lowest

depth region to deal with the present technology of the offshore with turbine.

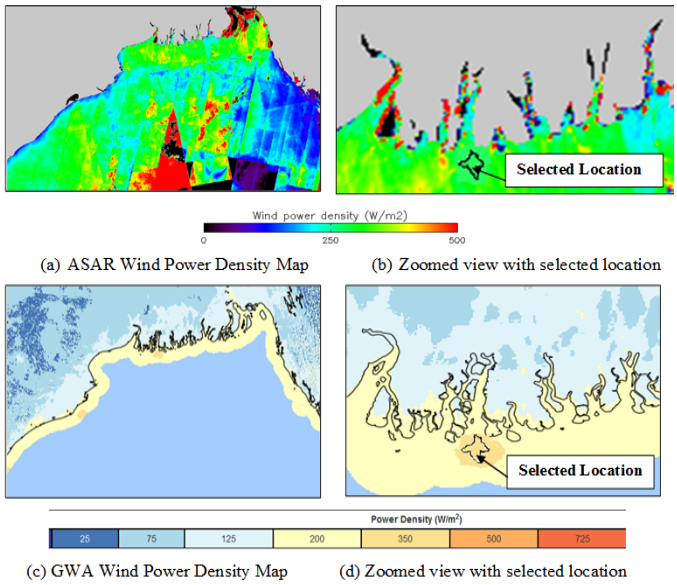


Fig 9. Selection of the suitable location

To determine a potential region for wind farm setting, selection of the highest wind energy density location is a must. For this purpose, the attained ASAR based Atlas and GWA power density maps were well analyzed to find out a specific zone with equivalent high power density (fig 9). In order to perform detailed analysis, the sample number of ASAR datasets were prioritiezed as a purpose of achieving more accuracy in wind energy estimation.

For a purpose of depth analysis of the Bay of Bengal, the bathymetry map has been downloaded in Shuttle Radar Topography Mission (SRTM) 30+ version of 30 arc second, Global. SRTM 3-arc second resolution digital elevation maps (DEM) have been merged with the vector map of the Bay of Bengal and divided in 6 zonal divisions. The target was to focus on the lowest depth zone and near to shore location.

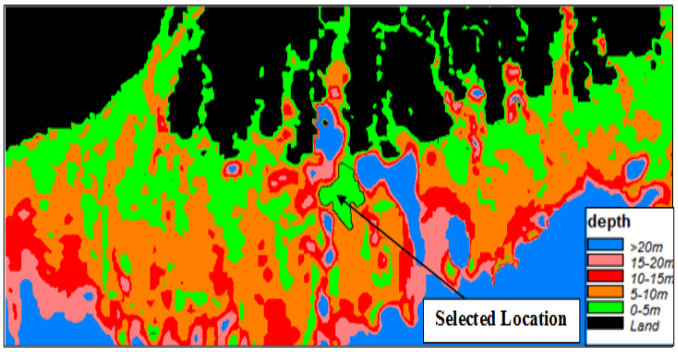


Figure 10. Bathymetrical segmentation of the Bay of Bengal

After investigating the figures of 10 and 11, such close to shore, higher sample number and lower depth areas with higher wind power density was found to be existed only at the northwest coast of Bangladesh. Therefore, the area has been chosen at the location of 21.5° Northern latitude and 88.667° Eastern longitude. This area is in the depth region of 0-5m, 5 km away from the coastline, comprising one of the highest sample numbers of 122 of ASAR datasets and predicted to have higher wind power density in an assumption from GWA and ASAR based wind datasets. In addition, ASCAT dataset was obtained for the chosen area which comprises a full year of 2015 measurements. The dataset contains 567 time series samples that fairly distributed in every single month and ascending-descending arrangement. This obtained ASCAT dataset is more frequent than ASAR datasets due to its technological aspects. Both of the satellite datasets were analyzed accordingly to get a broader interpretation.

B. Implying the statistical analysis

In our study, we investigated Weibull fitting by 4 methods as Empirical method (EM), Moment Method (MOM), Moment 3 method (MOM 3) and power density method (PDM) in the selected zonal area for wind farm. In addition, the relative accuracy test of Root Mean Square Error (RMSE), Chi-Square (χ^2) and Coefficient of determination (R^2) has also performed to predict the energy density uncertainty. Here, it should be mentioned that, RMSE and (χ^2) is considered more accurate when it is near to zero and (R^2) is considered more accurate when it is near to its highest value of 1.

TABLE 1. TWELVE SECTORAL STATISTICAL WEIBULL ANALYSIS OF ASAR DATASETS

Method	A	k	v_m	RMSE	χ^2	R^2
EM	7.39	2.09	6.54	0.00654101	0.104194	0.88102
MOM	7.39	2.07		0.00653913	0.103496	0.881089
MOM3	7.392	2.12		0.00655685	0.106151	0.880443
Epf	7.392	2.13		0.00656684	0.1071	0.880079

TABLE 2.TWELVE SECTORAL STATISTICAL WEIBULL ANALYSIS OF ASCAT DATASET

METHOD	A	K	v_m	RMSE	χ^2	R^2
EM	6.24514	2.21076	5.53095	0.003134	0.059732	0.914321
MOM	6.2453	2.19501		0.003104	0.058549	0.915931
MOM3	6.24533	2.18982		0.003095	0.058197	0.916435
EPF	6.24526	2.19938		0.003112	0.05886	0.915497

From the results, the MOM 3 method was chosen to be the best fit for the datasets and further WAsP analysis has been performed to predict the annual energy production of that particular area.

C.Reformation of Sectoral division

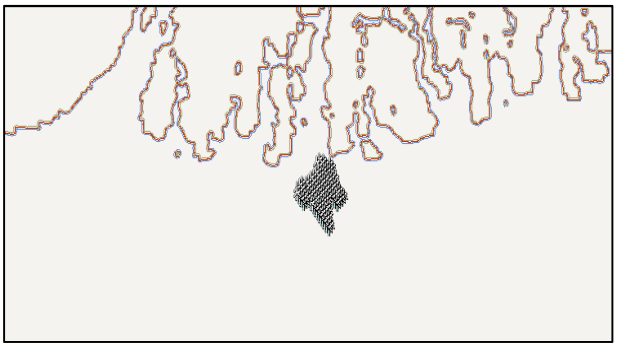
As the observed dataset of our study did not have efficient sample numbers for each sector, the samples were too few to count, that is why another proposed idea has been represented here. The proposed idea is to obtain more wind speed datasets in one sector, by reforming the sectoral division to 60 degree from 30 degree. So, in spite of having 12 sectors now, 6 sectoral representation has been investigated. As a result, the number of

samples for wind speed calculation increased in every sector and it shows a better result of Weibull fitting.

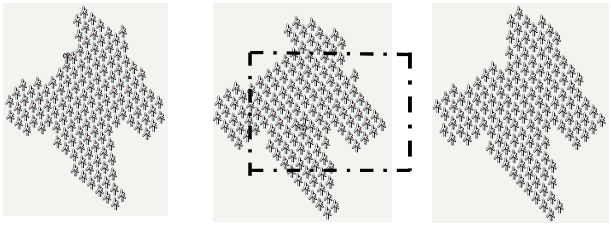
D.Wind Farm Layout

The selective zone was calculated and found to have 174.5 km^2. Based on this area, a minimum production from that extent was designed via allocating 134 turbines. These 134 turbines were chosen to have 3MW capacity individually. This results as estimation of 402 MW capacity offshore farm with 2.3 MW/ km^2 energy distribution per area. In our study, Vestas V112-3.0 MW model of turbine has been chosen which has a rotor diameter of 112 m and hub height of 84m. As at present, mostly used turbines are of around 3 MW for offshore wind technology, our turbine specification has been chosen accordingly. Turbines were set in an equidistant of 10 diameter from each other as a way to let the wind recover from turbine to turbine.

Turbines front sides were also considered depending on the dominant wind directions. It was found to be 229.459°, 218.015 for ASCAT and ASAR datasets respectively. The dominant direction for every sector has been determined by combining the wind speed and frequency generation of that respective sector and the opposite sector. In the below figure 11 the different wind farm layout based on the leading wind direction is presented where it shows a slight rotation of wind turbines formats.



(a) Selected area of turbine allocation



SAR (218.015 °) ASCAT (229.459°) GWA (232.2 °)

Figure 11. Wind farm layout for every datasets.

E.Calculation of Annual Wind Energy Production

After the wind farm insertion to the selective WAsP file, the calculation of Annual Wind Energy has been performed. The gross AEP is the energy production of a single turbine through corresponding wind distribution at the predicted hub height in the form of the supplied power curves. In gross AEP, the wake effect is not accounted. And for Net AEP case, the wake effect is counted as it the whole total annual energy of the farm. By the intricacy of this power curve and numerical wind distribution, each turbines annual energy production is calculated. Afterwards, by combing the each turbines annular energy and wake loses, the AEP is calculated for the whole wind farm.

TABLE 3. CALCULATION OF AEP USING MOMENT 3 METHOD

Source	Gross AEP (GWh)	Net AEP (GWh)	Wake Loss (%)
ASAR 12	1792	1693	5.52
ASAR 6	1793	1682	6.22
ASCAT 12	1385	1310	5.45
ASCAT 6	1379	1269	7.99
GWA	1352	1276	5.61

The above table (table 3) shows AEP estimation through MOM 3 method for all three datasets in 5 forms which are- ENVISAT ASAR for 12 sector and 6 sector, ASCAT for 12 sector and 6 sector and GWA. The results are quite similar among ASCAT and GWA, whereas ENVISAT ASAR datasets has an overestimation.

TABLE 4. ERROR CALCULATION OF SATELLITE DATASETS WITH RESPECT TO GWA.

Source	**Sector**	**Error Rate (%)**
ENVISAT ASAR	12	28
ENVISAT ASAR	6	27.6
ASCAT	12	2.6
ASCAT	6	0.5

While comparing the 12 sectors, it estimates a higher NET AEP, which leads to a higher error rate compared with GWA (table 4). On the other hand, when the sectors are enlarged (from 12 sector to 6 sectors), the error rate is less specially in ASCAT scenario. But the wake losses are found to be in rise when it is 6 sector

VIII. CONCLUSION

For a large offshore area, it is very difficult to define where should be focused on for planning of an offshore windfarm. Even we know, satellite measurements are new and has limited number of samples in any location, they are still the only representative of onsite measurements more than any of the models. Several models of satellites along with its higher number of samples can be used without any local measurements, although this can lead to very big uncertainties. But it can estimate a general region for further broad investigation.

Consequently, in this paper three (ENVISAT ASAR, ASCAT and GWA) remote sensing measurements have been discussed and those datasets were used primarily to generate a wind Atlas of the Bay of Bengal and by utilizing that Atlas a potential zone has been found at a near shore area of the Bay of Bengal. Formerly, a case study has been performed to predict the Annual Energy Production of that particular area. Among these datasets, two (ENVISAT ASAR and ASCAT)

are from satellite based and one (GWA) from long term mesoscale model measurement. The satellite datasets were compared with GWA, as no offshore wind farm has been established in the Bay of Bengal yet and there is no onsite data to have a suitable comparative study.

In this paper, the data analysis on the satellite datasets has been performed along with different statistical methodology selected empirical state of the art moment fitting function to have a better estimation. The initial results with the current dataset shows similar shape and scale parameters for climate values. Therefore it can be understood that, with the limited number of sample datasets, even the simple empirical method are accepted to use. Furthermore, in this study, a potential location has been chosen to define if these data can be used for modeling. As a result all the datasets have given similar production values without directional information. When the directional information is added, less number of data samples were seen to create less quality as it is expected.

The resulted calculation with their error is pointing towards the fact of directional view of wind (sectors) and lesser number of samples. Differences were accounted because of limited data in some sectors. So, reducing number of the sectors were suggested which showed improvement in Weibull fitting, though it has a slight increase in wake effects. Analysis shows that while there is not enough samples, all the methods shows similar results. One can understand from this result that, satellite data can be used to identify the highest density areas for a particular zone but it cannot be used exactly to estimate the highest region. As in our study, the satellite datasets were pointing the same location with GWA and one can make a preliminary analysis of the wind conditions by knowing that it can go up to 20% uncertainty.

From this paper, it can be said that, satellite gives us the ability to identify the higher power density area so that we can utilize it for met mast set up. Even to find a potential zone for establishing mat mast is challenging and time worthy. The foundation of this thesis project is that, satellite datasets can be used to identify potential zone and it has a promising outcome of calculation capability of annual energy production, if compared with any long time statistical result and bathymetry data combination. Furthermore, all three datasets used in this study comprises similar AEP (ASCAT better than ASAR) at the coastal area which indicates the selective location as a potential area and further researches regarding this study can lead towards a better future for the wind energy sector of Bangladesh.

REFERENCES

[1] Beaucage, P., Bernier, M., Lafrance, G., & Choisnard, J. (2008). Regional mapping of the offshore wind resource: Towards a significant contribution from space-borne synthetic aperture radars. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 1(1), 48-56.

[2] Kader, A., Chowdhury, M. Z. R., Uddin, M. M., Hoque, M. E., & Basak, S. C. (2013). Bathymetric study of the Bay of Bengal based on open source satellite and sounding data. International Journal of Geomatics and Geosciences, 4(1), 116.

[3] Christiansen, M. B. (2006). Wind energy applications of synthetic aperture radar. Institute of Geography, Faculty of Science, University of Copenhagen.

[4] Dagestad, K.-F., Horstmann, J., Mouche, A., Perrie, W., Shen, H., Zhang, B., . . . Lehner, S. (2012). Wind retrieval from synthetic aperture radar-an overview. Paper presented at the 4th SAR Oceanography Workshop (SEASAR 2012).

[5] Hasager, C. B., Mouche, A., Badger, M., Bingöl, F., Karagali, I., Driesenaar, T., . . . Longépé, N. (2015). Offshore wind climatology based on synergetic use of Envisat ASAR, ASCAT and QuikSCAT. Remote Sensing of Environment, 156, 247-263.

[6] Hasager, C. B., Astrup, P., Christiansen, M. B., Nielsen, M., & Barthelmie, R. (2006). Wind resources and wind farm wake effects offshore observed from satellite. Paper presented at the Proceedings of the European Wind Energy Conference and Exhibition (EWEC).

[7] Ricciardulli, L., Wentz, F.J. (Producer). (April 2016, 20 June,2016). Remote Sensing Systems ASCAT C-2015 Daily Ocean Vector Winds on 0.25 deg grid, Version 02.1. Retrieved from www.remss.com/missions/ascat

[8] Hasager, C. B., Bingöl, F., Badger, M., Karagali, I., & Sreevalsan, E. (2011). Offshore Wind Potential in South India from Synthetic Aperture Radar (8755039081). Retrieved from http://orbit.dtu.dk/fedora/objects/orbit:87302/datastreams/file_7749647/content

[9] Hasager, C. B., Mouche, A., Badger, M., Bingöl, F., Karagali, I., Driesenaar, T. Longépé, N. (2015). Offshore wind climatology based on synergetic use of Envisat ASAR, ASCAT and QuikSCAT. Remote Sensing of Environment, 156, 247-263.

[10] Peña, A., Hahmann, A., Hasager, C. B., Bingöl, F., Karagali, I., Badger, J., . . . Clausen, N.-E. (2011). South Baltic Wind Atlas: South Baltic Offshore Wind Energy Regions Project: Risø DTU-National Laboratory for Sustainable Energy.

[11] Bingöl, F., Hassager, C. B., Badger, M., & Badger, J. (2013). RZGM2013-14.

[12] Global Wind Atlas. (2015). Retrieved from http://globalwindatlas.com/tutorials.html

[13] Hasager, C. B., Astrup, P., Barthelmie, R., Dellwik, E., Mortensen, N. G., Nielsen, M. . . . Rathmann, O. (2002).

[14] ASCAT, Metop Meteorological Missions. (2006). Retrieved 28 June 2016, from European SpaceAgency(ESA), Retrieve from http://www.esa.int/Our_Activities/Observing_the_Earth/The_Living_Planet_Programme/Meteorological_missions/MetOp

[15]SAF, O. (2013). ASCAT Wind Product User Manual. Version, 1, 23.

[16] Kidmo, D., Danwe, R., Doka, S., & Djongyang, N. (2015). Statistical analysis of wind speed distribution based on six Weibull Methods for wind power evaluation in Garoua, Cameroon. Revue des Energies Renouvelables, 18(1), 105-125

.[17] Amaya-Martínez, P.-A., Saavedra-Montes, A.-J., & Arango-Zuluaga, E.-I. (2014). A statistical analysis of wind speed distribution models in the aburrá valley, Colombia. CT&F-Ciencia, Tecnología y Futuro, 5(5), 121-136

Electrical Energy Harvesting from the Foot Stress On Foot Overbridge Using Piezoelectric Tile

Tarik Bin Abdul Akib [1], Hasanuzzaman Mehedi [2], Md. Nazmuschayadat [3]
Dept. of Electrical and Electronic Engineering [1, 2, 3]
Bangladesh Army University of Engineering and Technology [1, 2, 3]
Rajshahi, Bangladesh [1, 2, 3]
abdulakibruhe@gmail.com [1], hzmehedi123@gmail.com [2], nazmuschayadat.eee@gmail.com [3]

***Abstract*—In this paper, Power generation from the foot overbridge using piezoelectric tile, which is constructed with the piezoelectric transducer, is being discussed. Often more, it is seen, people walk on foot overbridge. Installing piezoelectric tile (Piezotile), we can achieve the huge amount of power, a single piezotile, which has the capability to produce up to 7-watt power, eventually installing numerous piezotile, will provide us with much power, and so power generation from this way is an easy way out. To analyze the mechanical system of the piezotile convert to the electrical circuit and PID controller is used to making the system more stable and increase the accuracy in the piezoelectric transducer. It has the possibility to generate a huge amount of power and this power source has various applications in the home, street lighting, buildings, store energy, emergency power grid and so on. Energy is converted from one form of energy into another and most of the energy is lost in nature. The concept of green energy is to transform various energy into electrical energy. Piezotile is one of them, which convert foot stress into electric energy.**

Keywords*—Foot Overbridge; Footstep; Piezoelectric tile; Piezoelectric; PID; Power Generations; Renewable Energy.

I. Introduction

Usage of the unusable energy of foot stress with human motivity is immensely relevant and significant for dense countries. When the footbridge plane is mechanized with piezoelectric technology, the electrical energy produced by the foot stress is captured by piezotile and then reserved and used as an energy source. The piezoelectric effect has the sufficiency to generate electrical energy in repercussion for the applied foot stress [1]. Piezoelectric transducer has two properties that are defined as a direct and contrary effect. The direct effect is the property of some materials to promote electric charge on their plane when mechanized strength is put on them, while a contrary effect is the property of some materials to produce mechanized strength when an electric charge is induced. Nowadays, the fabrication of piezoelectric nanogenerators (NGs) was demonstrated using piezo materials such as ZnO, $BaTiO_3$, $ZnSnO_3$, $NaNbO_3$, (K, Na)NbO_3, PMN-PT, Ba(Zr, Ti)O_3–(Ba, Ca)TiO_3 and Pb(Zr, Ti)O_3 and so on. Normally, when mechanical stress is not initiated, piezoelectric crystals are neutral meaning their electrical charges are perfectly balanced. If a piezoelectric crystal is aggregated, the natural balance of positive and negative atoms is upset, causing net electrical charges to appear on the opposite, outer faces of the piezoelectric material [2]. The output voltage of a piezoelectric element is given by V = Q/C, where C is the capacitance, Q = d_{33}F is the charge induced by a force, F applied in an approach upright to the piezo electrode surfaces and d_{33} is the relevant piezoelectric coefficient [3]. Pavegen Systems, a technology company based out of London, recently introduced a piezoelectric based tile that harvests the kinetic energy created by a footstep. Their 17.1 by 23.6-inch tiles can produce an average of 7 watts per step. In the Pavegen Systems study, the piezoelectric tiles produce an average of 7 watts per footfall from a 150 pound, walking person [1] [4]. In Fig. 1 Piezoelectric transducer based single and assembled numerous piezotile are illustrated. Piezoelectric transducer based tiles is another prospect of renewable energy.

II. Proposed Structure of Piezoelectric Tile

The proposed design structure of piezotile is shown in Fig. 1. This theoretical piezotile is assembled with the numerous piezoelectric transducer. This total system is fabricated on the foot overbridge. When the tiles mechanized by the foot stress the tiles transfer their energy. Thus electrical power is produced and collected from the piezotile.

(a) (b)

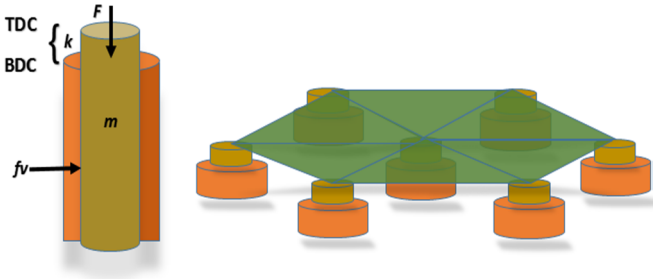


Fig. 1. The design structure of (a) single piezoelectric transducer; (b) and assembled numerous Piezotiles.

The mechanical model of piezotile, the foot stress (*F*), piezoelectric transducer mass (*M*), viscous damper (*fv*), and spring constant (*k*) interchange into the series connected dc voltage source (*v*), inductor (*l*), resistor (*r*) and capacitor (*c*) of the electrical equivalent circuit respectively.

(a) (b)

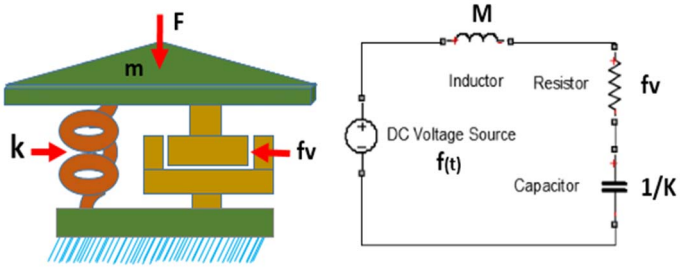


Fig. 2. The mechanical & Electrical model of the Piezotile.

Fig. 2 shows the mechanical model of piezotile transform into the electrical circuit and this is the best technique to analyze the piezotile. The transfer function (*tf*) of the electrical equivalent circuit can be written as:

$$tf = \frac{V(S)}{F(S)} = \frac{S}{MS^2 + f_v S + K} \quad (1)$$

A. Mathematical Representation of Piezotile

In arithmetically for input voltage $f(t)$ and current $i(t)$, the transient power of a piezoelectric transducer can be expressed as,

$$P_{transducer}(t) = f(t) * i(t) = \frac{1}{2} VI[\cos\phi + \cos(2\omega t - \phi)] \quad (2)$$

The effective power, which is the root mean square of the instantaneous power is as following [5],

$$P_{eff} = \frac{V^2}{2|Z|}\sqrt{\cos^2\phi + \frac{1}{2}} \quad (3)$$

Here $V = F_{step}/n$, where F_{step} is the amplitude of the foot stress force on the tile and n is the total number of footsteps on the tile respectively. Moreover, Z is the impedance. The effective input power can therefore be written as [5]:

$$P_{footstep} = \frac{F_{Step}^2}{2n^2|Z|}\sqrt{\cos^2\phi + \frac{1}{2}} \quad (4)$$

The piezoelectric transducer vibrates at the maximum amplitude at the resonance condition, when the inductive reactance, X_L equal to the capacitive reactance, X_C. Then the circuit impedance will be the smallest value and the circuit current will be maximum. When the load is purely resistive, no phase difference among voltage and current, then the output electrical power can be expressed as [5]:

$$P_{electrical} = \sqrt{\frac{3}{8}}\frac{V_{out}^2}{R} = \sqrt{\frac{3}{8}}\frac{SF_{Step}^2}{n^2 R(MS^2 + f_V S + K)^2} \quad (5)$$

From equation (4) and (5), the transient power and the efficiency of a piezotile can be observed easily. Eventually, the whole efficiency of a piezotile can be intended from Equation 6.

$$\eta_{tile} = \frac{p_{electrical}}{p_{footstep}}\% \quad (6)$$

B. Effects of Foot Stress on Proposed Piezotile

The force of gravity is mathematically represented as, $Fa = mg$. Where m is mass on the earth and g is the acceleration of gravity. The acceleration of gravity is approximately 9.8 m/s^2. Installing piezotile near or a little distance from the earth's surface, it does not put any impact on the acceleration of gravity, it's negligible. When we use piezotile on foot overbridge, the distance (4-5m) puts zero impact on the acceleration of gravity.

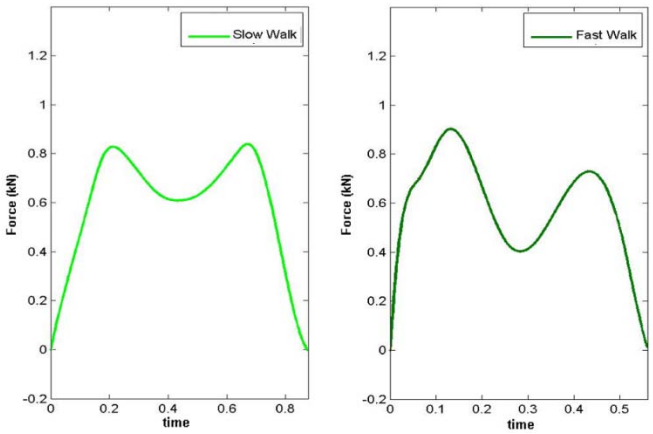


Fig. 3. Foot Stress on the piezotile at slow walking and fast walking time [6].

Fig. 3 shows the resulting force delivered to the piezotile by footstep. The waveforms observed when walking at a slow pace and a fast pace. The peak foot force is seen to approach 0.8kN at a slow pace. The peak force is seen to approach 0.9kN. As mentioned earlier, The 7 watts yielded per step stated by the Pavegen experiment was the output of the step pressure exerted by a 150-pound man (68.04kg) [1].With a gravitational acceleration of 9.8(m/s^2) that mass of the piezoelectric subject is,

$$M_p = 68.04 * 9.8 = 666.78\ N.$$

A study on the average walking speeds of subject's variable, found that the average speed of 4.52 feet per second and an average startup time of 3.39s to reach that speed, then the acceleration rate (a) is 1.33ft/s^2. The calculation can be expressed as [1]:

$$a = \frac{v_f - v_i}{t} = \frac{4.52\ ft/s - 0\ ft/s}{3.39s} = 1.33\frac{ft}{s^2} \quad (7)$$

Given the calculated mass and acceleration of walking on the piezotile subject, a force per step can be calculated as;

$$Force = M_p * a = 666.78\ N * 1.33\ ft/s^2 = 889.68\ N \quad (8)$$

So the assumption can be made that a walking person exerts 889.68N of force per step, which is similar to the force of a walking step shown in the time wave charts of Fig. 3 at around 0.8 to 0.9kN. So it is safe to say that 889.68 Newton of force generates 7 watts of energy or 127.10 Newton per watt [1].

III. Analysis and Assessment of Simulation

To analyze the mechanical system of piezotile in Fig. 2 converts to the electrical equivalent circuit [7]. This is the fair technique to analyze the piezotile. Equation (1), shows the transfer function of a single piezotile which is utilized to analyze the piezotile. The output voltage is observed by the response of the input signal by simulating the time response of the piezotile circuit model. For instance, we have totaled the output voltage for resistance value, $fv = 100m\Omega$, inductor value, $M = 1H$ and capacitor value, $K = 60mF$.

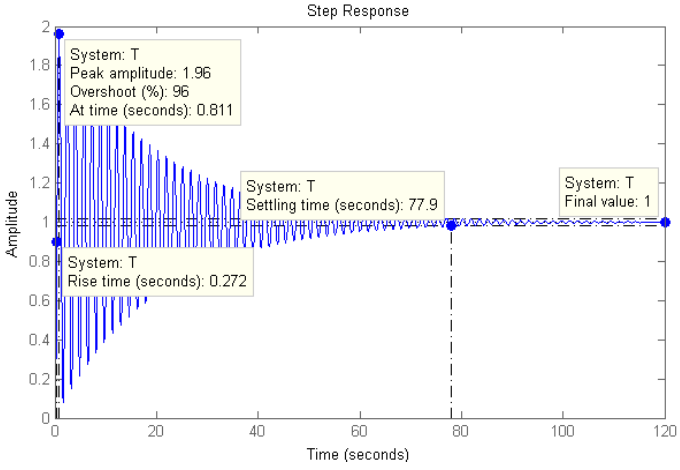


Fig. 4. The step response of the Piezotile.

Fig. 4 shows the plot of the step response of an arbitrary dynamic system model of the piezotile. The peak amplitude (A=1.96v), peak time (t_p=0.811s), rising time (t_r=0.272s), normal frequency (ω_n=1), Restraining ratio (ζ=0.0033), Relaxing time (t_s=77.9s), and percentage of overshoot (%os=96) is observed and simplified in Fig. 4.

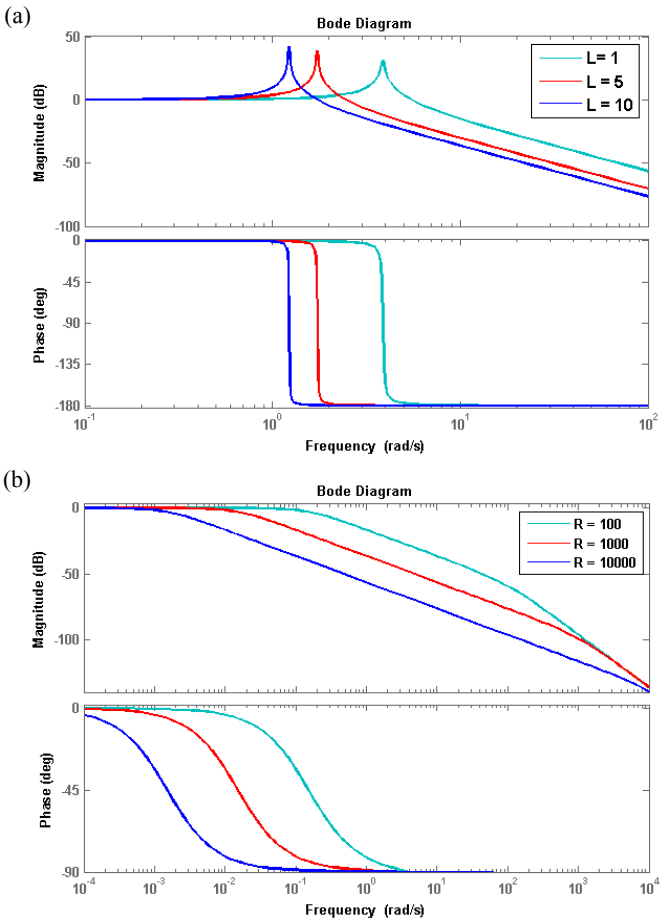


Fig. 5. The frequency response bode diagram of the electrical circuit model of piezotile with (a) several **L** values and (b) several **R** values.

In Fig. 5 shows the frequency response bode diagram of the electrical system, influenced by the different resistance (*R*) and the inductor (*L*) values. While the value of L is increased then the magnitude of the harvested energy is increased and while the value of R of the system is increased then the magnitude of the harvested energy is losses. R and L signifies the viscous damper of piezotile and mass via foot forced on the piezotile respectively. If the oscillatory system or damping is an underdamped condition ($\zeta<1$) and foot stress on the tile is increases then we get maximum output energy.

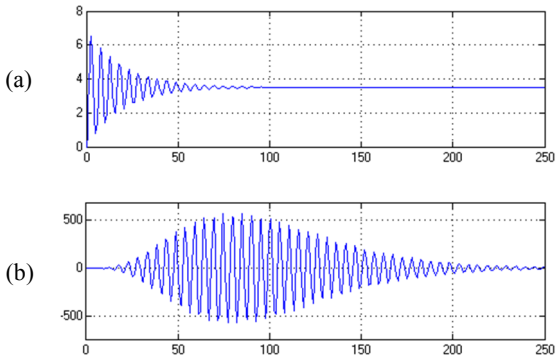


Fig. 6. The step response of (a) parallel connected and (b) series connected numerous piezotile.

Numerous piezoelectric transducer is generally connected in series or parallel connection. In a series connection, piezoelectric layers have opposite polarization directions, which produce a larger voltage output, but lower current output and parallel connection piezoelectric layers have the same polarization directions which produce in a larger current output, but a lower voltage output [8]. In figure 6, shows the step response of series and parallel connected numerous piezotile, between them parallel connected tile has more stable.

A. PID Controller for The Expeditious Stability of Piezotile.

PID control is a usual and easy control approach, which is used to make the system stable, overcome creep and vibrations, improvement of the steady state error, provide high gain feedback and enhance the tracking accuracy in the piezoelectric transducer based piezotile.

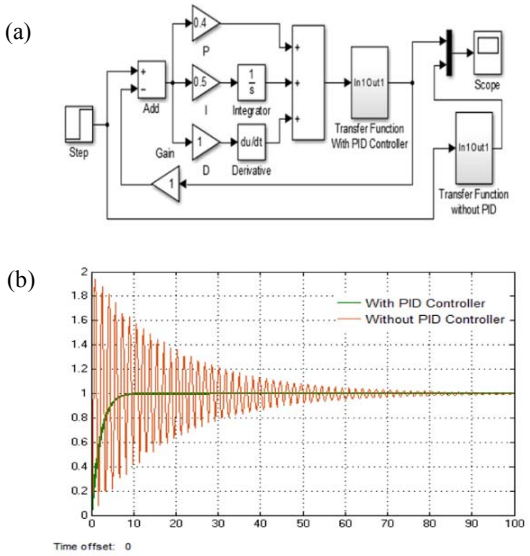


Fig. 7. The block of dynamic simulation model: (a) The entire model simulation in Simulink (b) Magnitude in time response with and without PID controller.

In PID controller, tune up the P, D and I gain for the steady oscillation and fixed the point for desired oscillations respectively. After using the PID controller, the system response is more stable than the underdamped conditions. As a result, the output power gain, the fastest responsibility, accuracy and reliability is increased. The results in Fig. 7 provide insight into the PID stabilization and offer convenience for the tuning of PID controllers.

B. Installation cost

Types of Piezoelectric technology were classified according to the uses, features and technical specification. Many types of piezoelectric technology used for harvesting energy.

TABLE I. PIEZOELECTRIC TECHNOLOGY TYPES AND PRICE [9].

Company Technology	Generated Energy	Price (US-Dollar)
Sustainable Energy Floor	30 W	$851
Waynergy Floor	10 W	$277
Pavegen Tile	7 W	$76
Electro-Active Polymers	1 W	Unknown
$BaTiO_3$	$7mW/cm^3$	Low Cost

POWER Leap PZT	0.5 mW	Unknown
Drum Harvesters Piezoelectric Ceramics	2.463 mW	$28

For the different feature, cost and technological reason pavegen tile is more capable and further cost-effective to generate energy as a piezotile. Pavegen tile is also lucrative and suitable to equate with other green energy and other power generation sources. Fig. 8 compare pavegen tile with other power generation sources.

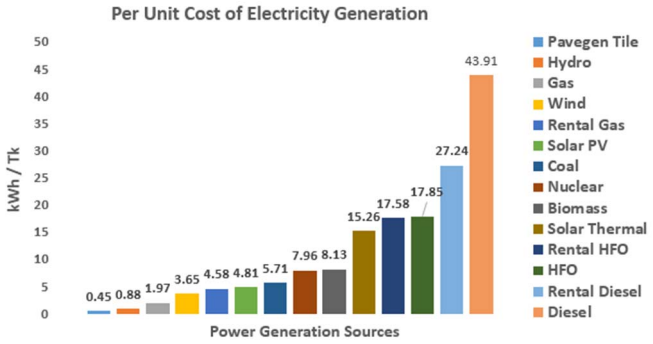


Fig. 8. Piezotile, green energy and other power generation sources per unit cost to generate electricity [10].

C. Observation of Foot Overbridge & Power Calculation

We observe a footbridge (e.g. Farmget Foot Overbridge, Bangladesh) length = 60m and width = 2.5m. The area of the foot overbridge,

$$A = Length * width = (60 * 2.5) = 150 \text{ sq. meters } (m^2).$$

TABLE II. FOOT OVERBRIDGE USER PER MINUTE [11]

Status	Number of people in (5min)	In every Minute
User	470	470/5=94
Non-User	3	3/5

We stated one piezotile for calculation purpose (e.g. Pavegen tile), One Pavegen tile being 17.1" x 23.6" inch. Then Area,

$$a = (17.1*23.6) \text{ Inches} = 403.56 \text{ sq. in}^2 \text{ or } 0.26 \text{ sq. m}^2.$$

So total piezotile need, $(150/0.26)$ sq. $m^2 = 576$ pieces.

Pavegen employees have been reported as saying the titles could cost as little as $76 in the future [9]. Initially, the installation cost is higher but output power will be more efficacious. The distance covered by a single step assuming a stride length of 0.762 meters [12]. One people covered 0.762 Meters equal to 1 Foot Step.

Therefore, one people covered 60 Meters = *79 Footsteps.*

When 94 people pass foot over bridge per minute (TABLE II), then total footsteps;

$$= (\text{Number of people} * \text{Footstep})$$

$$= (94*79) = 7{,}426 \text{ foot Steps/minute.}$$

Piezotile can produce an average of 7 watts per step then, Total Power,

$$P = 7{,}426 \text{ steps} * 7 \text{ watts} = 51.98 \text{ kW.}$$

Power developed for 60 minutes (1 hour),

$$P = 51.98 \text{ kW} * 60 \text{ minute} = 3.119 \text{ MW/h (approx.)}$$

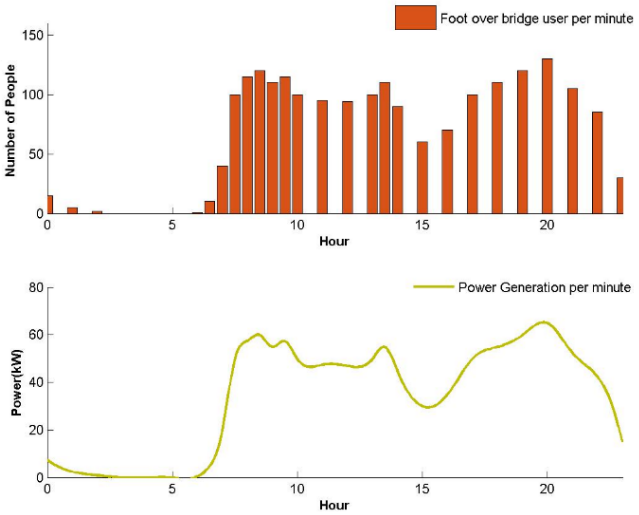


Fig. 9. The waveforms observed when (a) Number of people per minute (b) Power generated per minute.

It is an observed power generating from this curve with respect to time. Observe that from 6 a.m. to around 11 p.m. power is generating. The peak power is seen around from 7-10am, again from 1-2.30pm, and again 5-10pm, from this period maximum usage of the footbridge is observed. In this period, we get maximum power around 60kw and around minimum 30kW power is generated from the pavegen tile. In addition, average 35kW prevails.

D. Environmental Assessment

Piezotile is an environmentally amiable replacement of the burning of fossil fuels; liberate carbon dioxide, which is the leading greenhouse gas, contributor to global warming.

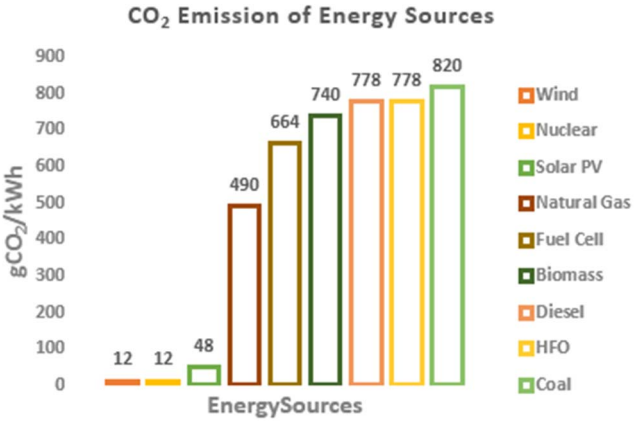


Fig. 10. Carbon Dioxide Emission of various Energy Sources

In various investigation on energy generating tiles piezoelectric transducer that energy produces in its lifetime can save approximately 0.28 ton of CO_2 [13] or emissions

savings of over 10,500 kgCO_2/year [14]. Fig. 10 shows the report on "Global warming potential and emission of CO_2 for selected electricity sources, IPCC, 2014" [15]. Piezotile is capable to produce energy which savings other sources CO_2 emission that is very important and essential for our Environment. In large crowded cities and potential locations where the tiles can be placed and more efficient in the field of renewable energy where electricity can be generated.

IV. CONCLUSION

Piezotile is superior in characteristic, a single piezotile that has the capability to produce enormous energy. It is shown the installation of piezotile on foot over bridge, which is an innovative product; capable of generating a huge amount of power is another way of Renewable energy. In a dense area, if piezotile based foot over bridge is implemented people will be highly benefited and the proper usage of foot stress waste energy.

REFERENCES

[1] J. Evans, "Energy harvesting through the piezoelectric effect at sports venues," Thesis, Humboldt State University, 2015.

[2] T. Gao *et al.*, "Highly oriented $BaTiO_3$ film self-assembled using an interfacial strategy and its application as a flexible piezoelectric generator for wind energy harvesting," Journal of Materials Chemistry A, vol. 3, no. 18, pp. 9965–9971, 2015.

[3] Kenji Uchino, "Piezoelectric Energy Harvesting Systems—Essentials to Successful Developments." Energy Technology, 6.5, 829-848, 2018.

[4] "Pavegen" http://www.pavegen.com/

[5] M. Goldfarb and L. D. Jones, "On the Efficiency of Electric Power Generation With Piezoelectric Ceramic," Journal of Dynamic Systems, Measurement, and Control, vol. 121, no. 3, pp. 566-571, Sep. 1999.

[6] Rod Cross, "Standing, walking, running, and jumping on a force plate." American Journal of Physics 67, 4, 304-309, 1999.

[7] Nise, N. S. "Control Systems Engineering- 7th Edition". Hoboken, NJ: Wiley, 2004.

[8] H. J. Lee, S. Zhang, Y. Bar-Cohen, and S. Sherrit, "High Temperature, High Power Piezoelectric Composite Transducers," Sensors, vol. 14, no. 8, pp. 14526–14552, Aug. 2014.

[9] A. M. Elhalwagy, M. Yousef M. M. Ghoneem, and M. Elhadidi, "Feasibility Study for Using Piezoelectric Energy Harvesting Floor in Buildings' Interior Spaces," *Energy Procedia*, vol. 115, pp. 114–126, Jun. 2017.

[10] A. H. Khan, A. Islam, M. Islam, T. E. Nur, and S. Ghose, "A Review on Electricity Generation and Evacuation in Bangladesh," *Energy Power*, vol. 5, no. 1, pp. 10–16, 2015

[11] S. Aminuzzaman, SH Shuvo, "Culture of Defying Laws: A Case Study on 'Foot Over-bridges'' in Dhaka City.,'" Deparment of Public Administration, University of Dhaka, 2015.

[12] "Steps to Meters | Kyle's Converter." http://www.kylesconverter.com/length/steps-to-meters/

[13] Z. L. Seow, S. T. Chen, and N. B. Khairudin, "An investigation into energy generating tiles : Pavegen," 2011.

[14] J. S. Partridge, "Human Energy Harvesting in the Urban Environment," Doctoral, UCL (University College London), 2015.

[15] "Life-cycle greenhouse-gas emissions of energy sources," Wikipedia.

Aspects of Numeral Value of Cu_2ZnSnS_4 (CZTS) Absorber Layer Fabricated in Sol-gel Process by Dip Coating Technique

Sabiha Dilruba[1], Tabassum Perveen Ananna[2], Afrina Sharmin[3], Zahid Hasan Mahmood[4]
[1,2]*Institute of Energy, University of Dhaka,* Dhaka, Bangladesh
[3]*IFRD, BCSIR,* Dhaka, Bangladesh
[4]*Department of Electrical and Electronic Engineering, University of Dhaka* ,Dhaka, Bangladesh
Email: sabihadilruba11@gmail.com, tabassum.perveen.ananna@gmail.com, afrina.sharmin@gmail.com, zahid@du.ac.bd

***Abstract*— This paper represents the analysis of structural and optical properties of CZTS films grown by dip-coating and its application as an absorber in thin film solar cells. The CZTS thin films were prepared in ordinary glass substrates for a different number of depositions. The optical band gap and absorption coefficient of a CZTS absorber layer related to the concentration of aqueous precursors containing copper (II) chloride dehydrate ($CuCI_2.2H_2O$), zinc (II) chloride ($ZnCI_2$), tin (II) chloride ($SnCI_2$) and thiourea ($CS(NH_2)_2$) were determined. The solution was prepared by dissolving cation metal salts and thiourea in 2-methoxyethanol and monoethanolamine. It is observed that two times deposited film shows good quality Kesterite structure of CZTS is revealed by the XRD. Optical identifications have been carried out using UV-Vis spectroscopy techniques. The direct band-gap of 1.48 eV was found in one sample which is very close to the optimal value as an absorber layer and absorption coefficient was found greater than 10^4 cm^{-1} which establish the characteristics of leading CZTS phase in the film.**

Keywords*— Cu_2ZnSnS_4, Thin films, Dip-coating method, Sol-gel, Annealing, Band gap, X-ray diffractometer.

I. INTRODUCTION

In Thin-film solar cells industry, among various compound material, Polycrystalline cadmium telluride (CdTe), copper indium diselenide (CIS), and copper indium (gallium) diselenide (CIGS) are being utilized for materialistic purpose. Due to restrictions on heavy metal usage for Cd and scarcity of In and Te have raised concern about limitations on fabrication capacity of the PV devices. To resolve this issue, it is necessary to develop another light absorber materials that are both nontoxic and simply available. In this situation, Cu_2ZnSnS_4 (CZTS) quaternary semiconductor compound has flourished as one of the promising candidate.[1] CZTS, as an auspicious thin-film solar-cell material, has drawn great attention in the past few years [2-4]. CZTS kesterite structure has a high absorption coefficient and a direct band-gap of (1.4-1.6 eV), which is reported in previous literature [5]. Furthermore, all the elements in CZTS are non-toxic and earth-abundant [6-9].

Sol process is correlated with the sol-gel method since after producing the solution, a consecutive gel can be formed in one-step. Definition of the sol-gel is as follows: "Formation of an oxide network through polycondensation reactions of a molecular precursor in a liquid" [10]. The final particles can be amorphous or crystalline. Generally, the sol or colloidal suspensions of particles are ruled by the Van der Waals forces or hydrogen bonds. The basis of sol-gel synthesis is to dissolve a precursor compound in a liquid and bringing it back as a solid in a regulated manner [10, 11]. Multi-component systems with a preferred stoichiometry by mixing sols of different compounds can be achieved by Sol-gel synthesis [12]; preventing problems of precipitation and intimate mixing of reagents at the atomic level by producing small particles. The conductivity of semiconducting junctions provides the basic diode mechanism, which coupled with the sensitivity to light and heat and variable resistivity is the main property used to build photovoltaic cells [13,14]. A traditional dyeing method of speedy deposition of polymer films over a large region is the Dip coating process. Dip coating means submerging of a substrate into a tank containing coating material, removing the piece from the tank, and allowing it to drain. The process is performed maintaining precise speed fulfilling controlled temperature and atmospheric conditions. The viscosity of the liquid, the dipping speed, the congeal content are some of the factors used to calculate the coating thickness. Because of simplicity, this method highly depends on automation. Many factors contribute to determining the final state of the dip coating of a thin film [15]. Heterogeneous types of dip-coated film structures and thicknesses can be fabricated by controlling many factors such as submersion time, the number of dipping cycles, solution composition, concentration and temperature, and environment humidity. The dip coating technique can give uniform, high-quality films even on bulky, complex shapes[16].

II. EXPERIMENTAL DETAILS

The sol-gel solution for CZTS precursors was prepared by dip coating and solution containing copper(II) chloride dehydrate ($CuCI_2.2H_2O$) (l.69M), zinc (II) chloride ($ZnCI_2$) (0.90M), tin (II) chloride ($SnCI_2$) (0.787M) and thiourea ($CS(NH_2)_2$) (6.665M), dissolved in 2-methoxyethanol. The mixture was stirred on the hotplate and magnetic stirrer at 45-55 °C for ½ hour. Meanwhile, the mono-ethanolamine were added (a few drops as a stabilizer). Then, the mixture incubated at room

temperature at least 24 hours to yield a homogenous and transparent yellow solution. Lastly, the yellowish mixture from the sol-gel process was fabricated on SLG (soda lime glass) substrates. The glass substrates were cleaned through methanol solution and after that DY water. Dip coater, PTL-MM01 was used to deposit the CZTS film on the cleaned soda lime glass substrates for withdrawing speed about 150 mm per minute. Then the wet films have been dried on an oven at 200 °C for 10 min. Sample is put into the environment for natural cooling after drying. The process was repeated about 1 and 2 times. The sample name is denoted by A-1 for one time dipped and A-2 for two times dipped. Fig. 1 represents sample A-1 and A-2.



Fig. 1. Fabricated samples of 1 and 2 times deposited CZTS thin Film

After that, the samples were annealed at 250°C for 15 min. The phases and crystal orientation analysis were identified by using X-ray spectrometer with Cu Kα1 radiation (λ = 1.5406 Å) collected from 2θ = 22°- 60°. The optical absorption and optical band gap is calculated by transmittance spectra on the wavelength range of 400 – 1200 nm measured by UV Vis spectrophotometer. The thickness of the samples was measured through Stylus profiler in this research work.

III. RESULTS AND DISCUSSION

A. Structural properties

Fig. 2 exhibits the XRD patterns of sample A-1 and A-2 of CZTS films prepared by dip coating technique. Broad peaks were obtained at the 2 Theta positions 28°, 47.40° and 55.40° for the films which corresponds to the (1 1 2), (2 2 0), (3 1 2) planes of CZTS thin films and they are distinct characteristic of tetragonal type kesterite structure of CZTS (ICDD card no: 26-0575)[17]. In Fig. 2, the films annealed at 250°C were identified by using XRD as a kesterite CZTS compound with Zn_2SnO_3[18] which corresponds to the (2 0 0) and (2 2 0)planes. The films revealed poor adhesion to the substrate because of the low temperature of annealing.

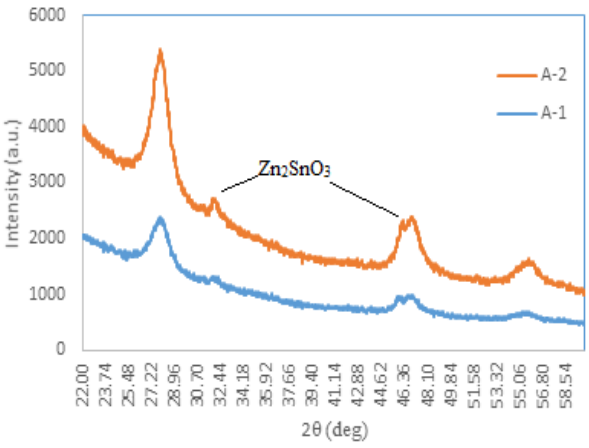


Fig. 2. XRD pattern of samples A for 1 and 2 times deposited CZTS film

B. Optical properties

The optical properties were calculated by taking on UV Vis spectrophotometer the transmittance spectra shown in Fig. 3 for 1 and 2 times deposited film which is recorded for wavelengths of 400 to 1200 nm. It has been observed that the two samples exhibit poor transmittance in the visible region for low-temperature annealing.

In the case of CZTS thin film absorber layer, low transmittance is desired so the samples variation were usually the same in transmittance vs wavelength curve. The other optical characteristics manifested variation so comparable result can be found.

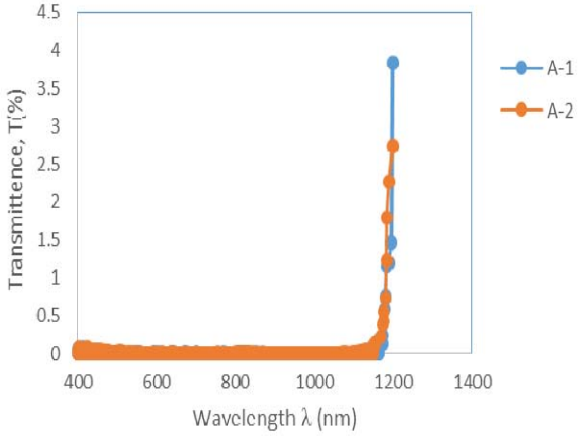


Fig. 3. Transmittance Vs Wavelength of 1 and 2 times deposited CZTS thin Film

The absorption coefficient (α), of CZTS films, was assessed from the transmittance data as demonstrated in Fig. 4[19].

$$\alpha=(1/d)\ln(1/T) \quad (1)$$

Where d is film thickness and T is transmittance. The absorption coefficient is $10.3x10^4cm^{-1}$ and $6.9x10^4cm^{-1}$ respectively of samples A-1 and A-2. The absorption

coefficient for different layer decreases with increasing depth.

The optical band gap of CZTS films is calculated as shown in Fig. 5 from the accordance to Tauc's formula [20].

$$(\alpha h\upsilon)=A(h\upsilon-Eg)^{m} \quad (2)$$

Where A is a constant, the exponent "m" can take values ½ for direct allowed, 2 for indirect allowed, 3/2 for direct forbidden and 3 for indirect forbidden. The band gap energy was enumerated by plotting $(ahv)^2$ versus the hv in Fig. 5, therefore, the direct band gaps (Eg) of CZTS annealed at 250°C during 15 minutes were estimated to be 1.40 eV and 1.48 eV respectively of samples A-1 and A-2. The band gap energy was found to be near 1.5 eV, which is in good agreement with previously reported value [21].

The band gap increases with the increase in thickness. It is found that the band gap of sample A-2 is near to the optimum for photovoltaic applications.

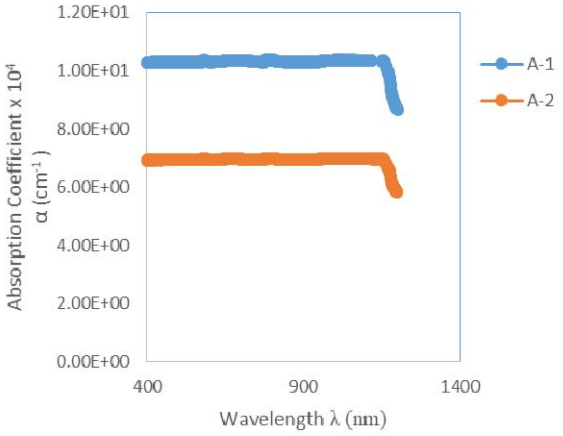


Fig. 4. Absorption coefficient of A-1 and A-2 layered CZTS thin films as a function of wavelength λ

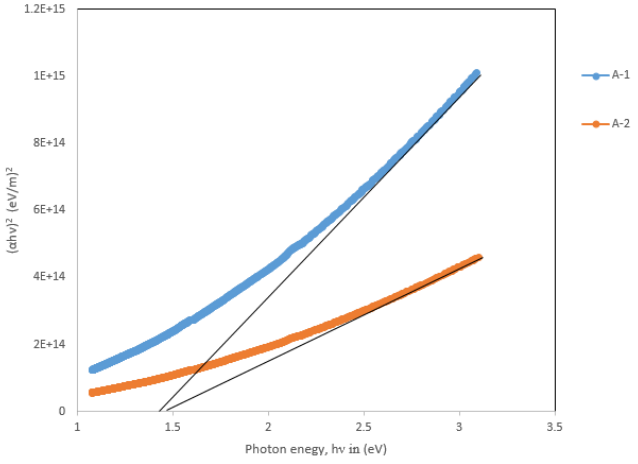


Fig. 5. The values of the energy gap $(\alpha hv)^2$ in terms of hv of CZTS thin films

TABLE I. OPTICAL PROPERTIES OF CZTS WITH VARIOUS LAYER

Sample name	A-1	A-2
Thickness (μm)	1.92	2.85
Optical band gap (eV)	1.40	1.48
Absorption coefficient(cm^{-1})	$10.3x10^4$	$6.9x10^4$

The extinction coefficient (k), was determined from α using the following relation [22]

$$\kappa=\alpha\lambda/4\pi \quad (3)$$

It is seen from Fig. 6 that the extinction coefficient ranges between the extremes of 0.228 and 0.917 for the samples of A-1 and A-2 layer in the wavelength range of 400- 1200nm. The behavior of curves appears to be related where the value of "k" increases with increasing wavelength. as "k" is directly proportional to the absorption coefficient (eq.3). Also, since the absorption coefficient of the CZTS is very high (>10^4 cm^{-1}), so "k" will be high.

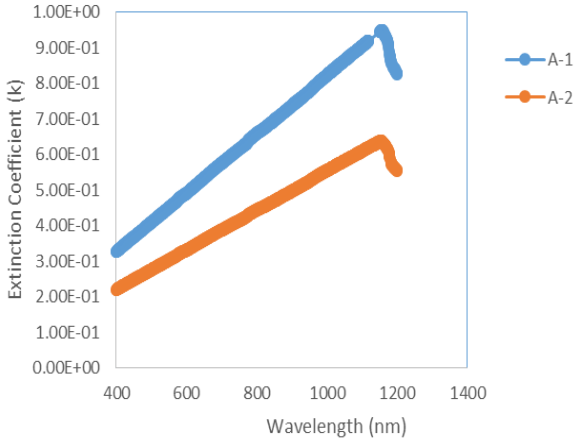


Fig. 6. Extinction coefficient of A-1 and A-2 layered CZTS thin films as a function of wavelength

IV. CONCLUSION

This research was observant to the synthesis and characterization of non-stoichiometric Cu_2ZnSnS_4(CZTS) absorber layers. Dip coating technique and anneal at air atmosphere had proclaimed to be convenient for kesterite CZTS thin film composition. The films exhibit poor adhesion to the substrates because of low annealing temperature. Several impurities are the cause of the secondary phase in the films. Here, the number of layers were varied and sample A-2 had a bandgap energy of 1.48eV which is suitable for photovoltaic applications. The X-ray diffraction displays that CZTS has a kesterite structure with a affluent orientation along the (112) plane [17]. It can be seen that the band gap increases with increase in thickness and absorption coefficient decreases at the same time. For this reason, the extinction coefficient decreases as α has proportional relation to k.

ACKNOWLEDGEMENT

This research was supported by the Solar cell Fabrication Lab of IFRD (Institute of Fuel Research and Development (IFRD), Bangladesh Council of Scientific and Industrial Research (BCSIR), Dhaka.

REFERENCES

[1] School of Engineering System, Queensland University of Technology, Brisbane, QLD 4001, Australia, “Progress in Thin Film Solar Cells Based on Cu2ZnSnS4” Hindawi Publishing Corporation International Journal of Photoenergy Volume 2011, Article ID 801292, 10 pages.

[2] T. K. Todorov, J. Tang, S. Bag, O. Gunawan, T. Gokmen, Y. Zhu, and D. B. Mitzi, AdvancedEnergy Materials 3, 34 (2013).

[3] D. B. Mitzi, O. Gunawan, T. K. Todorov, K. Wang, and S. Guha, Solar Energy Mater. Solar Cells 95, 1421 (2011).

[4] S. Siebentritt, N. Papathanasiou, J. Albert, and M. C. Lux-Steiner, Appl. Phys. Lett. 88, 151919 (2006)

[5] D. Seo , S. Lim "Effect of sulfur and copper amounts in sol- gel precursor solution on the growth, crystal properties, and optical properties of Cu2ZnSnS4 films" (2013) J Mater Sci: Mater Electron DOr 10.1007/sI0854-013-1314-z.

[6] S. Chen, A. Walsh, X. Gong, and S. Wei, Adv Mater 25, 1522 (2013).

[7] S. Chen, X. Gong, A. Walsh, and S. Wei, Appl. Phys. Lett. 96, 021902 (2010).

[8] S. Chen, J. Yang, X. Gong, A. Walsh, and S. Wei, Physical Review B 81, 245204 (2010).

[9] P. Xu, S. Chen, B. Huang, H. Xiang, X. Gong, and S. Wei, Physical Review B 88, 045427 (2013)

[10] G. Korotcenkov. Handbook of Gas Sensor Materials: Properties, Advantages and shortcoming for applications (2), new trends and technologies. Integreted analytical systems, (2), 396, (2014).

[11] R. W. Schwartz and M. Narayanan, Solution Processing of Chalcogenide Semiconductors via Dimensional Reduction, chapter 2, 33–76, John Wiley & Sons, Inc. (2008).

[12] R. D’Angelo, C. L. Azanza Ricardo, A. Mittiga, P. Scardi, M. Leoni. A water- and sulphurization- free solution route to Cu2-xZn1+xSnS4. J. Sol-Gel Sci. Technol. 72, 490-495, (2014).

[13] Feynman, Richard. Feynman Lectures on Physics. Basic Books, (1963).

[14] A. Walsh, S. Chen, S. H. Wei, and X. G. Gong. Kesterite. Thin-Film Solar Cells: Advances in Materials Modelling of Cu2ZnSnS4. Adv. Energy Mater. 2, 400–409, (2012).

[15] Rahaman, M.N. (2007). Ceramic Processing. Boca Raton: CRC Press. pp. 242–244. ISBN 0-8493-7285 2.

[16] Quéré, David (1999). "FLUID COATING ON A FIBER". Annual Review of Fluid Mechanics. 31 (1): 347–384. ISSN 0066-4189. doi:10.1146/annurev.fluid.31.1.347.

[17] ICDD, card no: 26-0575

[18] N. Nakayama, K. Ito, Applied Surface Science 92 (1996) 171-175.

[19] K. Zweibel, J. Mason, V. Fthenakis, "A Solar Grand Plan", Scientific American, Jan 2008. CdTe PV is the cheapest example of PV technologies and prices are about 16¢/kWh with US Southwest sunlight.

[20] Peng et al. (2013). "Review on life cycle assessment of energy payback and greenhouse gas emission of solar photovoltaic systems". Renewable and Sustainable Energy Reviews 19: 255–274. doi:10.1016/j.rser.2012.11.035.

[21] Tanaka K, Fukui Y, Moritake N, Uchiki H (2011) chemical composistion dependence of morphological and optical properties of Cu2ZnSnS4 thin films deposited by sol–gel sulphurization and Cu2ZnSnS4 thin films solarcell efficiency. Sol Energy Mater Sol Cells 95:838

[22] S. Islam, M. A. Hossain, H. Kabir, M. Rahaman, M. S. Bashar, M. A. Gafur, A. Kabir, M. M. R. Bhuiyan, F. Ahmed, N. Khatun, “Optical, Structural and Morphological Properties of Spin Coated Copper Zinc Tin Sulfide Thin Films”.Int. J. Thin. Fil. Sci. Tec. 4,No. 3, 155-161 (2015)

A Statistical Estimation of Solar Power for Energy Mix in Bangladesh

Sujoy Barua
Dept. of Electrical and Electronic Engineering
Chittagong Institute of Engineering and Technology
Chittagong, Bangladesh
sujoy.baruaa@gmail.com

Zainal Abedin
Dept. of Computer Science and Engineering
University of Science and Technology Chittagong
Foy's Lake, Chittagong, Bangladesh
zainal.abedin@ustc.ac.bd

Anik Nath
Dept. of Electrical and Electronic Engineering
Chittagong University of Engineering and Technology
Raozan, Chittagong, Bangladesh
anik.nathnu@gmail.com

Chitra Biswas
Dept. of Electronic and Telecommunication Engineering
University of Science and Technology Chittagong
Foy's Lake, Chittagong, Bangladesh
chitra.biswas86@gmail.com

***Abstract*— solar radiation (SR) is a significant parameter for producing solar power for the electric energy mix. Sun radiation data is hard to be accumulated at particular geographical location of a country due to measurement instruments is limited especially in developing country like Bangladesh. Hence, estimation of SR for a meteorological location is an important research to learn solar energy potentiality. In this regards, this paper presents a statistical inference model for the estimation of bright sunshine hour which is important metric for solar irradiance. Normal distribution model is inferred from the historical data of some important locations of Bangladesh. Mean and standard deviation are the parameters at 95% confidence interval which control the behavior of such distribution. In this statistical model, bright sunshine hour is estimated that expedite to find the potential of solar energy at particular location. The estimation model can be used as a decision making support to design photovoltaic power system for deployment of energy mix policy.**

Keywords—radiation, sunshine hour, statistical, solar energy, energy mix

I. Introduction

Power Generation mix refers to the combination of the various primary energy sources (fossil fuels, nuclear energy, renewable energy sources) using to meet electricity according to geographical region. The composition of the electricity generation mix depends on the availability of usable resources domestically or the possibility of importing it. Country's Policy for power generation defines historical, economic, social, demographic, environmental and geopolitical factors. These differences can be appreciated by taking a look at the generation and consumption of electricity information for a country.

Bangladesh is a developing country with 162.95 million populations. Population growth annually is 1.1% with density 1251.8 per sq. km of land area. In 2016 the Electric power consumption of per capita was 310 kWh for 0.46 metric ton carbon emission per capita. Information on Bangladesh is shown in Table I [1]. Fig.1 shows power generation mix of Bangladesh from the year 2001 to 2016 [1]. There is six power generation mix used (Gas, Oil, Hydro, Coal, Renewable and Nuclear) for generating electricity in Bangladesh but nuclear yet contribute. Gas is the dominant sources, above 80% of power generation-mix in Bangladesh. Presently, Contribution of coal is less significant but there are projects to be commissioned by 2021 [2]. The Nuclear power plant will be implemented and on an operation by 2024 [3] According to World Bank, renewable energy share is small amount into generation mix included hydropower plant in 2016.

TABLE I: COUNTRY INFORMATION IN 2016

Population, total (millions)	162.95
Population growth (annual %)	1.1
Surface area (sq. km) (thousands)	147.6
Population density (per sq. km of land area)	1,251.8
CO_2 emissions (metric tons per capita)	0.46
Electric power consumption (kWh per capita)	310
Access to electricity (% of the population)	75.92%

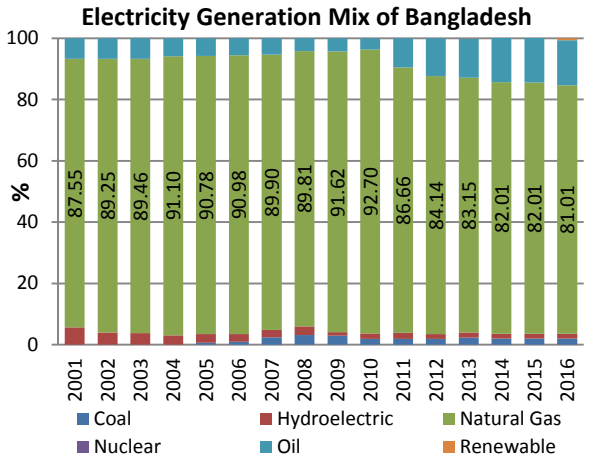


Fig. 1. Electricity Generation Mix (2001-2016) in Bangladesh

There are immense opportunities to contribute a large amount of solar energy as a "Power Generation Mix" in Bangladesh. Bangladesh is a south Asian country located in between 20.30°- 26.38° North latitude and 88.04° - 92.44° east longitude [4]. Therefore, it is an ideal subtropical location for solar energy utilization and 70% of sunlight is plentiful in the year [5]. The government has set a target to generate 2000MW of electrical power from renewable energy by 2021[6]. The new target of renewable energy would be 10% of the total electricity generation in 2021 and would increase to 20% percent by 2030 [7].

It is important to know the exact bright sun shine hour and solar energy potential at a particular geographic location. So that, solar information database can be built for research. In research, numerical and regression model was

established to calculate solar energy using sunshine hour [8] [9]. A comparative study of different types of correlation models had been done to find the solar radiation data [10] [11]. Numerical approaches using temperature base method and Weibull distribution were used to estimate solar radiation data at certain location [12]. In this paper, the historical data of bright sunshine hour at a particular geography location is analysed and applied to normal distribution to get stable bright sunshine hour. Standard deviation and error are determined how much the data stable at a particular geography location by utilizing statistical approaches. The Determinate bright sunshine hour has been applied to per kilowatt solar panel for calculating solar energy potential in the certain area.

II. Methodology

A. Setting Boundaries

Four different districts are selected for this analysis of solar energy potential. These four districts are having distinct geographical locations which are example of rest of districts in Bangladesh. These are shown in Table II [13]

TABLE II. District Geographical Location

Districts	Distinctness	Coordinates	Elevation
Hatia	Island	22°22' N , 91°7.5' E	4 m (13ft)
Rangamati	Hill Station	22°38' N , 92°12' E	11m (56ft)
Cox's bazar	Sea Beach	21°58' N , 92°01' E	3 m (10ft)
Dhaka	Town-Capital	23°42' N , 90°22' E	4m (13.12ft)

B. Research Methodology

The figure 2 presents the flowchart to simulate the research. Bright sunshine hour of different meteorological locations of Bangladesh as samples are considered to make an inference in building the statistical model. Some pre-processing such as normalization, outlier detection and missing values adjustment are applied to the input datasets in order to

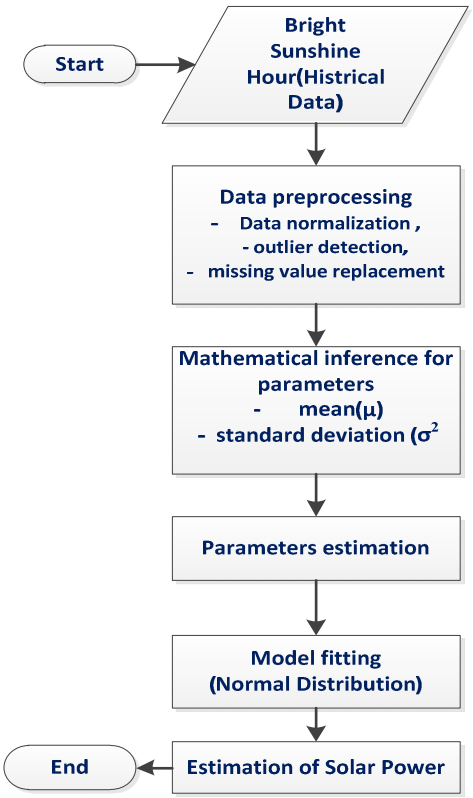


Fig. 2: Steps of Methodology

enhance the fitness of the data. To estimate the parameters (mean and standard deviation) of statistical model, mathematical inference is deducted. After estimation of those parameters, the probability distribution function (PDF) is established and fitted. Finally, solar power of every month can be estimated from the inferred model for different meteorological locations of Bangladesh.

C. Data Collection

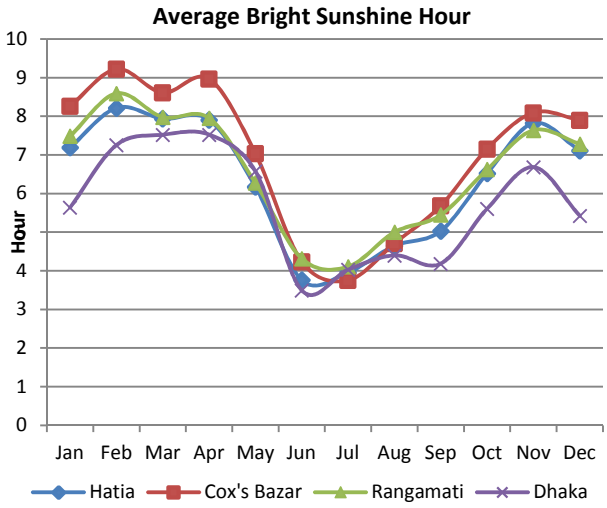


Fig. 3: Monthly average sun hour of 4districts

Bangladesh Agriculture Research Council (BARC) has developed climate information management system that is the website where are six types spatiotemporal weather data available. District wise bright sunshine hour data has been collected from climate information management system for last 13 years that is shown in figure 3 [14].

D. Statistical Parameter of Estimation [15]

Sunshine Hour (SH) is a decisive meteorological parameter for the prediction of solar energy. SH is a random variable with time which can by characterized by probability distribution function such as Gaussian or Normal distribution. The normal or Gaussian distribution is defined by the following distribution function:

$$F(SH) = \frac{1}{\rho\sqrt{2\pi}} e^{\frac{-(SH-\mu)^2}{2\sigma^2}} \quad \ldots\ldots\ldots\

Therefore, the likelihood of an outcome is mathematically defined by the following inequalities:

$$L(x_1,.....,x_n, \mu\ \sigma^2) = \prod_{i=1}^{n} f(x_i, \mu, \sigma^2)$$

$$=(\frac{1}{2\pi\sigma^2})^{\frac{n}{2}}\exp\{-\sum_{i=1}^{n}(x_i-\mu)^2/2\sigma^2\}$$

To maximise the log-likelihood, derivatives with respect to θ_1 and θ_2 are to be calculated and set to be zero and log-likelihood is given by the following equation:

$$ln(L) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}$$

Performing derivatives with respect to the μ and σ^2 and setting to zero, yields the following parameter estimations:

μ (estimated mean) = $\bar{x}$ (sample mean) ……. *(2)*

$$\partial^2 = \frac{\sum_{i=1}^{n}(x_i-\hat{\mu})^2}{n} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n} \text{ ……..... (3)}$$

A confidence interval (p) covers "plausible" values of the parameter. It is a simple arrangement of the point estimate along with estimated standard error. A confidence interval is connected with a confidence level, named 1−α and which express the confidence that the parameter μ essentially lies within the prescribed confidence interval. Confidence levels of 90%, 95%, and 99% are typically used, which correspond to α values of 0.10, 0.05, and 0.01, respectively. In our experiment, 95% confidence interval is considered.

F. Calculation of Energy Production

Solar Panel 1000W at Standard Testing Conditions (STC) needs sunlight to produce the maximum solar output. The solar panel is mounted facing south and in an area where is no shaded for any reason (such as building, tree). It is important that Solar panel can be produced maximum amount of energy during at peak hours.

A peak sun-hour, specifically, is an hour during which the intensity of sunlight is 1,000 watts per square meter during a typical day. And hours of daylight is time duration of light in a typical day. The amount of solar radiation, delivered by the sun varies throughout the day, based on the sun's position in the sky, clouds, and other atmospheric conditions. So availability of energy production is calculated using equation (4).

$$Energy\ (KWh) = bright\ sunshine\ hour \times capacity\ of\ Solar\ PV\ Pane \text{……..} (4)$$

III. EXPERIMENTAL RESULT

The normal probability distribution is characterised by two parameters (mean, μ and variance, σ). In regards, mean (μ) and variance (σ) represents average distributed sunshine hour and variation of this sunshine hour for the respective location respectively. In this work, the distribution of sunshine hour is fitted by the normal probability distribution function. The parameters of this distribution are estimated by using the equation (2) and equation (3). The samples are taken from historical mean sunshine hour (SH) recorded in different meteorologically locations.

TABLE III: NORMAL DISTRIBUTION OF BRIGHT SUNSHINE HOURS (μ) VARIANCE (σ) IN FOUR DISTRICTS OF BANGLADESH.

	Hatia		**Cox's Bazar**		**Rangamati**		**Dhaka**	
	μ	σ	μ	σ	μ	σ	μ	σ
Jan	7.19	1.06	8.26	0.72	7.48	0.90	5.64	1.03
Feb	8.22	1.16	9.22	0.83	8.59	0.86	7.25	0.90
Mar	7.94	0.99	8.61	0.94	7.98	0.87	7.52	0.71
Apr	7.91	1.04	8.97	0.78	7.95	0.75	7.52	0.83
May	6.18	1.19	7.04	1.19	6.28	0.91	6.58	1.22
Jun	3.76	1.13	4.24	1.08	4.32	0.77	3.49	1.07
Jul	3.96	0.97	3.75	1.23	4.11	0.89	4.03	0.90
Aug	4.65	1.15	4.71	1.05	5.00	1.13	4.40	1.06
Sep	5.03	0.75	5.68	1.07	5.45	0.65	4.18	0.84
Oct	6.52	0.65	7.15	0.91	6.62	0.47	5.61	0.74
Nov	7.82	0.87	8.09	1.07	7.64	1.17	6.68	0.87
Dec	7.12	1.09	7.90	1.18	7.28	0.72	5.42	1.30

Here Cox's Bazar, Dhaka, Hatia and Rangamati are considered to simulate for the statistical model. For each month, a statistical model (normal distribution fitting) is developed with the estimation of population mean and variance at 95% confidence interval. Fig. 4 to 7 illustrates the normal distribution fit with mean and variance of September month only for those locations. And table III shows mean (μ) value (estimation bright sunshine) and variance (σ) for four districts of Bangladesh. Finally, a statistical estimation is represented solar energy potentiality in different districts in Bangladesh using the equation (4). Table IV represents minimum and maximum solar energy potential in respective location in Bangladesh. Bright sunshine is taken for minimum (μ- σ) and maximum (μ+ σ) sunshine hour from the normal distribution model. And STC condition is considered 1000W solar Photovoltaic (PV) panel for calculation monthly average solar energy production in respective location.

TABLE IV: SOLAR ENERGY PRODUCTION (ESTIMATED KWH/DAY at STC) IN FOUR DISTRICTS OF BANGLADESH.

Month	**Hatia**		**Cox's Bazar**		**Dhaka**		**Rangamati**	
	Estimated Energy production at KW		Estimated Energy Production at KW		Estimated Energy Production at KW		Estimated Energy Production at KW	
	Min (kwh/day)	Max (kwh/day)	Min (kwh/day)	Max (kwh/day)	Min (kwh/day)	Max (kwh/day)	Min (kwh/day)	Max (kwh/day)
Jan	6.131	8.254	7.538	8.986	4.610	6.667	6.583	8.386
Feb	7.056	9.374	8.397	10.049	6.356	8.152	7.730	9.455
Mar	6.946	8.931	7.667	9.548	6.812	8.234	7.103	8.851
Apr	6.867	8.948	8.192	9.746	6.690	8.356	7.199	8.708
May	4.986	7.368	5.853	8.223	5.358	7.796	5.367	7.187
Jun	2.636	4.887	3.161	5.316	2.421	4.563	3.544	5.087
Jul	2.992	4.931	2.519	4.989	3.134	4.927	3.213	5.003
Aug	3.501	5.791	3.661	5.754	3.339	5.461	3.873	6.127
Sep	4.279	5.783	4.615	6.754	3.349	5.020	4.799	6.094
Oct	5.877	7.169	6.242	8.065	4.872	6.343	6.148	7.098
Nov	6.950	8.696	7.018	9.167	5.814	7.555	6.471	8.806
Dec	6.030	8.201	6.721	9.079	4.124	6.723	6.560	8.009

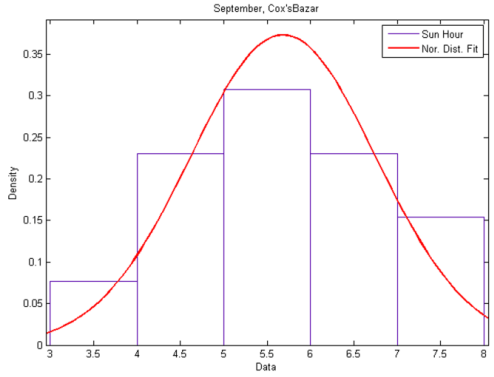


Fig. 4. Normal Distribution Fit of Sunshine Hour μ=5.68462 at standard error =0.296528 and variance σ=1.06915 at standard error =0.222933 of Cox'sBazar in September.

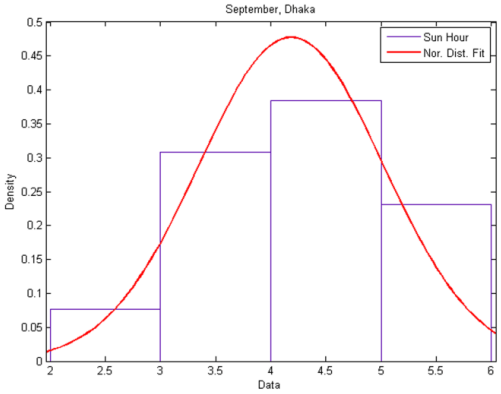


Fig. 5. Normal Distribution Fit of Sunshine Hour, μ= 4.18462 at standard error=0.231729 and variance σ=0.83551 at standard error =0.174216 of Dhaka in September.

IV. Conclusion

In solar energy technology, bright sunshine hour is the key meteorological parameter for harnessing solar energy by solar photo-voltaic (PV) panel. Estimation of solar energy from measured sunshine hour offers an important alternative in absence of measured solar radiation. This paper developed statistical model to figure out monthly approximation of sunshine duration at a particular location. The performance of each model is evaluated and compared using statistical error parameters.

Due to random behaviour of meteorological parameter, the statistical analysis has been chosen and inference models are developed from which sunshine hour can be predicted where no measuring instrument is available. This piece of knowledge could be utilized to assess solar energy potentiality of particular location. Here, four districts are randomly chosen from different geographical locations of Bangladesh to show numerical calculation of solar energy potential using MatLab. And similar study can be done for every district of Bangladesh. Government of Bangladesh is very concern now to generate power from numerous renewable energy sources such as solar, coal and nuclear as electricity generation mix policy. In this regard, this paper can be used as an artificial tool to make decisions in energy generation policy by the incorporation of solar power for green technology and, to reduce the consumption of natural gas.

REFERENCES

[1] World bank data Indicator, Available at databank.worldbank.org
[2] Bangladesh Power Development Board, "Annual Report 20016-2017" Available at: www.bpdb.gov.bd
[3] Power Reactor Information System. International Atomic Energy Agency. 2018-08-08. Retrieved 8 August 2018.

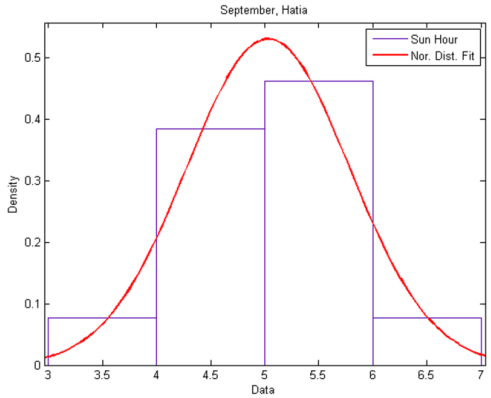


Fig.6. Normal Distribution Fit of Sunshine Hour μ=5.03077 at standard error= 0.208593 and variance σ= 0.752091 at standard error= 0.156822 of Hatia in September

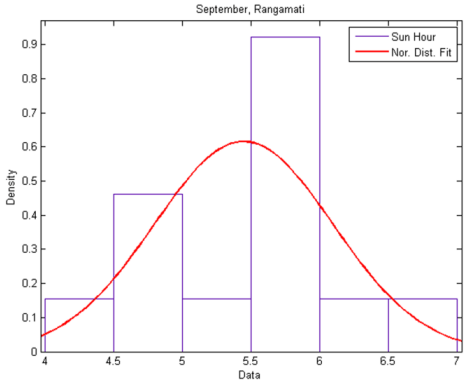


Fig.7. Normal Distribution Fit of Sunshine Hour μ= 5.44615 at standard error =0.179606 and variance σ = 0.647579 at standard error= 0.13503 of Rangamati in September.

[4] Sujoy Barua and M. M. Hasan, "A Study of Socioeconomic Aspects of Motor Assisted Pedal Rickshaw and its PV Solution", IEB Journal of Electrical Engineering, Vol. EE 40, No. I & II, pp.32-36, June & December, 2014
[5] Md Niaz Murshed Chowdhury. 2015. "Present Scenario of Renewable Energy and Non Renewable Energy Sources in Bangladesh: A Compact Analysis" Business and Economics Journal.
[6] "Target to Generate 2,000 MW Renewable Energy in 2021", The Independent Published 24 Jan 2016. Available:www.theindependentbd.com/printversion/ details/31438 [access: 17 Sep 2018]
[7] Shariful Islam Sharif, Md. Anisur Rahman Anik, Md. Al-Amin, Md. Abu Bakr Siddique, "The Prospect of Renewable Energy Resources in Bangladesh: A Study to Achieve the National Power Demand", Energy and Power 2018, 8(1): 1-6
[8] A. De Miguel, J. Bilbao, S. Salson and A. Lage, Solar radiation and sunshine hour maps in Castilla and Leon region (Spain), Renewable Energy, Vol. 4. No. 8, pp.933-940, (1994).
[9] K. K. Gopinathan, A general formula for computing the coefficient of the correlation connecting global solar radiation to sunshine duration. Solar Energy, Vol. 41. No. 6, pp. 499-502. (1988).
[10] Shuvankar Podder, Raihan Sayeed Khan, Shah Md. Ashraful Alam Mohon, Md. Jubair Hussain, and Enamul Basher , "Solar Radiation Approximation Using Sunshine Hour at Patenga, Bangladesh" 8th International Conference on Electrical and Computer Engineering 20-22 December, 2014, Dhaka, Bangladesh.
[11] A.Q. Jakhrani, A. K. Othman, A. R. H. Rigit and S. R. Samo, "A simple method for the estimation of global solar radiation from sunshine hours and other meteorological parameters", IEEE ICSET 2010 6-9 Dec 2010, Kandy, Sri Lanka.
[12] Majid Vafaeipour, Mohammad H. Valizadeh, Omid Rahbari, Mahsa Keshavarz Eshkalag "Statistical Analysis of Wind and Solar Energy Potential in Tehran" International Journal Of Renewable Energy Research Vol.4, No.1, 2014
[13] Bangladesh Republic of Bangladesh site: www.bangladesh.govt.bd [Accessed: 10 Sep 2018]
[14] Climate information management System under Bangladesh Agriculture Research Council, site: climate.barcapps.gov.bd [accessed: 07 Sep 2018]
[15] Anthony Hayter, "Probability and Statistics for Engineers and Scientists", Fourth Edition, ISBN-13:978-1-111-82704-5

A Comparative study on Waste-to-energy (WTE) technologies for municipal solid waste (MSW) management in Chittagong, Bangladesh

Md. Rashidul Islam
Dept. of Electrical & Electronics Engineering
Internatinal Islamic Universiity, Chittagong.
Chittagong, Bangladesh
rashidul05eee@yahoo.com

Jobair Al Rafi
Dept. of Electrical & Electronics Engineering
Internatinal Islamic Universiity, Chittagong.
Chittagong, Bangladesh
mjrafi7234@gmail.com

Md Sajjad Hossain
Dept. of Electrical & Electronics Engineering
Internatinal Islamic Universiity, Chittagong.
Chittagong, Bangladesh
sajju151040eee@gmail.com

***Abstract*— Currently, energy security and environmental cleanliness are becoming an important concern to the civilized world due to increasing population and global warming, which motiving the researcher towards finding the alternative energy resources instead of traditional energy source for power generation. Now, Waste-to-Energy technology (WTE) has earned a great deal of attention both the developed and developing countries for an effective waste management system. This study aiming at finding a more suitable and effective method among WTE technologies in perspective of Bangladesh's economy, energy efficiency, and environmental aspects that can sustain with our future energy demand. In this paper we initially compared different types of emerging WTE technologies and then these technologies are applied in Chittagong, a city of Bangladesh in different aspects and finally finding out that Incineration-based WTE technology is the most suitable method because of its operation flexibility, lower cost, higher efficiency, less greenhouse (GHG) emissions, and simplicity in design on here. Thus, the energy produced in this method can contribute significantly to reduce our electricity shortage with an effective waste management system and can be a building block of a green-smart city.**

Keywords*— Energy security, Environmental cleanliness, WTE technology, Bangladesh, Incineration, Green-smart city.

I. INTRODUCTION

According to the International Energy Agency, about 22% of the global population lacks energy access [1]. Also, an average of 34% of people in developing countries lack of waste collection service and the waste generation rate is expected to double by 2025 [2]. Moreover, Bangladesh is a developing country with a large population where only 75.92% of people have electricity access and facing load shedding problem due to its gradually increasing power demand [3].

Waste generation in big quantities and lack of appropriate waste management system is one of the most important issues in the present world. Waste-to-Energy technology (WTE) is an emerging technology of the currently developed country. It can also be a good solution for the scarcity of electricity as well as the management of solid waste in Bangladesh. We know that waste everywhere needs urgent disposal which requires manpower, cost, and proper space; so waste management is a challenge for the government and city authority. Many developed countries such as the UK, Middle East, USA, and Venezuela, etc. have already adopted WTE technology for waste management and to produce significant energy from waste [4-5]. Waste-to-Energy (WTE) technologies consist of any waste treatment process that creates energy in the form of electricity, heat or transport fuels from a waste source [6]. WTE technology produces energy using waste as a feedstock. Waste is generated at a large extent in our day to day life. We just throw away this wastes and the governing authority collects them. Instead, if these wastes are utilized using WTE technology then a large amount of energy can be produced. This generated energy can contribute to the growing power demand in a highly dense populated developing country like Bangladesh where waste generation rate is very high. The waste management system should not be only limited to energy and heat production but it should also be environmentally sustainable and cost-efficient. Therefore, determining a more suitable WTE technology in the perspective of Bangladesh is a great challenge.

In this research paper we initially compared different types of emerging WTE technologies and then these technologies are applied in Chittagong in different aspects and finally finding out that Incineration-based WTE technology is the most suitable method because of its feasible surroundings in economy, energy efficiency, and environmental circumstances in perspective of the present conditions of Chittagong city that can sustain with future energy demand. Thus, the energy produced in this method can contribute significantly to reduce our electricity shortage with an effective waste management system and can be a building block of a green-smart city.

II. OVERVIEW OF EMERGING WTE TECHNOLOGIES

WTE refers to the production of Electricity, Heat and Biomass gas from Municipal Solid Waste. Currently emerging WTE technologies used worldwide for energy generation are mainly [7] —

(a) Incineration
(b) Gasification
(c) Anaerobic Digestion
(d) Landfilling.
(e) Pyrolysis.

The main features of these WTE technologies are summarized in the Table I [8].

Within the given WTE technologies, there have certain differences among themselves. We consider here based on the design maximization, requirement of oxygen, treated flue gases and material handling and preparation processes. The major Key differences between WTE technologies are stated in Table II [9].

Net Energy Generation Potential per Ton MSW of Incineration, Gasification, Anaerobic Digestion and Landfilling is given in the Table III [10].

TABLE I. MAIN FEATURES OF WTE TECHNOLOGIES

Incineration	Gasification	Anaerobic Digestion	Landfilling	Pyrolysis
Thermal Conversion	Thermo-chemical Conversion	Biochemical Conversion	Biochemical Conversion.	Thermal-chemical Conversion
Compete oxidation of waste matter	Partial Oxidation of waste matter	Enzymatic decomposition of waste matter	Partial Oxidation of waste matter.	No oxidation process is used.
End product is steam which is used to produce electricity	End product is syngas. Syngas is primarily CO and H_2. Syngas is used to produce electricity	End product is Methane gas which is used to produce electricity	The end product is ignitable methane gas which converts into gas engine and finally into electricity.	End product is Char, bio-oil and syngas. Syngas is used to produce electricity.

TABLE II. MAJOR DIFFERENCES AMONG WTE TECHNOLOGIES

Incineration	Gasification	Anaerobic Digestion	Landfilling	Pyrolysis
Designed to maximize the conversion of feedstock to CO_2 and H_2O	Designed to maximize the conversion of feedstock to CO and H_2	Designed to maximize the conversion of feedstock to CH_4	Designed to maximize the conversion of feedstock to CH_4 to CO_2.	Designed to maximize the conversion of volatile organic product into low molecular weight gas & char.
Large quantities of oxygen needed	Limited quantities of oxygen	No oxygen needed	No oxygen needed.	No oxygen needed.
Treated flue gas discharged to atmosphere	Treated syngas used for chemical production and power production	Treated flue gas is used at some extent but CO_2 is discharged to atmosphere	Treated flue gas and discharged CO_2 to the atmosphere.	Treated syngas used for power production.
Material handling and preparation is complex	Material handling and preparation is simple	Material handling and preparation problems can arise	Material handling and preparation is complex enough due to the large area requirement compared with others.	Material handling and preparation is complex.

TABLE III. ENERGY GENERATION POTENTIAL PER TON MSW

Technology	Energy Potential (KWh per ton MSW)
Incineration	585
Gasification	660
Anaerobic Digestion	250
Landfilling	105
Pyrolysis	660

III. CASE STUDY-CHITTAGONG, BANGLADESH

In WTE technologies, Municipal Solid Waste (MSW) is used as feedstock. In this study, Chittagong city of Bangladesh is considered for implementation of WTE technology. Chittagong is the second largest city and commercial capital of Bangladesh with a large population. The area of Chittagong City Corporation is 155.4 Sq. km located in between 22°13' and 22°27' north latitudes and in between 91°40' and 91°53' east longitudes. The total Population of Chittagong city corporation was 25, 82,401 in 2011 including the floating population 5, 42,965 according to the information from Chittagong City Corporation's website [11] and in 2015 this amount exceeds on 32, 09, 650 [12]. In the Chittagong area, the demand for power is around 1,000 MW, but the grid provides only 700 to 750 MW [13]. So it is very much necessary to produce more power in order to meet this demand. For the design topology boost up we consider the Municipal Solid Waste for five days at a stretch of Chittagong City in 2018 which is shown in "Figure 1".

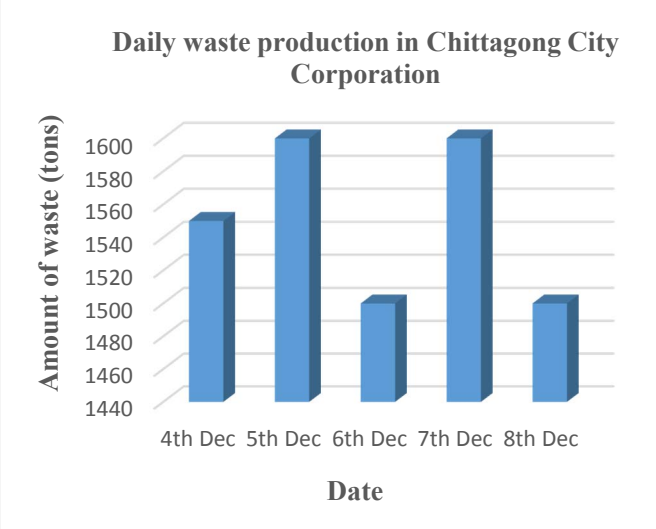


Fig. 1. MSW production in Chittagong city for five days

" Figure 1" provides that the average amount of waste produced in Chittagong city is near about 1550 tons in which we collected the real-time waste production data physically. Satyajit Roy Das et.al. Showed a clear practical analysis about the actual amount of Solid Waste management named after as "Solid Waste Management in Commercial City of Bangladesh" by the Chittagong City Corporation authority and from their analysis we consider here the amount of waste collection is 958 tons per day [12]. There are several types of waste are produced in Chittagong city. The food wastes are the main components of them. The characteristics of MSW of Chittagong city is presented in Table IV [14-16].

"Figure 2" shows the pie chart of typical MSW characteristics of Chittagong city based on the values of wet weight function from Table IV.

MSW generation can be calculated for the perspective city using equation (1) as Historical Trend [17]:

$$PWG = \frac{(PBY + PBY * AGP) * (PCWB + PCWB * AGW) * 365 * 1}{1000} \ldots\ldots (1)$$

Here, PWG is represented as projected wastes generation in a year (tons). Also, PBY, AGP, PCWB, and AGW represent the population in the baseline year, an annual growth rate of population, per capita wastes generation in the baseline year (kg/cap/day) and average growth rate in the per capita waste generation respectively.

TABLE IV. TYPICAL MSW CHARACTERISTICS IN CHITTAGONG CITY

Material	Wet weight fraction (%)	Moisture content (%)	Dry weight fraction (%)
Food wastes	70	74.86	17.60
Plastic	5	0.61	4.97
Paper	10	3.60	9.64
Grass & straw	2	34.18	1.32
Glass & ceramic	1	0	1.00
Metals	4	0	4.00
Textiles	5	7.3	4.64
Others	3	7.24	2.78

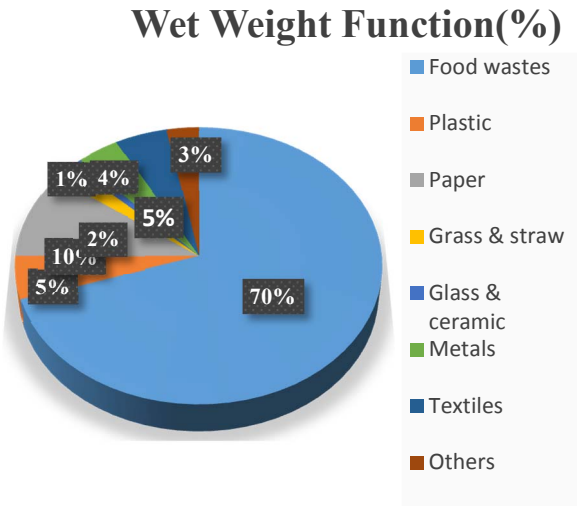


Fig. 2. Amount of different types MSW production of Chittagong city.

Table V shows the differences among various approaches of WTE technologies applied in Chittagong city considering the Case study information, Total cost, and MSW conversion and emission factor. In terms of case study information, we consider waste feed 958 tons/day here in every technology from produced 1550 tons/day according to the information from Satyajit Roy Das et.al [12]. Besides this average distance from transfer point to hub consider here 7 km and also the data's related to vehicles which are used to convey the wastes are also taken from that analysis. In terms of cost parameter, we consider three types of cost here including Average plant Capacity cost (USD/ton waste), Operation and Maintenance Cost (USD/ton waste), Transportation cost (USD/km). Having said that we just have to contemplate the transportation cost in landfilling technologies, the existing method. The major feature we analyzed in Table V is the production of electricity and heat from per ton MSW for different technologies in which we can perceive that there is no electricity or heat produced from the current landfilling process.

IV. RESULTS AND DISCUSSION

By analyzing different technologies including the case study information, Total cost, MSW conversion and emission factor from Table V we determine the results for each scenario which shown in Table VI based on the study done by Satyajit Roy Das et.al [12]. They investigated only the overall data of waste production and transportation costs from each site and finally collected waste disposals to the land. But they didn't propose any effective solutions to the collected waste after dumping them to the land or any other better clarifications than landfilling. Our offerings here to provide an effective method for the management of solid waste in Chittagong city and also by producing electricity, heat and other effective matters from that disposed waste based on Energy, Economy and Environmental strictures analyzed through various emerging technologies and to mark waste as a source of energy. We also analyzed the cost from the waste collection stage to the electricity production stage.

The Advantages and disadvantages of Incineration, Landfilling, Gasification, and pyrolysis are also considered here for the analysis of energy, cost and environmental impact [18-20]. According to the information of Worcester polytechnic institute, we used here, from 1-ton waste Anaerobic Digestion method produces 1.2 m3/ton MSW [21] and for per ton CO2 production in pyrolysis methods are 1682.6 (tCO2/d) [20].

This study emphases on the management of solid waste in Chittagong city and its proper utilization by analysis the comparative studies based on the amount of Energy production, Cost perspective, and Environmental impact. We can easily understand the necessity of changing the existing waste management method known as "Landfilling" method as because it results in a huge negative cost impact of 76,140.4 USD/d including the costs of transportation, tipping fee, operation & maintenance, and also emitted near about 1000 t CO2/d. To overcome this problem we can use WTE method which is much profitable and environmentally friendly too. Based on the five consequences of WTE method applicable in Chittagong in Table-V we could prefer incineration technology in terms of energy generation which includes the production of electricity and heat. Incineration technology is capable to produce a total of 560.43 MWh/d of electricity and 1668.84 GJ/d of heat which trailed by Anaerobic Digestion is shown in "Figure 3". Though the Gasification and Pyrolysis method of technologies provides a little more power than the Incineration technology they don't produce any heat from their process. In terms of cost analysis, the favorable method is Anaerobic Digestion technology because of its lower operating and maintenance cost and the total cost is 84,369.41 USD/d. The second preferred method is incineration technology which slightly higher than anaerobic digestion technology and is near about 94,925.73 USD/d. But the other methods analyzed here is much higher than these two methods exposed in "Figure 4".

However, in considerations of environmental issues we should prefer the Gasification technology due to its higher amount of carbon mitigation and produce only 219.6 tCo2/d and it follows the Anaerobic Digestion, Incineration, and Pyrolysis which is shown in "Figure 5". Whereas Anaerobic Digestion and Incineration technology produces a marginally higher amount of CO2 than Gasification technology.

TABLE V. PARAMETERS FOR THE VARIOUS TECHNOLOGIES CONSIDERED FOR SOLID WASTE MANAGEMENT IN CHITTAGONG.

Parameter	Method				
	Landfilling	**Incineration**	**Gasification**	**Anaerobic Digestion**	**Pyrolysis**
Case study information					
Waste feed (t/d)	958	958	958	958	958
Average distance from transfer point to hub (Km)	7	7	7	7	7
Average vehicle Capacity (t/truck)	3.82				
Number of trips by vehicles	376				
Costs					
Average plant Capacity cost (USD/t waste)	14	2.1074	1.522	6.088	8.595
Operation and Maintenance Cost (USD/t waste)	10	95	115	80	115
Transportation cost (USD/km)	3.45				
MSW conversion factor					
Electricity production (MWh/ t MSW or MWh/m^3)		0.585	0.66	0.25	0.66
Heat production (MWh/ t MSW or MWh/m^3)		1.742		0.2976	
Ash production (t/t MSW)		0.1			
Biogas Production (m^3/t MSW)				1.2	
Digestate production (t/t MSW)				0.3	
Emission factor					
CO_2 emission from processing (t CO_2/t MSW)	0.791	0.28	0.2	0.253	1.7
CO_2 emission from transportation (t CO_2/km)	0.014	0.014	0.014	0.014	0.014
CH_4 emission (t/t MSW)	0.023				

TABLE VI. RESULTS FOR EACH SCENARIO BASED ON ENERGY, COST AND ENVIRONMENTAL EMISSION

Parameter	Landfilling	Incineration	Gasification	Anaerobic Digestion	Pyrolysis
Energy					
Electricity (Mwh/d)		560.43	632.28	287.4	632.28
Heat (Mwh/d)		1668.84		342.12	
Biogas (m^3)				1149.6	
Costing (USD/d)					
Capital cost	13,412	2,018.89	1,458.12	5,832.57	8,234.28
Processing (O & M) cost	9,580	91,010	110,170	76,640	110,170
Transportation cost	16,560	1,896.84	1,896.84	1,896.84	1,896.84
Total cost (USD)	39,552	94,925.73	113,524.96	84,369.41	120,274.12
Environment Emission (tCO_2/d)					
Processing Emission	757.778	268.24	191.6	242.38	1628.6
CH_4	22.034				
Transmission Emission	342	62	28	28	28
Total	1121.812	330.24	219.6	270.38	1656.6

But pyrolysis technology produces a much higher amount of CO_2 and is near about 1656.6 t CO_2/day. Analyzing the above discussion we are proposing Incineration Technology is the best solution for MSW management in Chittagong city based on the three major disputes i.e. Energy, Cost and Environment.

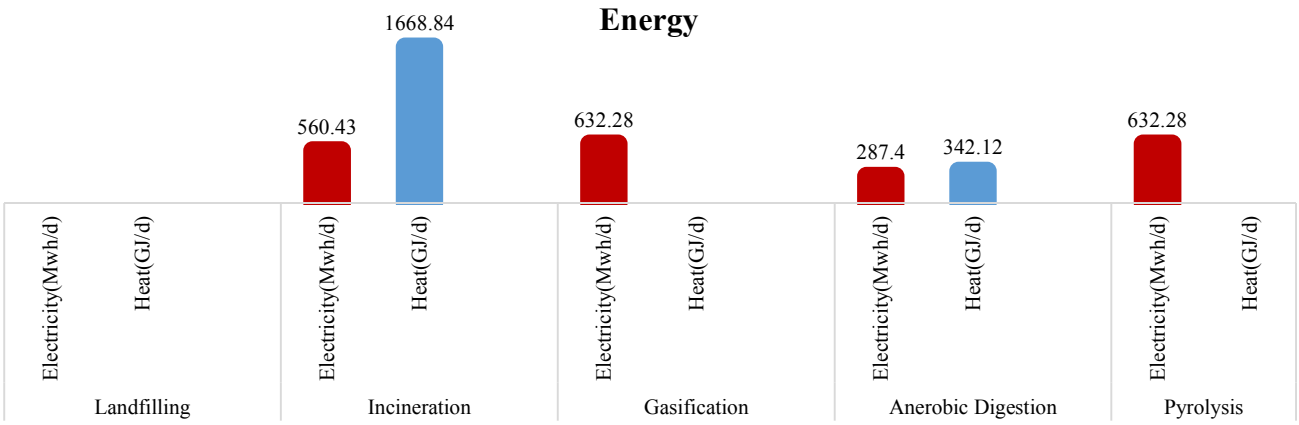


Fig.3. Energy production analysis in n different scenarios of waste management in Chittagong City.

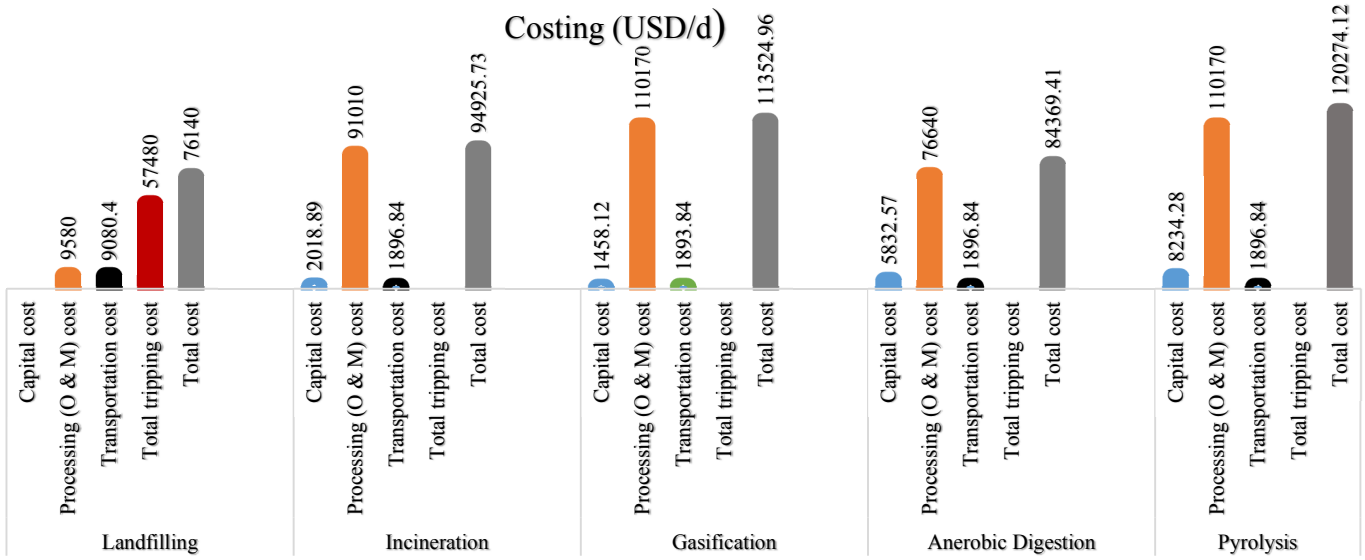


Fig.4. Cost analysis in different scenarios of waste management in Chittagong City.

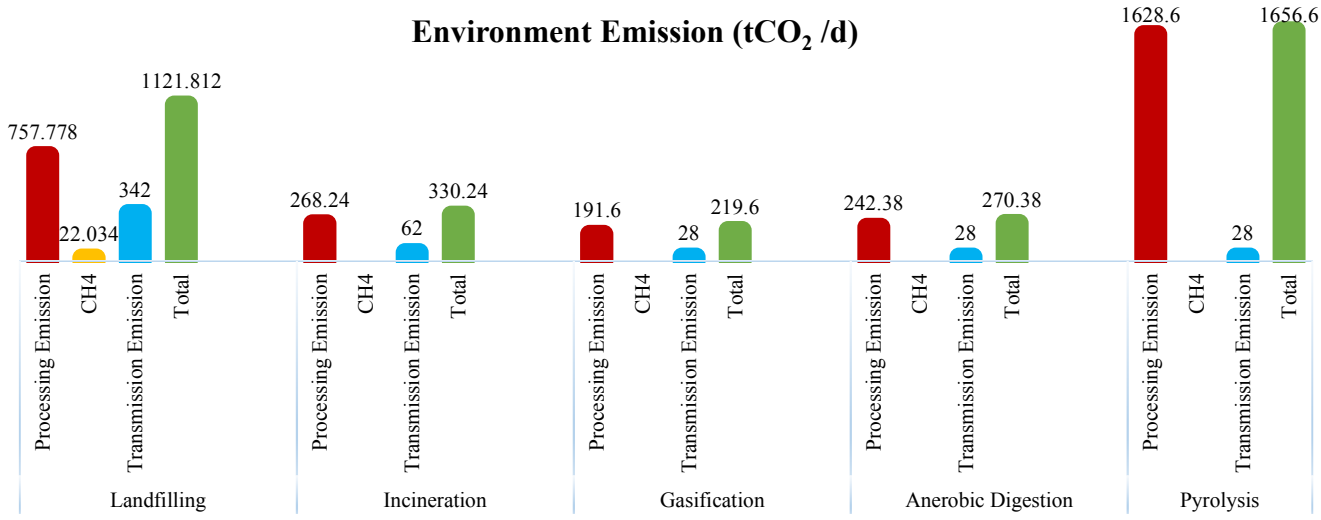


Fig.5. Environment Emission analysis in different scenarios of waste management in Chittagong City.

V. CONCLUSIONS

The municipal waste management system is always a great concern for the authority of any city and also it is much needed for them to clean it properly. In the prospect of Bangladeshi cities, the proper waste management system is still a great challenge. Besides this, the shortage of electricity in our country is still at a very high rate. According to our proposed method, the MSW can easily be managed and also the wastes can be used as a source of energy. According to the overall analysis, we can say that among all of the WTE technologies in perspective of Bangladesh Incineration is the most cost-effective, efficient, and environment-friendly method of waste-to-energy (WTE) technology. Moreover, for our future study, we are thinking to utilize the renewable sources with WTE technology to make a smarter micro grid which can be a building block for a green-clean smart city. We hope that adopting this method will help us to produce electricity using waste for our growing power demand as well as to fulfill our vision of making a green and smart city.

REFERENCES

[1] World Energy Outlook, Online, Available: https://www.iea.org/newsroom/news/2010/november/world-energy-outlook-2010.html

[2] Hoornweg, Dan & Perinaz, B.T.. (2012). What a waste: a global review of solid waste management. Urban Dev Ser Knowl Pap. 15. 87-88.

[3] Bangladesh-Access to electricity (% of population), Online, Available: https://tradingeconomics.com/bangladesh/access-to-electricity-percent-of-population-wb-data.html

[4] H.K. Jeswani, A. Azapagic, Assessing the environmental sustainability of energy recovery from municipal solid waste in the UK, Waste Management, Volume 50, 2016, Pages 346-363, ISSN 0956-053X, https://doi.org/10.1016/j.wasman.2016.02.010.

[5] Marisabel Cuberos Balda, Takaaki Furubayashi, Toshihiko Nakata,Integration of WTE technologies into the electrical system for low-carbon growth in Venezuela,Renewable Energy,Volume 86, 2016, Pages 1247-1255, ISSN 0960-1481, https://doi.org/10.1016/j.renene.2015.09.052.

[6] Angelidaki, Irini & Batstone, Damien. (2010). Solid Waste Technology & Management, Volume 1 & 2. 10.1002/9780470666883.ch37.

[7] Technology Options for Energy Recovery from Waste, Online, Available: http://www.eai.in/ref/ae/wte/pro/process.html

[8] Münster, M (2009) Energy Systems Analysis of Waste to Energy Technologies by use of Energy PLAN, Danish Technical Research Counsel; ISBN 978-87-550-3719-9

[9] Zaman, Atiq. (2010). Comparative Study of Municipal Solid Waste Treatment Technologies Using Life Cycle Assessment Method. International Journal of Environmental Science and Technology. 7. 10.1007/BF03326132.

[10] Analyzing Different Waste-to-Energy Technologies, Online, Available:https://bioenergyconsult.wordpress.com/2011/09/24/comparison-of-different-waste-to-energy-processes/

[11] Media Brief of CCC, Online, Available: http://www.ccc.org.bd/media_brief

[12] Das, Satyajit & Hossain, Md & Talukder, Snigdha & Hossain, Mohammed. (2015). Solid Waste Management in Commercial City of Bangladesh. 3. 60-70. M. Alamgir and A. Ahsan, "Characterization of MSW and nutrient contents of organic component in Bangladesh," Electronic Journal of Environmental, Agricultural and Food Chemistry, vol. 6, no. 4, pp. 1945–1956, 2007.

[13] Load-shedding back in town, Online, Available: http://m.theindependentbd.com/post/95770

[14] Alamgir, Muhammed & Ahsan, Amimul. (2007). Characterization of MSW and nutrient contents of organic component in Bangladesh. 6.

[15] Alamgir, Muhammed & Ahsan, Amimul. (2007). Municipal solid waste and recovery potential: Bangladesh perspective. Iranian Journal of Environmental Health, Science and Engineering (ISSN: p-ISSN: 1735-1979) Vol 4 Num 2. 4.

[16] M. M. Uddin and M. S. S. Mojumder, "Biogas production from municipal waste: prospect in Bangladesh," Journal of Selected Areas in Renewable and Sustainable Energy, pp. 1–5, 2011.

[17] BBS, Statistical Pocketbook of Bangladesh-2013, Bangladesh Bureau of Statistics (BBS), Statistics and Informatics Division (SID), Ministry of Planning, Government of the Peoples's Republic of Bangladesh, Dhaka, Bangladesh, 2014

[18] What in Incineration? Online, Available: https://www.conserve-energy-future.com/advantages-and-disadvantages-incineration.php

[19] What are the advantages and disadvantages of using landfilling for trash disposal? Online, Available: https://www.quora.com/What-are-the-advantages-and-disadvantages-of-using-landfills-for-trash-disposal

[20] Parker, C., and Roberts, T.. *Energy from waste: An evaluation of conversion technologies*. United States: N. p., 1985. Web. Online, Available: https://www.osti.gov/biblio/7002418

[21] Methane Creation from Anaerobic Digestion, an Interactive Qualifying Project Report by WORCESTER POLYTECHNIC INSTITUTE, Online, Available: https://web.wpi.edu/Pubs/E-project/Available/E-project-042810-154622/unrestricted/DigestionIQP2

Modeling and Performance Analysis of an Γ-Model Based DFIG Wind Turbine with Duel Crowbar Protection

Mehedi Azad Shawon
Dept. of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka, Bangladesh
mehedi.shawon@aiub.edu

S. M. Imrat Rahman
Dept. of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka, Bangladesh
imratrahman@aiub.edu

Naheed Hossain
Dept. of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka, Bangladesh
naheed@aiub.edu

***Abstract*—This paper presents a detailed mathematical model of a doubly fed induction generator (DFIG) based wind turbine system. The controllers associated with this DFIG model are also derived in synchronously rotating reference frame considering Γ-Model of induction generator. A PSCAD/EMTDC simulation model has been implemented based on the proposed DFIG model. The effectiveness of this model is further verified through different case studies where the dynamic response of the implemented PSCAD model is compared with the theoretical model response. The simulation results show satisfactory performance of the proposed model and can be used for evaluation of transient stability of DFIG wind turbine system.**

Keywords—Doubly fed induction generator, grid side converter, rotor side converter, DC-link.

I. INTRODUCTION

The increasing demand of electricity is one of the greatest challenge that the modern world is facing. At the same time, the source of electrical energy has become an important concern in terms of environmental pollution and global warming which are mainly caused by burning fossil fuel in power generation. Thereby the renewable energy sources are taking over the place of conventional fossil fuel based energy generation system for obvious reasons. Among all renewable energy sources, the wind energy has become one of the prominent and dominant technology since last decade [1]. Today variable speed wind turbines equipped with doubly fed induction generator (DFIG) are widely used in wind turbine industries since energy efficiency of a DFIG based wind turbine (WT) system is a few percentage units higher than the wind turbine system controlled by full power converter [2]. In addition, the output power variations generated by wind gust can be significantly reduced in DFIG WT system compared to fixed speed wind turbine system [3]. With the large penetration of DFIG WT system, the study of reliability, ease of controllability, fault ride through capability of WT have become an important area of study. For many years the DFIG WT systems have been investigated for better performance [4][5]. In most of these cases, the simplified converters are used as controllable current sources or voltage sources which make it impossible to investigate the performance of the converters accurately. In this paper a DFIG based WT model is developed mathematically with decoupled control law both for the grid side converter (GSC) and rotor side converter (RSC). A duel overcurrent protection is implemented for the rotor windings and DC-link for damping of excess energy from the rotor shaft during grid disturbances [6]. The proposed DFIG WT model has been implemented in PSCAD/EMTDC simulation tool. The validation of investigated model is further done by the simulation results of different case studies.

II. MODELING OF DFIG BASED WIND TURBINE SYSTEM

A. Components of DFIG Wind Turbine

In a DFIG based wind turbine, the turbine shaft is connected with the generator shaft through a gearbox. The stator windings are connected to the grid via coupling transformer, while the rotor windings are connected with the grid through a back-to-back voltage source converter (VSC) system as depicted in Fig. 1. The DFIG feeds power to the grid from both its stator (approx. 2/3 of P_n) and rotor (approx. 1/3 of P_n) windings.

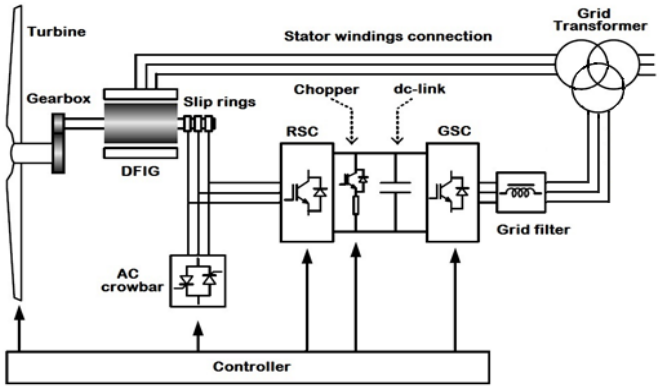


Fig. 1. Components of a DFIG WT system

The purpose of the RSC is to control the rotor currents in such a way that; the desired torque is developed at the shaft of the DFIG. The GSC regulates the dc-link voltage across the DC-link capacitor by exchanging active power with the grid. It can also control reactive power exchanged between the GSC and the grid. Here, both the VSCs operate in 4 quadrants for bidirectional power flow between rotor windings and the grid network [7].

B. Mathematical Model of DFIG

The DFIG model presented in this paper is based on the inverse gamma model (Γ-model). The stator dynamics of the DFIG is ignored in this Γ-model representation. The rotor dynamics of a DFIG are predominantly controlled by the RSC & GSC for stable operation. Therefore, for transient stability analysis, the Γ-model representation of induction generator as shown in Fig. 2 is considerably reasonable choice for mathematical modeling of DFIG WT system [3].

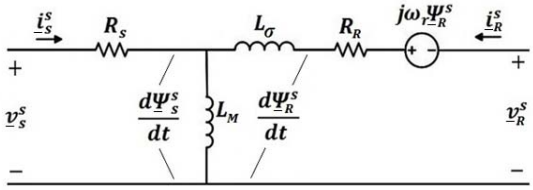


Fig. 2. Γ-model representation of IG in stator coordinates

This model is described by the following differential equations in synchronous coordinates with the help Clark's and Perk's transformation.

$$\underline{v_s} = R_s\underline{i_s} + \frac{d\underline{\Psi_s}}{dt} + j\omega_k\underline{\Psi_s} \tag{1}$$

$$\underline{v_R} = R_R\underline{i_R} + \frac{d\underline{\Psi_R}}{dt} + j\omega_{slip}\underline{\Psi_R} \tag{2}$$

Where v, i and Ψ denote the voltage, current and flux quantity respectively and ω_r is the frequency of the rotor. The subscripts s and R refer to the quantities of the stator and rotor respectively. Moreover, the superscript s is used to identify space vector that refer to the stator coordinate system. Finally, the vectors with no superscript represent quantities refer to synchronous coordinate system. The term ω_k and ω_{slip} represent the synchronous and slip frequencies respectively. The stator flux, rotor flux and electromechanical torque are given by (3), (4) & (5) respectively.

$$\Psi_s = L_M(i_s + i_R) \tag{3}$$

$$\Psi_R = (L_M + L_\sigma)i_R + L_M i_s = \Psi_s + L_\sigma i_R \tag{4}$$

$$T_e = \frac{3}{2}n_p \mathrm{Im}[\Psi_s {i_R}^*] \tag{5}$$

$$\frac{J}{n_p}\frac{d\omega_r}{dt} = T_e - T_m \tag{6}$$

Where L_M is the magnetizing reactance, L_σ is the leakage inductance. Finally, the mechanical dynamics of the induction generator can be described by (6) where, J is the inertia of single mass WT model and T_m is the mechanical torque.

The conversion of different parameters of DFIG from the classical T-model into Γ-model is done as per the following relations as given in [2] where, the "r" in subscript represents the rotor quantities in classical T-model.

$$v_R = \gamma v_r \qquad i_R = \frac{i_r}{\gamma} \qquad \Psi_R = \gamma\Psi_r \qquad \gamma = \frac{L_{s\lambda} + L_m}{L_m}$$

$$R_R = \gamma^2 R_r \qquad L_\sigma = \gamma L_{s\lambda} + \gamma^2 L_{r\lambda} \qquad L_M = \gamma L_m$$

C. Series Filter Modeling for GSC

The GSC is connected to the grid through AC filter in order to reduce current transient. A *L*-filter or *LCL*-filter is mainly used as grid filter for voltage source converter [2]. For simplicity, *L*-filter is used in this DFIG WT model. The series filter impedance ($Z_f = R_f + jX_f$) has been chosen as 10% of base impedance of GSC. Filter losses has been evaluated as 10% of the filter reactance. By evaluating with numerical values, the GSC series filter is found as follows:

$$L_f = 0.0869\ mH \qquad R_f = 3.279 m\Omega$$

III. Control of DFIG Based Wind Turbine System

The control algorithms of the RSC are derived in the stator flux oriented d-q coordinate system. For coordinate transformation of RSC quantities, the phase angle of the stator flux is estimated in an open loop manner by measuring grid voltage and the stator current instantaneously.

$$\underline{\Psi_s^s} = \int (\underline{v_s^s} - R_s\underline{i_s^s})dt = |\Psi_s^s| \angle\mu \tag{7}$$

The output of an integrator (as estimated stator flux) is further passed through a band pass band filter with low cut-off frequency for avoiding the DC offsets, as suggested in [8].

To perform the coordinate transformation, the phase angle of the voltage vector must be known which can be estimated through a PLL in practical applications [9]. Since the GSC control system is aligned with the grid voltage vector, here the PLL is tracking the grid voltage E_g applied to the GSC in the secondary of the coupling transformer yielding $E_g = E_{gd}$ and $E_{gq} = 0$ in steady state condition. The law governing the PLL is considered as the one proposed in [10] where the controller gains are defined as

Proportional gain, $K_{p,PLL} = 2\,\alpha_{PLL}$

Integrator gain, $K_{i,PLL} = {\alpha_{PLL}}^2$

The bandwidth of the PLL for GSC is selected to be low as $\alpha_{PLL} = 34\ rad/s$. Thus the higher order transient of the grid voltage is rejected when the PLL operates with a reduced bandwidth [11]. Also the stability of the WT system improves with low bandwidth of PLL during grid disturbances [12].

A. Control of Rotor Side Controller

To utilize the advantage of variable speed operation of a DFIG based wind turbine, the tracking of optimum torque-speed curve is essential [13]. The speed of the rotor can be adjusted to a desired value by controlling the torque. The optimum power point is tracked by following a predefined torque-speed characteristic of the wind turbine. The turbine output power reference is varied according to the measured

rotor speed. In this DFIG WT model, the following relation between optimum power (P_{opt}) and turbine speed (ω_t) is used.

$$P_{opt} = K_{opt}{\omega_t}^3$$

Where, $K_{opt} = \frac{1}{2}\rho A C_{p_{opt}} \left(\frac{R}{\lambda_{\text{opt}}}\right)^3$ is constant. Also the terms $\rho, A, C_{p_{opt}}$ and λ_{opt} represent air density, swept area, power coefficient and optimum tip speed ratio of the turbine respectively.

Current Control

In order to derive the rotor current control law, (2) can be re-written in terms of state variable i_R and the input signal v_R by substituting the value of (4) into (2) as follows:

$$\underline{v_R} = (R_R + R_s + j\omega_{slip}L_\sigma)\,\underline{i_R} + L_\sigma \frac{d\underline{i_R}}{dt} + \underline{E} \tag{8}$$

Here, E is the back EMF and the term $j\omega_{slip}L_\sigma \underline{i_R}$ is the cross-coupling term. Assuming perfect parameter estimation and cancellation of back EMF the rotor current dynamics formed by introducing active damping (R_a) and cross-coupling term as follows:

$$\underline{v_R}' = (R_R + R_s + R_a)\underline{i_R} + L_\sigma \frac{d\underline{i_R}}{dt} \tag{9}$$

Therefore, applying internal model control (IMC) method [11], (9) yields a PI controller of gain parameters as follows:

Proportional gain, $K_{p,P} = \alpha_{rsc} L_\sigma$

Integrator gain, $K_{i,P} = \alpha_{rsc}(R_R + R_s + R_a)$

The bandwidth of the current controller is selected to be $\alpha_{rsc} = 3000\, rad/s$. Finally, the rotor current law can be rewritten as below and the controller diagram is shown in Fig. 3.

$$\underline{v_R} = \left(K_{p,P} + \frac{K_{i,P}}{s}\right)(\underline{i_R}^{ref} - \underline{i_R}) + \left(j\omega_{slip}L_\sigma - R_a\right)\underline{i_R} + \underline{E} \tag{10}$$

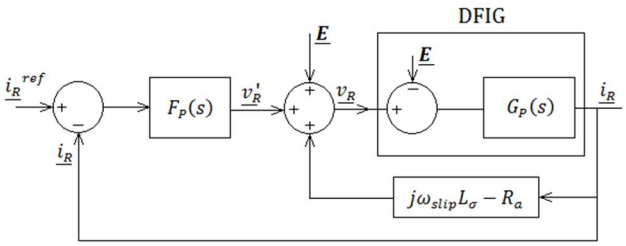


Fig. 3. Diagram of RSC current controller.

Power and Torque Control

In a stator flux coordinated system, $v_{sd} = 0$ assuming the stator resistance is negligible. The active and reactive power of the stator of a DFIG can be expressed in terms of rotor current quantities of synchronously rotating reference frame as follow:

$$P_s = -\frac{3}{2}(v_{sq} i_{Rq}) \tag{11}$$

$$Q_s = \frac{3}{2} v_{sq}(i_M - i_{Rd}) \tag{12}$$

It can be seen that, P_s and Q_s can be controlled independently by regulating the rotor current components, i_{Rq} and i_{Rd} respectively. This model control the torque in an open loop manner by following the electro-dynamical torque T_e expression as follows:

$$T_e = \frac{3}{2} n_p \text{Im}[\underline{\Psi_s}\,\underline{i_R}^*] \approx -3\underline{\Psi_s} i_{Rq} \tag{13}$$

The pole-pair of this DFIG machine under consideration in the PSCAD simulation model is $n_p = 2$ [14].

Speed Control

Based on the mechanical dynamics of the induction machine, a proportional controller is used where the gain parameter is as follows:

Proportional gain, $K_{p,S} = \frac{J\alpha_S}{n_p}$

$$T_e^{ref} = K_{P,S}(\omega_r^{ref} - \omega_r) \tag{14}$$

In this model, the bandwidth of the speed controller is selected as, $\alpha_S = 150\, rad/s$.

B. Control of Grid Side Controller

The GSC control scheme consists of two cascaded control loops. The inner current control loop regulates independently the d-axis and q-axis of the filter current, i_{fd} and i_{fq} respectively. The outer loop regulates the dc-link voltage and the reactive power exchange between the grid and GSC.

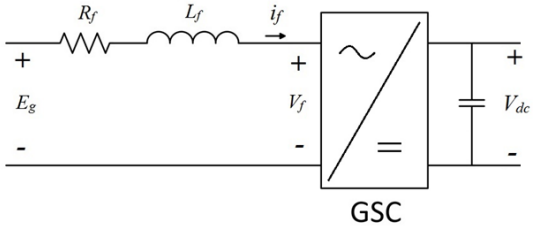


Fig. 4. GSC with L-filter

Fig. 4 shows the simplified model of GSC. The dynamics of the grid filter in synchronously rotating reference frame can be expressed as (15) where ω_g is angular frequency of the grid voltage and $L_f\omega_g \underline{i_f}$ is the cross coupling term.

$$\underline{v_f} = -(R_f \underline{i_f} + L_f \frac{d\underline{i_f}}{dt}) - jL_f\omega_g \underline{i_f} + \underline{E_g} \tag{15}$$

Current Controller

Applying IMC method on (15) yields a PI controller as follows:

$$\underline{v_f} = -\left(K_{f,P} + \frac{K_{f,P}}{s}\right)(\underline{i_f}^{ref} - \underline{i_f}) - j\omega_g L_f \underline{i_f} + \underline{E_g} \quad (16)$$

Where, the GSC current controller gains are defined as

Proportional gain, $K_{p,g} = \alpha_{gsc} L_f$

Integral gain, $K_{i,g} = \alpha_{gsc} R_f$

Bandwidth of GSC current controller,

$$\alpha_{gsc} = 3000\ rad/s$$

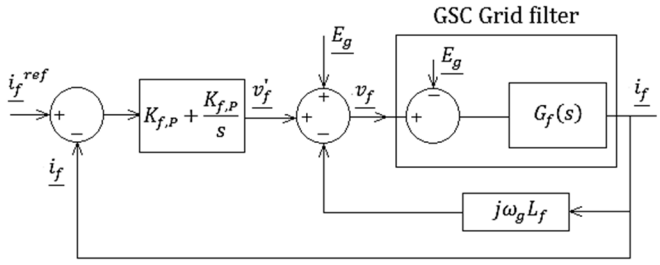


Fig. 5. Diagram of GSC filter current controller.

DC-link Controller

In the simplified DC-link model as shown in Fig. 6, the amount of the power that flows through the DC-link capacitor depends on the power delivered to the DC-link from GSC, P_{gsc} and the power delivered to the rotor of the DFIG, P_r.

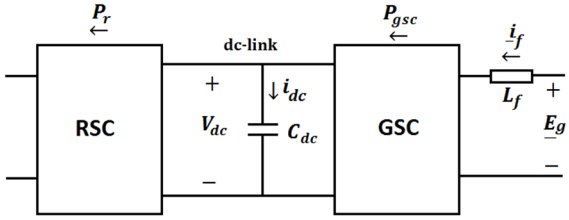


Fig. 6. Simplified DC-link model

Under the assumption that, the GSC is lossless and the DC-link capacitor is constant in time, the power flowing in the DC-link P_{dc} must be equal to the energy stored in the DC-link capacitor, yielding

$$P_{gsc}(t) - P_r(t) = P_{dc}(t) = \frac{dE_{dc}(t)}{dt}$$

$$= \frac{1}{2} C_{dc} \frac{d(V_{dc})^2}{dt} \quad (17)$$

To derive the control law, (17) can be linearized as suggested in [15]

$$\frac{1}{2} C_{dc} V_{dc,0} \frac{d\Delta V_{dc}}{dt} = P_{dc} = \Delta P_{gsc} - \Delta P_r \quad (18)$$

Observed that, the closed loop dynamics depends on the operating point of the dc-link voltage, $V_{dc,0}$. This inconvenience can be avoided by using the square of the dc-link voltage $(V_{dc})^2 = W$ in the control law as suggested in [15]. However, applying IMC the (17) yields a proportional controller with gain $K_{p,dc} = \frac{\alpha_{dc} C_{dc}}{2}$. Therefore, the complete control law is given as

$$P_{gsc}^{ref} = \left(K_{p,dc} + \frac{K_{i,dc}}{s}\right)(W^{ref} - W) + P_r \quad (19)$$

Observed from (19) that, the active power delivered to the rotor (P_r) is included in control law as a feed forward term. Finally, using amplitude invariant transformation, the d-axis grid filter current reference i_{fd}^{ref} can be derived as

$$i_{fd}^{ref} = \frac{2}{3} \frac{P_{gsc}^{ref}}{E_{gd}} \quad (20)$$

The bandwidth of the outer controller should be 5 to 10 times smaller than the inner controller [4]. Here, the bandwidth of the DC-link controller is selected to be $\alpha_{dc} = 300\ rad/s$.

IV. DFIG WT Protection System

In spite of having a number of advantages, a major drawback of DFIG based wind turbine system is their lower LVRT capability. Any voltage dip at grid induces a serious over-voltage as well as over-current in the rotor windings. Moreover, a major part of this over-current will flow through the dc-link capacitor which results an over-voltage in the DC-link [16].

In this DFIG model, an AC crowbar system is directly connected to the rotor windings to prevent the rotor transient over-current to pass through the rotor side converter valves. The selection of the crowbar resistance is made according to (21) as suggested in [16].

$$R_{AC-cb} < \frac{V_{dc,nominal}}{\sqrt{6} I_{rsc,max}} \quad (21)$$

Where, $V_{dc,nominal}$ is the nominal dc-link voltage and $I_{\text{rsc,max}}$ is the maximum allowable rotor current. Normally, the IGBTs in the converter can handle twice as much as the nominal current of the converter for 1 ms. For this reason, the typical set point for the maximum allowable rotor current is considered as 1.8 times the nominal rotor current as suggested in [15]. Therefore, the AC crowbar resistance can be found as

$$R_{AC-cb} < \frac{0.94\ kV}{\sqrt{6}\ (1.8 * 0.84\ kA)} < 0.254\ \Omega$$

In this model the AC crowbar resistance is set to be $R_{AC-cb} = 0.25\ \Omega$.

For the dc-link over-voltage protection, an IGBT controlled DC crowbar is placed across the dc-link buses. When the DC-link bus voltage reaches the maximum allowable voltage set point, the IGBT will act and thereby, the excess energy of the DC-link capacitor will be discharged through the DC crowbar resistance, $R_{dc-link}$ until the DC-link voltage become lower than the maximum set point. In this DFIG based WT model, the maximum dc-link voltage is set at 1.05 pu and the dc-link chopper resistance is chosen as $R_{dc-link} = 1\Omega$. This DC crowbar system will improve the voltage ride through capability of the DFIG WT system which compromises its high installation cost [17].

V. Simulation Results

A. Performance of DC-link Controller

One of the important analysis on the stability of a DFIG WT control system is the performance of the DC-link controller. The DC-link voltage must be kept constant within a predefined limit to prevent the rotor shaft from over-speed caused by undamped excess energy in case of any grid disturbance. Thus the dynamics of the DC-link controller must be fast enough to ensure that the actual DC-link voltage follows the reference in all operating conditions. The performance of the DC-link controller used in this model is analyzed by two different step response of the DC-link voltage reference (a negative step at 3s and a positive step at 3.7s) as shown in Fig. 7.

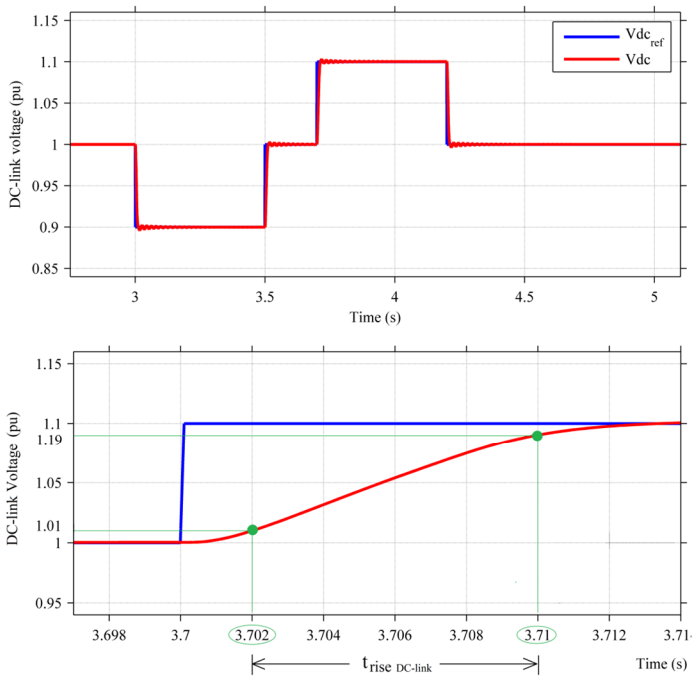


Fig. 7. (a) Performance of the dc-link voltage controller against step variations of the DC-link voltage reference; (b) Rise time of DC-link voltage.

It can be observed from the Fig. 7(a) that, there are small oscillations in the DC-link voltage upon step changes. Although the DC-link model is a first order system, the addition of the small integrator gain in the controller makes the close loop system as a second order system. Moreover, there is a small time delay (1 ms) between the measured GSC power and the GSC power reference ($P_{gsc}{}^{ref}$) imposed by the controller. This is an important factor that causes a very small overshoot and oscillation in the DC-link voltage. However, the DC-link controller is able to damp the oscillation within 150-200 ms, so that, the dynamic performance of the DC-link controller can be considered suitable for the investigated DFIG WT model. Moreover, Fig. 8(b) shows that, the rise time of the actual DC-link voltage is 8 ms; where, the theoretical rise time of the modeled DC-link voltage controller is 7.3 ms $\left(\frac{\ln 9}{\alpha_{dc}}\right)$ according to the selected DC-link controller bandwidth.

B. Verification of DFIG WT Model

Case Study 1

Depending on the mode of operation (i.e. Sub-synchronous or Super-synchronous), the power flow direction between the rotor windings and the grid changes as shown in Fig. 8. Neglecting the losses in the generator and power converter, the relation between the stator power (P_s) and rotor power (P_r) can be seen by [18]

$$P_r = -sP_s \tag{22}$$

where, s is the slip of the DFIG defined as $s = \frac{\omega_k - \omega_r}{\omega_k}$

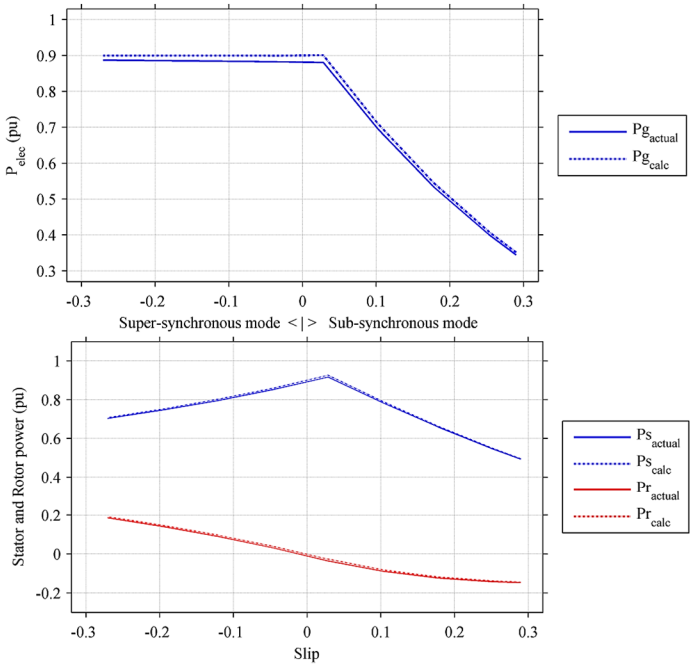


Fig. 8. Steady state power flow of DFIG at different slip: (a) total generated power; (b) stator and rotor power.

Fig. 8 shows the steady state power flow performance with respect to different slip conditions where the incident wind velocities are taken as constant. It can be seen from the plot (a) that, the actual generated power (Pg_{actual}) from the wind turbine at different slip is almost identical to the calculated power (Pg_{calc}). Similarly in plot (b), the actual stator and the rotor power follows the calculated stator and rotor power respectively at different operating conditions. Moreover, the polarity of rotor power changes exactly at slip zero, which further validates the implemented DFIG WT model.

Case Study 2

The dynamic performance of the rotor current of the DFIG model is illustrated in Fig. 9. The DFIG WT initially operates with a rotor speed at $\omega_r = 0.9\ pu$. From $t = 4\ sec$ the speed reference is gradually increased from 0.9 pu to 1.1 pu in 6 seconds which means, the operation mode of the DFIG switches from sub-synchronous mode into super-synchronous mode. During this ramp up time, the wind speed is assumed to be constant.

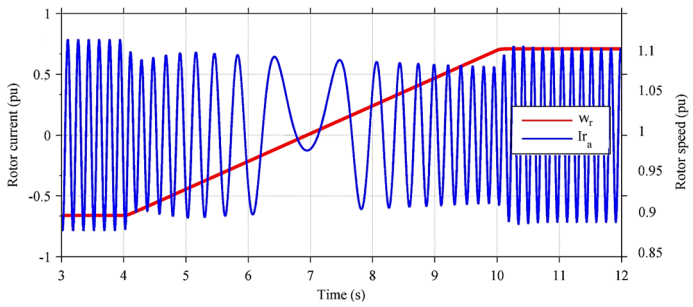


Fig. 9. DFIG rotor acceleration through synchronous speed.

A smooth transition of the rotor current is observed at $t = 6.9\ sec$ from sub-synchronous to super-synchronous mode without being distorted by higher order harmonic which signifies the effectiveness of the RSC controller used in this DFIG WT model.

Case Study 3

In this case study, a symmetrical voltage dip is emulated during which the remaining voltage at the point of common coupling (PCC) is controlled at 0.2 pu. for a period of 200ms as depicted in Fig. 10(a). Before this event, the wind turbine is generating nominal active power of 0.9 pu with unity power factor and with nominal slip equal to 2.7%.

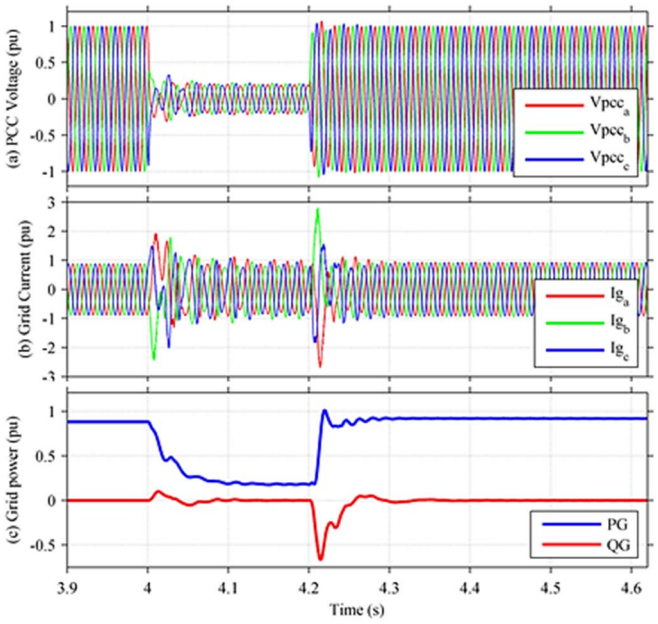


Fig. 10. Responses during symmetrical voltage dips: (a): three-phase terminal voltages; (b): three-phase current (I_g); (c): active and reactive power

As shown in Fig. 10(b), the maximum fault current from the wind turbine reaches 2.9 pu during the recovery of the grid voltage whereas a current peak of 2.4 pu is produced during the first 30ms of the voltage dip. Fig. 10(c) shows that, after the voltage recovery, the active power rises towards its pre-fault value due to the increasing voltage and current transient. As a result the DFIG WT reaches steady state within two electrical periods after voltage recovery. Finally, the reactive power is controlled in zero at steady state.

VI. Conclusion

This paper presents a complete model of DFIG WT with all control systems. The mathematical model is developed in PSCAD/EMTDC simulations. The dynamic response of simulated model with decoupled control of RSC and GSC are found consistent with the mathematical model through different case studies. This model can be used for investigating the transient stability and voltage ride through capability of DFIG based WT systems.

References

[1] L.N. Widanagama Arachchige, A.D. Rajapakse, D. Muthumuni, "Implementation, Comparison and Application of an Average Simulation Model of a Wind Turbine Driven Doubly Fed Induction Generator", *Energies*, vol. 10, issue 11, 2017

[2] Wei Qiao, "Dynamic Modeling and Control of Doubly Fed Induciton Generators Driven by Wind Turbines," in IEEE/PES Power System Conference and Exposition, PSCE '09, 2009.

[3] Massimo Bongiorno and Andreas Petersson, "Development of a Method for Evaluation of Wind Turbines Ability to Fulfil Swedish Grid Codes," ELFORSK, Stockholm, 2009.

[4] S. Chondrogiannis. "Technical aspects of offshore wind farms employing doubly-fed induction generators", PhD Thesis, University of Manchester, 2007.

[5] J. B. Ekanayake, L. Holdsworth, X. Wu, and N. Jenkins. "Dynamic modeling of doubly fed induction generator wind turbines", IEEE Transaction on Power Systems, Vol. 18, 2003

[6] T.Lei, M. Ozakturk. Modelling and analysis of DFIG wind turbine system in pscad/emtdc. 6th IET International Conference on Power Electronics, Machines and Drives (PEMD), pp. 1-6, 2012.

[7] Wei Qiao, "Dynamic Modeling and Control of Doubly Fed Induciton Generators Driven by Wind Turbines," in *IEEE/PES Power System Conference and Exposition, PSCE '09*, 2009.

[8] B. Hopfensperger, D.J. Atkinson, and R.A. Lakin, "Stator-flux-oriented control of a doubly-fed induction machine with and without position encoder," *IEE. Proc.-Electr. Power Appl.*, vol. 147, No. 4, July 2000.

[9] R Teodorescu, M Liserre, and P Rodriguez, "Grid synchronization in three phase power converters," in *Grid Converters Photovoltaic Wind Power Systems*, New York, 2011, pp. 169-204.

[10] Massimo Bongiorno and Torbjorn Thiringer, "A Generic DFIG Model for Voltage Dip Ride-Through Analysis," *IEEE Transactions On Energy Conversion*, May 2012

[11] Xiao-Qiang Guo, Wei-Yang WU, and He-Rong GU, "Phase Locked Loop and synchronization methods for grid interfaced converters: a review," in *Key Lab of Power Electronic for Energy Conservation and Motor Drive*, 2011.

[12] L. Ruikuo, Y. Jun, P. Jinxin, W. Xuewei, S. Peng, G. Xiaolong, "Dynamic Stability Analysis of the Weak Grid-Connected DFIG-based Wind Turbines under Severe Symmetrical Faults", International Conference on Power System Technology (POWERCON), ISBN: 978-1-5386-6461-2, Guangzhou, 2018.

[13] Badrul H. Chowdhury and Srinivas Chellapilla, "Doubly-fed induction generator control for variable speed wind power generation," *ELSEVIER*, pp. 786-800, December 2005.

[14] ABB, New standard slip ring generator series for the doubly fed 2.1 MW turbine, 2010, Technical Brochure.

[15] Abram Perdana, "Dynamic Model of Wind Turbines," Chalmers University of Technology, Goteborg, Sweden, Ph.D thesis 2008.

[16] SUN Li-ling, Yang Pu, and Wang Yi, "Simulation Research for LVRT of DFIG Based on Rotor Active Crowbar Protection," in International Conference on Sustainable Power Generation and Suppply (SUPERGEN, 2012), Hangzhou, 2012, pp. 1-7.

[17] R. Monteiro Pereira, A. Pereira, C. Ferreira, F. Barbosa, "Influence of Crowbar and Chopper Protection on DFIG during Low Voltage Ride Through", *Energies*, doi. 10.3390/en11040885, 2018.

[18] John Fletcher and Jin Yang, *Paths to Sustainable Energy*, 2010, ch. Introduction to Doubly-Fed Induction Generator for Wind Power Applications, pp. 258-278.

Electrical Properties of CSS Deposited CdTe Thin Films for Solar Cell Applications

N. K. Das
Dept. of EEE, Chittagong University of Engineering and Technology
Chittagong, Bangladesh,
nipudas@cuet.ac.bd

S. A. Razi
Dept. of EEE, Chittagong University of Engineering and Technology
Chittagong, Bangladesh,
razicuet@gmail.com

K. S. Rahman
Institute of Sustainable Energy
Universiti Tenaga Nasional
Kajang, Selangor, Malaysia
sajed_cuet_eee@yahoo.com

M. A. Matin
Dept. of EEE, Chittagong University of Engineering and Technology
Chittagong, Bangladesh
imamatin@yahoo.com

N. Amin
Institute of Sustainable Energy
Universiti Tenaga Nasional
Kajang, Selangor, Malaysia
nowshad@ukm.edu.my

***Abstract*— CdTe is a very potential absorber material for thin film solar cell application. In this work CdTe thin films (TF) were deposited on CdS thin films by close-spaced sublimation (CSS) technique at different source and substrate temperature in inert gas condition. To bring out the optimum temperature set for CSS deposited CdTe films, several experiments were done. The process pressure of the chamber during deposition was maintained at 1.5 Torr in a dynamic condition. The effects of deposition temperature on the electrical properties of the as-deposited CdTe films were investigated by Hall Effect measurement. The films were deposited at 610 ˚C, 630 ˚C, 650 ˚C and 670 ˚C source temperatures. The first three films showed p-type conductivity while n-type conductivity appeared in the film deposited at 670 ˚C. The hole concentration of the as-grown p-type CdTe films followed an upward trend with the increase of source temperature and it reached a peak value at 650 ˚C. The highest hole mobility was observed for the lowest source temperature. However, the resistivity of the CdTe films was found increasing with the increase of source temperature. Thus, the CdTe thin film deposited at 650 ˚C showed better electrical properties for solar cell applications.**

Keywords—CdTe thin film, Close-spaced sublimation, Electrical properties, Hall Effect measurement, Solar cells.

Introduction

CdTe is considered as a leading polycrystalline materials for the fabrication of highly efficient thin film (TF) hetero-junction solar photo-voltaic (PV) cells. Some unique physical properties make it an ideal p-type absorber material for solar PV cells. These properties include a direct band gap of about 1.45 eV which is optimum for the absorption of solar irradiation, an absorption coefficient which can be higher than 5×10^5 cm^{-1} at incident wavelengths of about 700 nm and the highest lattice parameter in the II-VI semiconductor group. Besides these, its formation energy is negative and the lowest among the members of this group. The CdS/CdTe based thin-film solar PV cells have already attained the efficiency over 22% in research level and over 18% in module level [1-2]. There are several techniques to deposit CdTe thin-films on CdS layer in which small-area laboratory solar cell demonstrate efficiency over 10% [3]. The most remarkable ones are CSS, electro-deposition, vacuum evaporation, sputtering, Vapor Transport Deposition (VTD), Metal Organic Chemical Vapor Deposition (MOCVD), Molecular Beam Epitaxy (MBE), and screen printing. All these techniques allow the deposition of high-quality thin films of CdTe which is polycrystalline in nature. However, the most efficient CdTe TF solar cells have been found with CdTe absorber layers deposited using CSS technique till to date [4]. Moreover, this technique is attractive among the researchers because of its high deposition rate, the simplicity of operation, and lower vacuum level during deposition and easiness of fabricating large volume commercial modules [5-6]. The most vital deposition parameters of CSS equipment are included; the temperature of source and substrate, the chamber pressure, the deposition time and the deposition atmosphere [7-11]. The photovoltaic conversion efficiency depends on the microstructure, optical and electrical properties of the as-grown CdTe thin film absorber layers which can be regulated by the deposition parameters. Researchers around the world are trying to find out the relationship between these properties and the deposition parameters [12]. The built-in bias voltage at equilibrium open circuit voltage V_{bi} is related to the carrier concentration of absorber and window material according to (1) [13].

$$V_{bi} = \frac{k\,T}{q} \ln\left(\frac{N_d N_a}{n_i^2}\right) \tag{1}$$

Where k is Boltzmann constant, N_a, N_d, and n_i are acceptor, donor and intrinsic carrier concentration per cm^3 of absorber and window layer, q is the charge of electron and T is the temperature. Larger built-in potential is desirable for separating light generated electron-hole pair in the absorber material. The short circuit current density Jsc is related to the parasitic series (R_s) and shunt resistance (R_{sh}) according to (2) [13].

$$J_{sc} = J_o\left[e^{\frac{(q(V-R_sJ_{sc})}{AkT}} - 1\right] - J_L + \frac{(V - R_sJ_{sc})}{R_{sh}} \tag{2}$$

Where A is the diode quality factor and J_o is the saturation current density. In this work, the CdTe thin films were deposited on 100 nm thick CdS thin films for four different source and substrate temperature. The temperature difference (ΔT) between source and substrate of the CSS equipment was kept at 30 ^{0}C. The source temperature was gradually increased from 610 ^{0}C to 670 ^{0}C in a steps of 20 ^{0}C. The dependence of several electrical properties such as carrier concentration, resistivity and carrier mobility of the as-deposited CdTe thin films were investigated by Hall Effect measuring system.

I. Experiment

The 3 cm × 3 cm borosilicate glass substrate was used to deposit CdTe TF by CSS techniques due to its better optical transmittance in the visible wavelength and higher strain point than soda lime glass [15]. The substrates were subject to cleaned by standard cleaning procedure. The substrates were scratched by the pen cutter in one side for marking and then scrubbed manually for 30 s each side with a nylon brush in soapy water. After that, the substrates were rinsed with warm deionized water and then cleaned ultrasonically by methanol, acetone, methanol and deionized water respectively at a temperature 50 ^{0}C for 15 minutes each. The clean substrates were dried with N_2 gas and then air anneal in a hot plate at 100 ^{0}C for 5 minutes. Finally, four cleaned substrates were mounted to the substrate holder in an RF sputtering machine. The distance between target and substrate was kept 7 cm. A 100 nm thick CdS film was deposited on each clean substrate by RF sputtering under 3.9 mTorr working pressure for 30 minutes using 99.99% pure CdS target of 2-inch diameter. During deposition Ar gas flow was maintained to 16 sccm through MFC controller. The RF power of the plasma gun was kept to 30 W while the substrate temperature was at 100 °C. The as-deposited CdS film shown in Fig. 1.

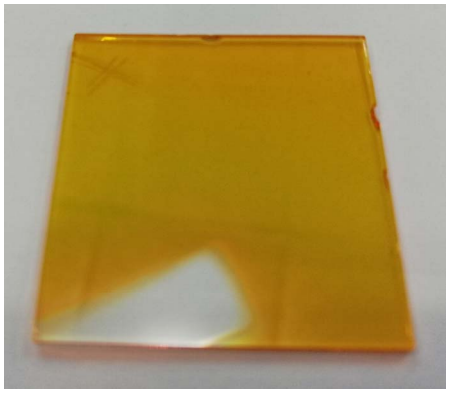


Fig. 1 RF sputtered as-deposited CdS films

The thickness of the CdS film was measured by the Quartz Crystal-deposition Monitor (QCM). The deposited CdS films were stored in a vacuum desiccator for avoiding the detrimental effects on the film-quality. The schematic diagram of traditional CSS equipment is shown in Fig. 2. The CSS equipment deposits CdTe film on one CdS film at a time. The 4N analytical grade CdTe powder was used as source material.

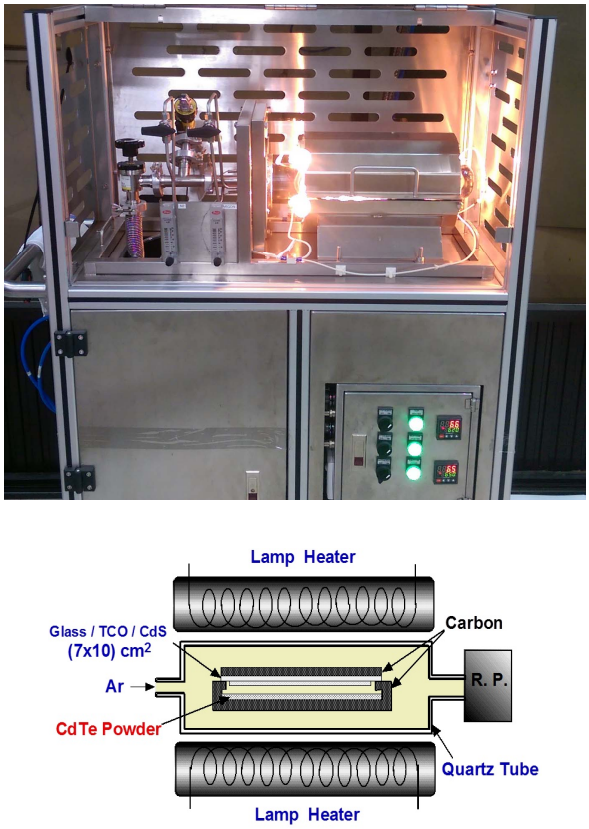


Fig. 2 CSS equipment with schematic diagram of chamber

The CdTe powder was poured into a graphite susceptor having 2.85 cm × 2.85 cm width, and 0.6 mm depth. The source powder was then compacted by a 2 cm× 5 cm glass slider. Finally, the CdTe powder was sintered to form a hard CdTe compound at 700 ^{0}C temperature for 30 minutes in Ar ambient. After that, the CdS film faced down towards the CdTe source as shown in Fig. 3.

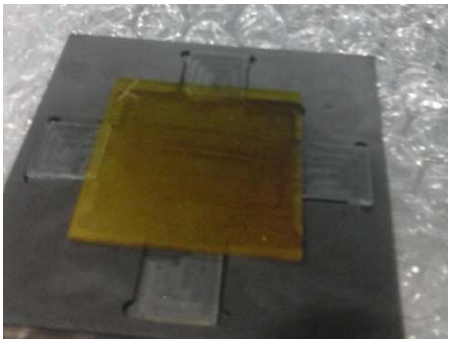


Fig. 3. CdS Films facing towards the CdTe Solid.

The important process parameters for depositing high-quality CdTe TF include the temperature of the source (T_{sou}) and substrate (T_{sub}), the spacing between source and

substrate, process pressure and process gas (Ar, He, or Mix) [3, 5]. The spacing between source and substrate was fixed at 2 mm and the process temperature was monitored by inserting thermocouple into the graphite susceptor and controlled by a PID controller. Two 2 kW tungsten halogen lamps were used, one at the bottom and the other at the top of the quartz tube for heating the source and substrate as illustrated in Fig. 2. The chamber was kept at a vacuum of 10 mTorr by a rough pump before deposition and then purged with highly pure N_2 gas. After that Ar gas was applied to the chamber. During the deposition, the chamber pressure was monitored by Instrutech pressure gauge and maintained at 1.5 Torr by regulating a throttle valve connected between the rough pump and the chamber. The CdTe films were deposited at four different source temperatures starting at 610 °C and then increased to 670 ^{0}C in three steps of 20°C equal increment while keeping the difference between the source and substrate temperature constant at 30 °C. The deposition time of each films were maintained to10 minutes. The as-deposited CdTe TF is shown in Fig. 4.

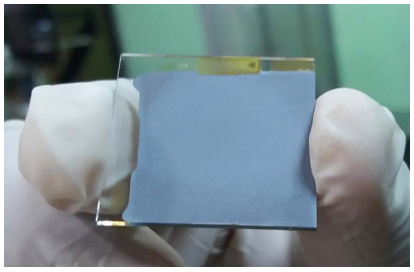


Fig. 4 As-deposited CdTe films on CdS films

II. Results and Discussion

The three major electrical properties including carrier concentration, mobility and resistivity of the as-deposited CdTe thin films were observed at room temperature by Hall/resistivity measurements. The ECOPIA Hall effect measurement system (HMS-3000) which is integrated with a resistivity measurement tool was used for this purpose. The system uses the van der Pauw method which is a four-probe method for the Hall measurements. During the measurements, a constant magnetic field was produced by using a pair of permanent magnets and then applied orthogonally to the upper surface of the films. Meanwhile, a DC (direct current) was made to pass through the thin film samples by using one of the two diagonal pairs of gold electrodes, electrically connected to a constant dc current source. The induced hall voltage was detected by the remaining electrode pair. The magnitude of the applied magnetic field was constant at 0.51 T. The type (p or n) of the CdTe films was denoted by the system using a positive (+) or negative (-) sign preceding the numerical value of the carrier concentration. The details results are discussed in the following subsections.

A. Effect of Temperature on Carrier Concentration

The calculation of the carrier concentration and mobility was done by the system itself using the magnitude of the induced hall voltage. Fig. 5 shows the effects of source temperature on the majority carrier concentration in the CdTe films. It is evident from Fig. 5 that the carrier concentration increases with the increase in source temperature. The values of carrier concentration were found 6.38×10^{11} cm^{-3}, 4.2×10^{12} cm^{-3}, 7.7×10^{12} cm^{-3}, and -6.78×10^{12} cm^{-3} for the samples deposited at source temperatures of 610 ^{0}C, 630 ^{0}C, 650 ^{0}C and 670 ^{0}C respectively. The highest carrier concentration was found in the sample deposited at 650 ^{0}C source temperature. As the photovoltaic conversion efficiency is directly related to the carrier concentration, the CdTe films deposited at this source temperature are expected to yield the more efficient solar cells than the films deposited at other source temperature which is very consistent with the results reported by the other researchers [4-6].

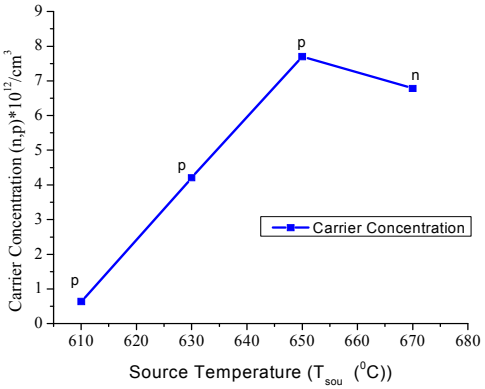


Fig. 5 Effect of temperature on CdTe TF carrier concentration

Another interesting observation is that, the CdTe films deposited at source temperatures higher than 650 ^{0}C become n-type semiconductor which is not expected in case of CdTe thin film solar cells because the CdTe is usually used as the p-type absorber layer. These might be happened for source temperature higher than 650 ^{0}C, where the sublimation rate of the cadmium atoms on the CdS layer surpasses that of tellurium atom. As a result, the percentage of cadmium becomes higher than the tellurium in the CdTe films and thus the films become n-type.

B. Effect of Temperature on Carrier Mobility

The variation of majority carrier mobility with the source temperature is shown in Fig. 6.

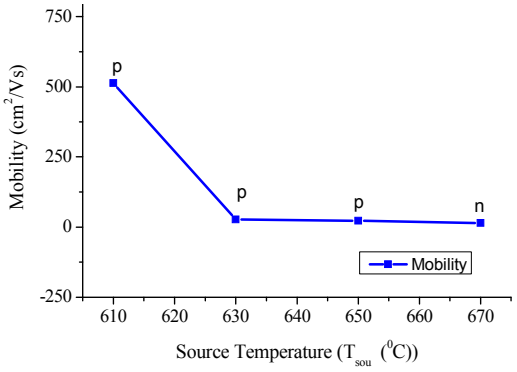


Fig. 6 Effect of temperature on CdTe TF carrier mobility.

In the figure it is observed that, hole mobility is highest (512 cm^2/Vs) in the sample deposited at 610 ^{0}C source temperature. As the source temperature increases from 610 ^{0}C to 630 ^{0}C, the hole mobility decreases significantly from

512 cm^2/Vs to 25 cm^2/Vs. After that, the rate of decrease in mobility become not significant. For the source temperature of 630 ^{0}C and 670 ^{0}C, the mobility was found to be 22 cm^2/Vs (hole) and 15 cm^2/Vs (electron) respectively. The decrease in carrier mobility with the increase in source temperature is agreeable with the theory. As the source temperature increases, the carrier concentration also increases. As a result the mean free path of the carriers decreases and the total number of collisions per carrier to travel a certain distance increases consequently decreasing the carrier mobility.

C. Effect of Temperature on Resistivity

The Fig. 7 demonstrates the change in film resistivity with the increase in source temperature is shown. It is clear from Fig. 7 that the resistivity shown an increasing trend with the increase in source temperature.

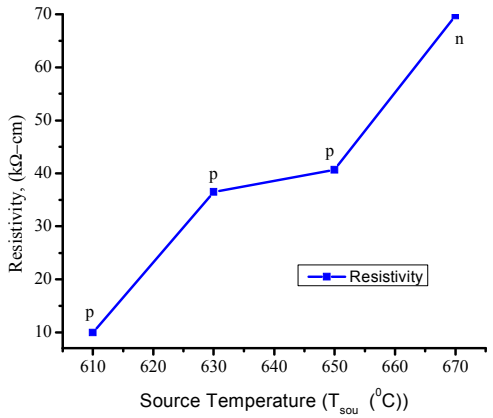


Fig. 7 Effect of temperature on CdTe TF resistivity

The resistivity was found 9.96 kΩ-cm, 38.6 kΩ-cm, 40.7 kΩ-cm and 70 kΩ-cm for the samples deposited at source temperature of 610 ^{0}C, 630 ^{0}C, 650 ^{0}C and 670 ^{0}C respectively. Among the p-type CdTe films, the highest resistivity was found in the sample for 650 ^{0}C source temperature. The resistivity of as-deposited CdTe films were lower than the results reported by the other researcher [15-18]. The n-type CdTe thin film offers the highest resistivity among the four CdTe TF samples. This may be because of the lowest mobility found in n-type CdTe as-deposited films.

In general, a higher carrier concentration, lower resistivity and higher mobility are desirable for the absorber layer to increase open-circuit voltage V_{oc}, fill factor FF and short-circuit current density J_{sc} of the solar cell. It is very difficult to achieve all these suitable properties in a single film. Therefore, in practice rigorous optimization is recommended for this purpose. Considering all the electrical properties finally it is concluded that, the CdTe film deposited at 650 °C source temperature is suitable as the absorber layer in solar cell applications.

III. CONCLUSION

The performance parameters of CdTe solar cell are directly depends on the micro-structural, electrical and optical properties of different functional layers. In CSS technique the direct phase transition occurs between the CdTe solid to gas state congruently. The main focus of the work was to investigate the temperature dependence of electrical properties of as grown CdTe TF. Both the p-type and n-type conductivity of CdTe TF were found for different source and substrate temperatures. The carrier mobility of the CdTe TF dramatically decreases for the source temperature increasing from 610 °C to 630 °C, whereas the resistivity of the CdTe TF increase with source temperature. The effects of the source and substrate temperature on the electrical properties of CdTe films were found consistent with the other related published works. The electrical properties of the CdTe TF at 650 °C source temperature were found to be better than the counterparts deposited in dynamic condition.

REFERENCES

[1] First Solar, "First Solar Series 6." [Online]. Available:http://www.firstsolar.com/Modules/Series-6. [Accessed: 19-Feb-2019].

[2] M. A. Green et al., "Solar cell efficiency tables (Version 53)," *Prog. Photovoltaics Res. Appl.*, vol. 27, no. 1, pp. 3–12, 2019.

[3] C.S. Ferekides, U. Balasubramanian, R. Mamazza, V. Viswanathan, H. Zhao, D.L. Morel, "CdTe thin film solar cells: device and technology issues," *Solar Energy*, vol. 77, pp. 823–830, 2004.

[4] J. Sites, A. Munshi, J. Kephart, D. Swanson, and W. S. Sampath, "Progress and challenges with CdTe cell efficiency," *2017 IEEE 44th Photovolt. Spec. Conf. PVSC 2017*, no. June, pp. 1–4, 2018.

[5] J. Britt and C. Ferekides, "Thin-film CdS/CdTe solar cell with 15.8% efficiency," *Appl. Phys. Lett.*, vol. 62, no. 22, pp. 2851–2852, 1993.

[6] X. Wu, "High-efficiency polycrystalline CdTe thin-film solar cells," *Sol. Energy*, vol. 77, no. 6, pp. 803–814, 2004.

[7] A. Seth, G.B. Lush, J.C. McClure, V.P. Singh, D. Flood, "Growth and characterization of CdTe by closespaced sublimation on metal substrates," *Solar Energy Materials & Solar Cells*, vol. 59, pp. 35-49, 1999

[8] A. Bosio, G. Rosa, and N. Romeo, "Past, present and future of the thin film CdTe/CdS solar cells," *Sol. Energy*, vol. 175 no. Nov 2018, pp. 31–43, 2018.

[9] Benjamin R Wakeling, " Close space sublimation of CdTe for solar cells and the effect of underlying layers," PhD Thesis, Dept. of Applied Science, Security and Resilience, Cranfield University, August 2009

[10] B.A. Korevaar , A. Halverson, J. Cao, J. Choi, C. Collazo-Davila, W. Huber, "High efficiency CdTe cells using manufacturable window layers and CdTe thickness," *Thin Solid Films*, vol. 535, pp. 229-232, 2013.

[11] Naba R. Paudel, Yanfa Yan, "Fabrication and characterization of high-efficiency CdTe-based thin-film solar cells on commercial SnO2:F-coated soda-lime glass substrates," *Thin Solid Films,* vol. 549, pp. 30–35, 2013.

[12] Russell M. Geisthardt, Marko Topic, James R. Sites, "Status and Potential of CdTe Solar-Cell Efficiency, *IEEE Journal of Photovoltaics*, vol. 5, no. 4, July 2015.\

[13] J. Nelson, *The physics of Solar Cells*, 1st ed. World Scientific Publishing Company, 2003.

[14] Samuel H. Demtsu, "Impact of Back-Contact Materials on Performance and Stability of CdS/CdTe Solar Cells," PhD thesis, Dept. of Physics, Colorado State University Fort Collins, Colorado, June, 2006.

[15] Bülent M. Başol, Brian McCandless, "Brief review of cadmium telluridebased photovoltaic technologies*," Journal of photonics for Energy*, vol. 4, pp. (040996) 1-11, Jun. 27, 2014.

[16] M.A. Flores Mendoz, R. CastanedoPe´rez, G. Torres Delgado, J. Ma´rquezMarı´n, A. Cruz Orea, O. Zelaya Angel, "Structural, morphological, optical and electrical properties of CdTe films deposited by CSS under an argon–oxygen mixture and vacuum," *Solar Energy Materials & Solar Cells*, vol. 95, pp. 2023–2027, 2011.

[17] F. M. T. Enam, K.S. Rahman, M. I. Kamaruzzaman, K. Sobayel, M. Aktharuzzaman, N. Amin, "An Investigation on Structural and Electrical Properties of Close-Spaced Sublimation Grown CdTe Thin Films in Different Growth Conditions," *Chalcogenide Letters*, vol. 14, no. 4, pp. 125 – 131, April 2017.

[18] Peng Gu, Xinghua Zhu, Jitao Li, Haihua Wu, Dingyu Yang, " Influence of sputtering power on structural, optical and electrical properties of CdTe thin films prepared by DC magnetron sputtering," *Journal of Materials Science: Materials in Electronics*, vol. 29, pp. 14635–14642, 2018.

Layer thickness effect on power conversion efficiency of a P_3HT:PCBM based organic solar cell

Raian Islam
Dept. of Electrical and Electronic Engineering
Ahsanullah University of Science and Technology.
raianislam16@gmail.com

Murad Mehrab Abrar
Dept. of Electrical and Electronic Engineering
Ahsanullah University of Science and Technology.
mehrababrar@gmail.com

Farhatul Hassan
Dept. of Electrical and Electronic Engineering
Ahsanullah University of Science and Technology.
farhatulhassan1710@gmail.com

Sifat Adnan
dept. of Electrical and Electronic Engineering
Ahsanullah University of Science and Technology.
sifatmimo19@gmail.com

***Abstract*—In recent years Organic Solar Cells have become a prominent topic of research to achieve an optimum efficiency at low cost. By using the GPVDM software on this paper, we have analyzed the effect on power conversion efficiency, changing simultaneously both the polymers and blending layer thickness of P3HT:PCBM based solar cell. Our main goal is to find which layer change keeps the power conversion efficiency output better. After comparing them, the result shows that setting the active layer (P3HT:PCBM) in an optimum fixed value and varying the polymer layer (PEDOT:PSS) gives a better output of PCE. In our paper, from the data in both cases, The highest efficiency is 4.50 percent where P_3HT:PCBM layer thickness is 2.2×10^{-7} m and PEDOT:PSS layer thickness is 1×10^{-7} m.**

***Keywords*—P3HT:PCBM, PEDOT:PSS, PCE, Organic solar cell, active layer, blending layer, polymer layer.**

I. INTRODUCTION

Organic bulk heterojunction (BHJ) solar cells will have a very large impact on renewable energy in near future. Being flexible and semitransparent, organic solar cells are different from the inorganic solar cells[1]. In organic solar cells, organic polymers are used to absorb the light and for transporting the charge in order to produce electricity. They can cover larger range of surfaces than inorganic solar cells and the manufacturing costs are significantly lower than the conventional inorganic cells[2],[3]. They also can be processed easily than inorganic solar cells [7]. It is called bulk heterojunction solar cells as electron acceptor and donor materials are blended in a disordered orientation which limits its Power Conversion Efficiency (PCE)[4],[5].

In order to improve the performance and power conversion efficiency, we have changed the layer thickness and observed the characteristics curve which we have generated using General-Purpose Photovoltaic Device Model (GPVDM)[6]. It is a free general-purpose software (a potential device simulator) for photovoltaics where the model has been tested in steady state time domain (usfs time scales and fx domain) by doing various experiments. Drift diffusion model, Maxwell's equation based optical model, SRH (Shockley-Read-Hall) model which has been used in this software, has been crosschecked with analytical solutions. In our work, by changing the layer of thickness of PEDOT: PSS (poly(3,4-ethylenedioxythiophene)) and the active layer P_3HT: PCBM (poly (3-hexylthiophene): phenyl-C61-butyric acid methyl ester) we have investigated the power conversion efficiency.

II. DEVICE SIMULATION PARAMETERS

A. Structure of GPVDM Software model

To see the changing effect of layer thickness, we have established a simulation model using GPVDM. The model has two parts, I) The graphical interface, II) The back-end solver. Graphical interface is where we edit all the input parameters. It also shows us the result by using python program. Back-end solver is the place where all the calculations of GPVDM is done. It is written in C program.

B. Electrical Model

To perform the device simulation, the built-in potential, charge carrier transport, carrier tapping, free to free carrier recombination of the device are needed to be calculated. In GPVDM software, solar cell simulation is done by solving these basic equations. Poisson's equation is given by-

$$\frac{d}{dx}\,\varepsilon_0\varepsilon_r\,\frac{d\phi}{dx} = q(n-p) \qquad (1)$$

Where ε_0 is the permittivity of free space, ε_r is relative permittivity, ϕ is the voltage profile, q is the elementary

charge of electron, n is the free electron concentration and p is the free hole concentration. Equation (2) and (3) are the bipolar drift and diffusion equations for electrons and holes respectively[6].

$$J_n = q\,\mu_e\, n\frac{dE}{dx} + qD_n\frac{dn}{dx} \qquad (2)$$

$$J_p = q\,\mu_p\, p\frac{dE}{dx} - qD_p\frac{dp}{dx} \qquad (3)$$

Where J_n is the electron current density, μ_e is the electron mobility, D_n is the electron diffusion coefficient, J_p is the hole current density, μ_p is the hole mobility, D_p is the hole diffusion coefficient and E is electric field. Equation (4) and (5) are the electron and hole continuity equations in 1D and time domain.

$$\frac{dJ_n}{dx} = q\left(R_n - G + \frac{dn}{dt}\right) \qquad (4)$$

$$\frac{dJ_p}{dx} = -q\left(R_p - G + \frac{dp}{dt}\right) \qquad (5)$$

Where, R_n is the net recombination rate for electrons, R_p is the net recombination rate for holes and G is the free carrier generation rate. The details of the above equations can be found.

The 3-D graphical representation of the planner heterojunction architecture of our solar model is shown in figure-1. Here P_3HT:PCBM is the active layer. The layers configuration used in this simulation are respectively ITO/PEDOT:PSS/P3HT:PCBM/Al. Here, ITO (Indium Tin Oxide) is the oxide layer, PEDOT:PSS is the polymers blending with P3HT:PCBM and Al is the metal layer. Initially the layer thickness of PEDOT:PSS is preset to 1×10^{-7} m, P_3HT:PCBM to 2.2×10^{-7} m, ITO and Al both are preset to 1×10^{-7} m by default in the simulator. The values are given below in Table-1.

TABLE I: SIMULATION PARAMETER

Layer Name	Thickness (m)	Layer Type
ITO	1×10^{-7}	Contact Layer
PEDOT:PSS	1×10^{-7}	Polymer Layer
P_3HT:PCBM	2.2×10^{-7}	Active/Blending Layer
Al	1×10^{-7}	Contact Layer

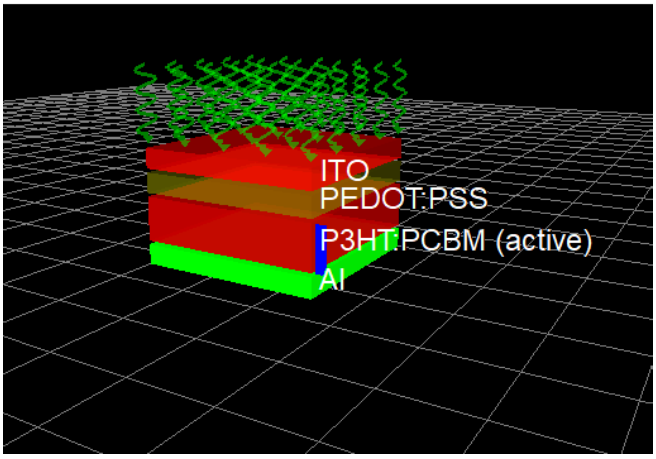


Fig-1: The planner heterojunction architecture of P_3HT:PCBM based solar cell.

In our work, we have varied both the active layer and the polymer layer thickness. We first kept PEDOT:PSS layer fixed at 6 different thicknesses and varied P_3HT:PCBM layer under each thickness (case-I). Then again we kept P_3HT:PCBM layer fixed at 6 different thicknesses and varied PEDOT:PSS layer (case-II). For individual data sets, we have obtained PCE curves.

III. Result And Discussion

In general studies, a common model P_3HT:PCBM (only active layer) structure is used [5]. But here, in case-I, we varied P_3HT:PCBM (active layer) for 6 different values of PEDOT:PSS (polymer layer). Varying P_3HT:PCBM layer thickness from 1.2×10^{-7} m to 10.2×10^{-7} m while PEDOT:PSS layer is fixed at 1×10^{-7} m we have obtained the maximum efficiency curve. From our data we observed that P_3HT:PCBM layer thickness of 2.2×10^{-7} m and PEDOT:PSS layer thickness of 1×10^{-7} m gives us the maximum PCE point of 4.50% (which is by default set in GPVDM), with voltage at maximum power = 0.468738 V, an open circuit voltage of = 0.602238 V and a short circuit current of = -111.634936 A/m^2. Figure-2 shows the graphical comparison among different layer thickness and their effects on PCE. It can be seen from Figure-2 that when PEDOT:PSS layer is fixed at 2×10^{-7} m, there is a sudden increase in PCE at 8.2×10^{-7} m layer thickness of P3HT:PCBM, but not so significant (2.4%). In case-II, we kept P_3HT:PCBM layer fixed at 6 points and varied PEDOT:PSS layer from 1×10^{-7} m to 10×10^{-7} m and plotted the values following the same procedure. Figure-3 shows the PCE curve of data from case-II.

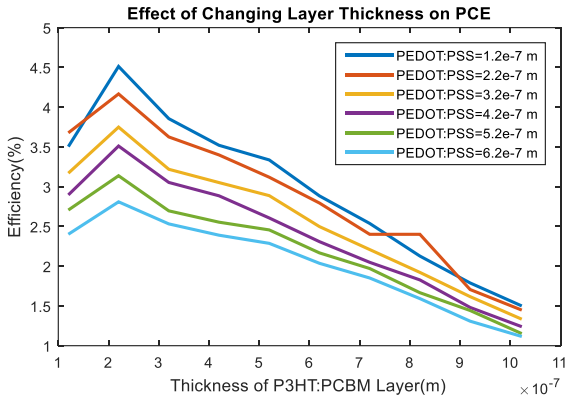


Fig-2: case-I where PEDOT:PSS layer is kept fixed and P_3HT:PCBM is varied for each thickness-PCE curve.

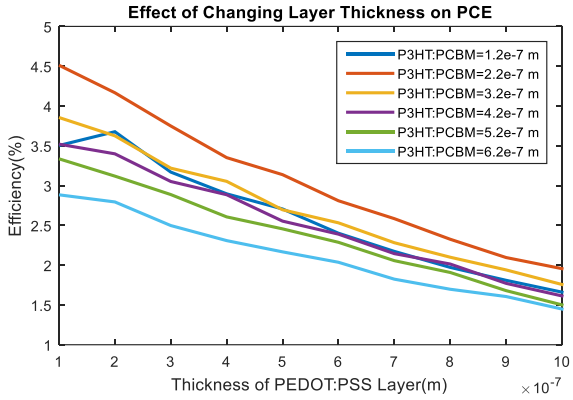


Fig-3: case-II where P_3HT:PCBM layer is kept fixed and PEDOT:PSS is varied for each thickness-PCE curve.

In case-II we can note that when P_3HT:PCBM layer thickness is 2.2×10^{-7} m and PEDOT:PSS layer thickness is 1×10^{-7} m PCE reaches to the maximum of 4.50% which is similar as case-I, then decreases gradually. If we compare the two curves, we can distinguish the better PCE output easily. It is clear from both curves that the rate of decrease of PCE in case-I is higher than case-II. The lowest PCE in case-II is 1.449% which is higher than the lowest PCE found in case-I, 1.121%. The I-V characteristics curve at maximum efficiency point is given in Figure-4.

IV. Conclusion

We have analyzed the Power Conversion Efficiency of a P_3HT:PCBM based BHJ solar cell in our work. The results show that a proper choice of active layer and blending layer thickness can give us an optimum output and increase the efficiency. From the layer thickness- PCE curve, we can note that increasing layer thickness can drop the PCE ratio drastically, as BHJ solar cells deal with thin films.

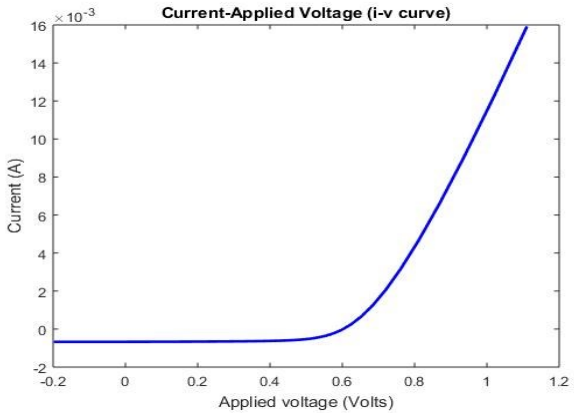


Fig 4: I-V characteristics curve at maximum efficiency point where PEDOT:PSS layer thickness is 1×10^{-7} m and P_3HT:PCBM layer thickness is 2.2×10^{-7} m

We also can conclude from our analysis that changing the polymer layer (PEDOT:PSS) thickness keeping the active layer (P_3HT:PCBM) fixed can give a better PCE output rather than changing the active layer thickness. So, if we keep the active layer in an optimum value and change the polymer layer, better PCE can be obtained. In near future more PCE improvement can be done by changing the active material layer and polymer layer.

References

[1] Akihiro Kojima, Kenjiro Teshima, Yasuo Shirai and Tsutomu Miyasaka, Organometal halide perovskites as visible-light sensitizers for photovoltaic cells, *Journal of the American Chemical Society*, Vol. 131, no. 17, pp. 6050-6051,2009,.

[2] Roderick C. I. MacKenzie, Thomas Kirchartz, George F. A. Dibb, and Jenny Nelson, Modeling Nongeminate Recombination in P3HT:PCBM *Solar Cells, J. Phys. Chem. C*, Vol. 115, no. 19, 2011, pp 9806–9813.

[3] R. Hanfland, M.A. Fischer, W. Brütting, U. Würfel, R.C.I. MacKenzie, The physical meaning of charge extraction by linearly increasing voltage transients from organic solar cells, *Appl. Phys. Lett*., Vol. 103, no. 6, 2013, pp 063904.

[4] F. Deschler, D. Riedel, B. Ecker, E. von Hauff, E. Da Como, R.C.I. MacKenzie, Increasing organic solar cell efficiency with polymer interlayers, *Phys. Chem. Chem. Phys*., Vol. 15, no. 3, 2012, pp 764-769.

[5] R.C.I. MacKenzie, C.G. Shuttle, M.L. Chabinyc, J. Nelson, Extracting microscopic device parameters from transient photocurrent measurements of P3HT:PCBM solar cells, *Adv. Energy Mater*., Vol. 2, no. 6, 2012, pp 662-669.

[6] A. Hima1, A. Khechekhouche, I. Kemerchou, N. Lakhdar,B. Benhaoua, F. Rogti, I. Telli, A. Saadoun,GPVDM simulation of layer thickness effect on power conversion efficiency of CH3NH3PbI3 based planar heterojunction solar cell, *International Journal of Energetica (IJECA)*, Volume 3. Issue. 1. 2018

[7] Nazerdeylami, Somayeh & Dizaji, Hamid. (2014). Effect of Active Layer Thickness on Characteristics of P3HT:PCBM based Organic Solar Cells.

An Efficient CZTS Solar Cell from Numerical Analysis

E.M.K. Ikball Ahamed[1, 2*], A.K. Sen Gupta[1, 2], M. Qamruzzaman[1, 2], M.A. Matin[1, 2]

[1]Department of Electrical and Electronic Engineering, Chittagong University of Engineering and Technology, Chattogram-4349, Bangladesh

[2]Renewable Energy Laboratory (REL), Chittagong University of Engineering and Technology, Chattogram-4349, Bangladesh

*ikballahamed@cuet.ac.bd

***Abstract*—The Copper Zinc Tin Sulfide (CZTS) is considered as one of the promising semiconductor materials in Kesterite family for potential solar cells absorber material. This work focuses on the numerical model and simulation of efficient CZTS based thin-film solar cell. Popular thin-film solar cell simulator wxAMPS was used for simulation of (Al:ZnO/i-ZnO/CdS/CZTS/Mo)structured solar cell and the effects of different layer properties on cell performance parameters were analyzed. An optimized cell with 15.84% (%FF=72.86, *Voc*=0.78Volt and *Jsc*=27.98 mA/cm^2) energy conversion efficiency has proposed. The temperature coefficient (TC) of 0.02%/°C had shown the higher stability of the proposed cell. All of the simulated results had encouraged towards fabrication of high performance thin-film CZTS solar cells.**

Keywords*— *Thin-film solar cell, CZTS, Numerical Simulation, wxAMPS, Conversion efficiency

I. Introduction

The world is approaching towards big crisis of energy as most of the conventional energy resources are going to be depleted. In this situation, solar photovoltaic is one of the most effective ways for the direct power generation from the sunlight. In recent years, photovoltaic keeps on being one of the quickest developing ventures with development rates beyond 40% per year [1].

Among the photovoltaic technologies, Thin-film Solar Cells (TFSCs) have gained popularity for its material usage in a lower scale. Popular TFSCs are Chalcopyrite based polycrystalline binary semiconductor material CdTe (Cadmium Telluride) and quaternary semiconductor material $Cu_2In_xGa_{1-x}Se_2$ (CIGS) is commercially successful. However, there are some restrictions on the usage of limited material Indium (In) and Tellurium (Te) which are raised problems about restriction on generation limit of Photovoltaic (PV) cells.

For solving the issues, Cu_2ZnSnS_4 (CZTS), a quaternary semiconductor material consists of earth abundant material Zn and Sn is one of the hot candidates for PV technology. Abundance of Sn and Zn in earth is very much higher than In (45 times and 1500 times correspondingly) and the price of Indium is almost in the order of twice higher than Sn and Zn [2]. Bandgap energy of CZTS is in the range of 1.4eV to 1.5eV which is suitable for photovoltaic conversion [3]. Furthermore high absorbance capability (absorption coefficient, α is more than 10^6/m) ensures the good possibility of photovoltaic conversion.

Multiphase defects or impurities like ternary and quaternary compounds are more easily formed than CZTS formation. So, pure CZTS crystal formation is a very challenging task and it is formed in a narrow region [2-3].

Highest conversion efficiency of CZTS based photovoltaic solar cell is 10% (confirmed) and 11% (notable exceptions) were fabricated by UNSW [4], Australia whereas William Shockley and S. J. Queisser gives the limiting value of conversion efficiency of single junction solar cell is 30% [5-6]. So there is a huge scope to explore the performance of CZTS based photovoltaic solar cells.

In this research paper, Numerical simulation of CZTS based solar cells has been done with solar device simulator wxAMPS. The cells are simulated by varying the properties of different layers and the important performance parameters were analyzed.

II. Device Structure and Materials

A. Basic Format of CZTS solar Photovoltaic Cell

The photovoltaic effect in heterodiode consists of transparent conducting oxide: Cadmium Tin Oxide (CTO) and CZTS thin film is first reported by Ito and Nakezawa [7]. P-CZTS material is the main absorber layer in CZTS solar cell. The general structure of CZTS based solar cell (Aluminum doped ZnO/intrinsic-ZnO/CdS/CZTS/Mo/Glass substrate) is shown in Fig.1.

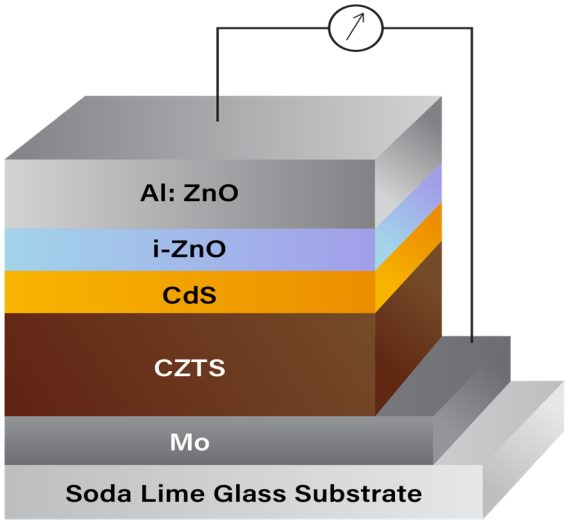


Fig.1. CZTS solar cell structure (thickness is not scaled)

CdS and CZTS are the buffer and absorber layer respectively form the p-n junction and constitute the key part of the device. Aluminum doped ZnO is transparent conducting oxide which creates an n-n+ junction before metal semiconductor contact junction, ensures a low resistive contact for electron transport. The whole structure should be grown layer by layer into a substrate to form a stable structure, because the layer thickness is in the order of nm to µm scale, the cell has no rigidity or mechanical strength of

its own. Usually cell deposited on glass substrate for mechanical strength.

B. Device Modeling and Simulation

Before any type of expensive fabrication, the performance of a cell can be estimated by numerical simulation. For simulation, electrical modeling of layers is needed. Popular solar cell simulator wxAMPS (Analysis of Microelectronic and Photonic Structures with wxWidgets) was used in this work. The wxAMPS is a refresh variant of well-known simulator AMPS (Analysis of Microelectronic and Photonic Structures) created by S. Fonash and refreshed by Angus Rockett [8]. Calculation of wxAMPS is altered by consolidating Newton and Gummel techniques and is a decent apparatus to break down various sorts of solar photovoltaic cells, for example, CdTe and CIGS thin films and different materials. Intra band tunneling model in wxAMPS works extremely well for solar cells of low deformity, for example, CIGS and CdTe Solar cells [9].

In this paper we consider a basic structure of CZTS absorber material based TFSCs consists of CdS heteropartner as described in Fig.1. Material properties utilized in this simulation, which depended on literature, experiments [10-13] and in some case, based on reasonable estimations. Electrical properties of different layers are shown in TABLE I.

TABLE I: OPTIMIZED ELECTRICAL PROPERTIES OF DIFFERENT LAYERS OF CZTS PV CELL

	Layers			
Layer Properties	Al:ZnO	i-ZnO	CdS	CZTS
Thickness (μm)	0.1	0.06	0.06	1.8
Energy Bandgap, E_g (eV)	3.3	3.3	2.4	1.5
Electron Affinity, χ (eV)	4.4	4.4	4.2	4.3
Relative Permittivity, ε_r	9	9	10	7
Density of states in CB, N_c (cm^{-3})	2.22×10^{18}	2.22×10^{18}	2.22×10^{18}	2.22×10^{18}
Density of states in VB, N_v (cm^{-3})	1.78×10^{19}	1.78×10^{19}	1.78×10^{19}	1.78×10^{19}
Electron and Hole Mobility, μ_n and μ_p ($cm^2/v/s$)	100 and 25	100 and 25	100 and 25	100 and 25
Doping Density, N_a or N_d (cm^{-3})	10^{20} (D)	10^{18} (D)	10^{18} (D)	10^{16} (A)

The symbol D means donor and A means acceptor in TABLE I.

The simulation works are done by indicating the device structure and electrical and optical parameters of materials as input contentions of the test system. At that point determined the energy band diagram, recombination profile and transport properties are determined based on the Poisson equation just as continuity conditions. Recombination in profound mass dimensions and their occupation are portrayed by Shockley-Read-Hall (SRH) recombination concept.

III. RESULT AND DISCUSSIONS

Simulations were done in wxAMPS environment to examine the effect of different layer properties on cell performance. Further the results were analyzed to fit an optimized result.

A. Effect of thickness of Al:ZnO(AZO)

To investigate the effect of Aluminum doped ZnO (AZO) layer thickness on the performance parameters of the cell, simulations were done in wxAMPS for different thickness of AZO varying from 50 nm to 500 nm while all other properties of different layers of the cell were kept constant. The simulation results are shown in Fig.2.

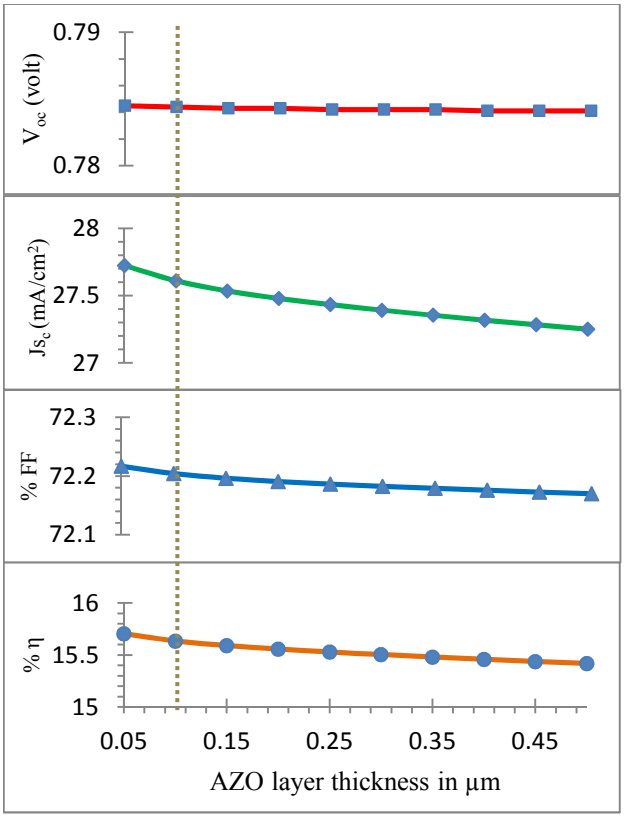


Fig.2. Performance parameters versus AZO layer thickness

The simulation result shows that V_{oc}, J_{sc} and FF are shrinked with the growth of AZO due to the higher absorption of light in the low wavelength region of sun spectrum. So the AZO thickness has been chosen in the lower value (100 nm) which can be easily deposited by popular deposition technique sputtering.

B. Effect of thickness of CdS layer

Cadmium Sulfide (CdS) having a bandgap energy of around 2.42eV, guarantee most of the incident light assorted towards junction. Hypothetically, n-layer thickness ought to be minimal to recoil the series resistance of the device. In the meantime a decreased thickness of CdS allows low absorption in the blue region of the sun spectrum as well as improved short circuit current density, J_{sc}.

Simulations were done by varying the thickness of CdS from 50 nm to 200 nm and the effects in performance parameters were observed. Obtained simulated results are shown in Fig.3.

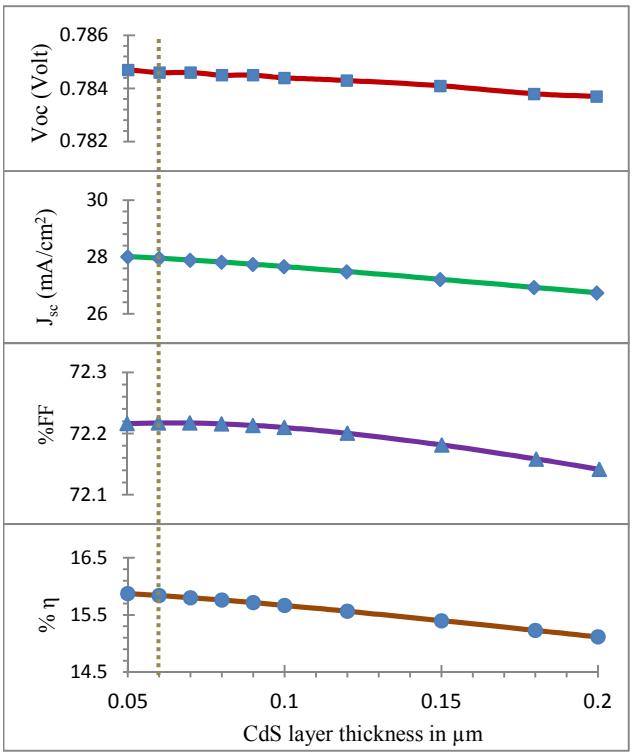


Fig.3. Effect of CdS layer thickenss on cell performance

The simulation results shows that V_{oc}, J_{sc} and FF are shrinked with the growth of CdS as described in the previous paragraph. So the CdS thickness has been chosen in the lower value (60 nm) which can be easily deposited by very popular and cheap deposition technique named Chemical Bath Deposition (CBD).

C. *Effect of absorber layer (CZTS) thickness*

To get acceptable photo conversion efficiency with thin absorber layer is one of the main challenges of thin-film PV cells. The reliance of performance of CZTS solar PV cell on the absorber thickness is considered by changing the CZTS layer thickness differs from 0.1µm to 3 µm while other parameters of the layer as well as other layers are kept unaltered. The results are shown in Fig.4.

It is noted that the V_{oc} and J_{sc} increased with absorber layer thickness due to the large amount of absorption in the longer wavelength region of the sun spectrum and more electron-hole pair generation. Due to the increment of both V_{oc} and J_{sc}, efficiency has been also improved with increasing thickness of the absorber layer. For CZTS thickness greater than 1.8µm, efficiency has shown a decreasing trend due to the bulk recombination and then the decreasing trend of fill factor dominates for thicker absorber layer. Bulk use of materials means more costly device, then the cost of device plays an important factor and a trade of between efficiency and cost should be considered in design.

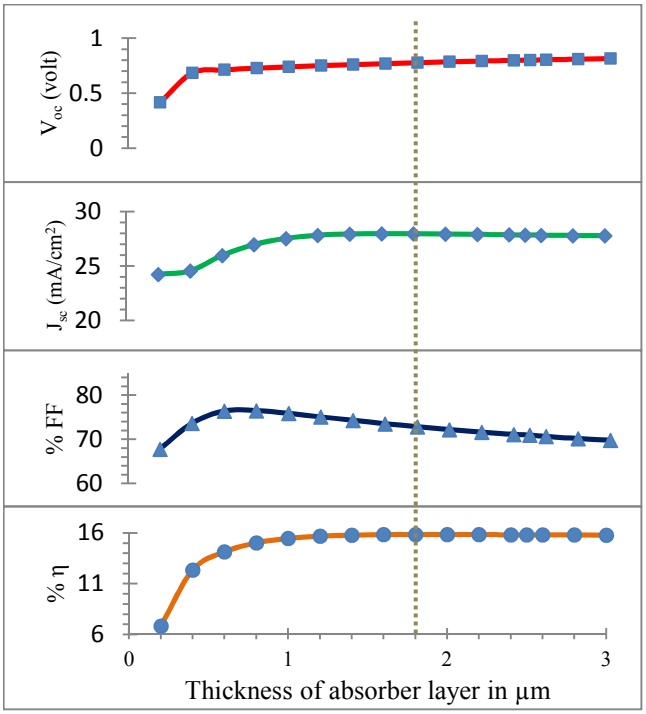


Fig.4. Dependency of Performance parameters on absorber layer thickness

D. *CZTS doping density*

Further simulations were done in wxAMPS to investigate the effect of absorber layer doping profile from 10^{14} cm^{-3} to 10^{17} cm^{-3} in the performance parameters while all other properties of CZTS as well as other layers were kept unchanged. The results are shown in Fig.5.

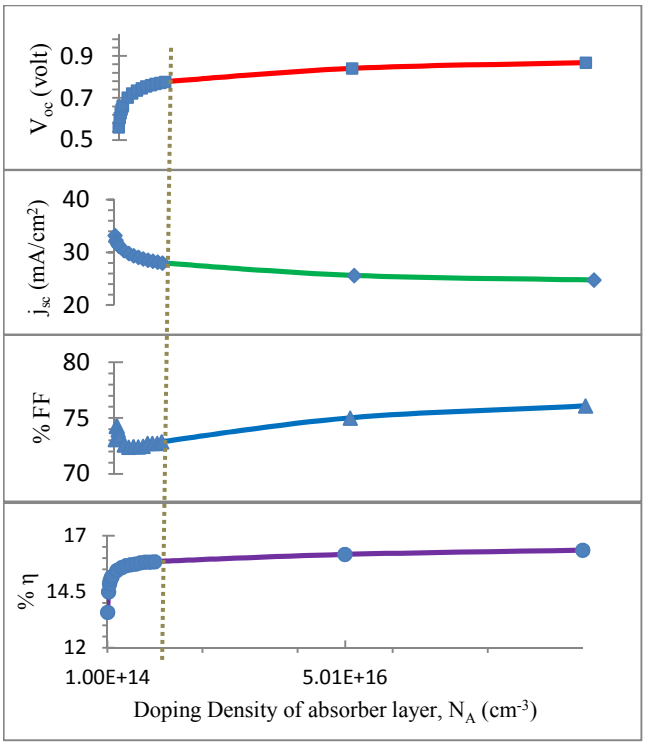


Fig.5. Performance parameters versus absorber layer doping profile

Fig.5. illustrates that the open circuit voltage has increased while the short circuit current density has decreased with the doping profile of CZTS. Fig.6. is an alternative representation of the dependency of V_{oc} and J_{sc} with doping profile of absorber layer. A series of J-V curves are illustrated in a same plot for different doping profile of CZTS absorber.

Dependency of V_{oc} with absorber doping (N_A) can be analyzed as in (1), where I_L is photo generated current and I_o is diode saturation current which is inversely proportional to doping concentration, N_A.

$$v_{oc} = \frac{KT}{q} \ln\left(\frac{I_L}{I_o} + 1\right) \quad (1)$$

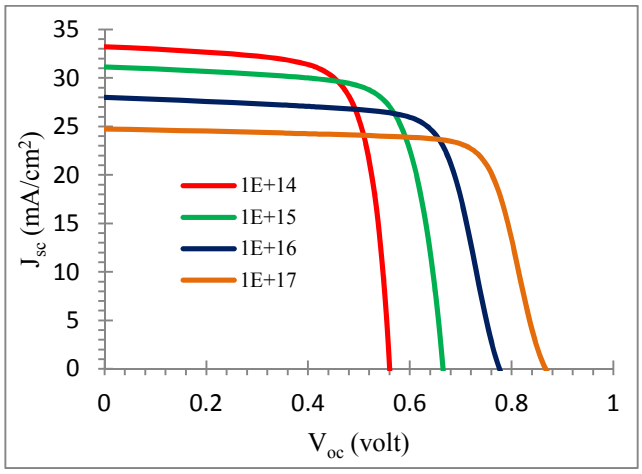


Fig.6.Representation of J-V curves for different doping profile of CZTS absorber layer

It is clear from Fig.6 that I_o is decreased with larger doping profile of the absorber layer and further reduction in I_o causes higher V_{oc} for higher N_A according to the equation 1 as well as simulation results show the same trend of V_{oc} as shown in Fig.5. and Fig.6. Short circuit current density, J_{sc} is decreased with increasing doping density due to the enhanced recombination probability. This will create an obstacle in the path of electron transport and hence reduces the J_{sc}. Disturbance in carrier transport as well as recombination also reduces the external quantum efficiency. The value of N_A had set at 10^{16} cm^{-3} because fabrication of CZTS with heavy doping is very tough.

E. Effect of absorber layer defect density in cell performance

There are so many studies in literature demonstrated that the donor type imperfections were formed strongly than the acceptor type in CZTS material [10]. So the single acceptor type defect is introduced to analyze the defect state of CZTS absorber. The effect of defect density of CZTS in performance is presented in Fig.7.

It is observed that the V_{oc} and J_{sc} are not affected strongly in low level defect but for the defect level 10^{14} and higher, the V_{oc} and J_{sc} were abruptly affected by the defect states as well as the cell performance was strongly diminished. The possible reason is a large no of trap states which further causes huge recombination of photo-generated carrier as well as reduction in J_{sc}, V_{oc} and finally the conversion efficiency.

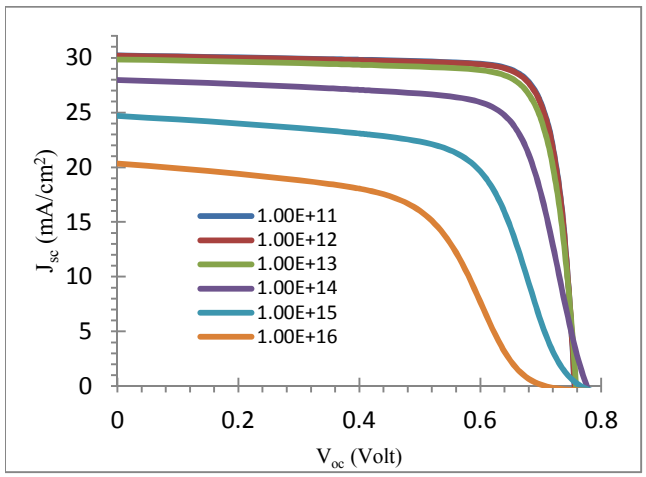


Fig.7. A series of J-V curves with variety of CZTS defect density

F. Effect of absorber layer thickness on Quantum Efficiency

Fig.8. shows a series of external quantum efficiency curves for different thickness of CIGS absorber layer. Visible change in quantum efficiency has occurred up to 1.5µm thickness of CZTS absorber. Variation in CZTS layer thickness affects the external quantum efficiency because higher absorber layer thickness results in more photon generated carrier which offers good possibility to get more carriers in the external circuit.

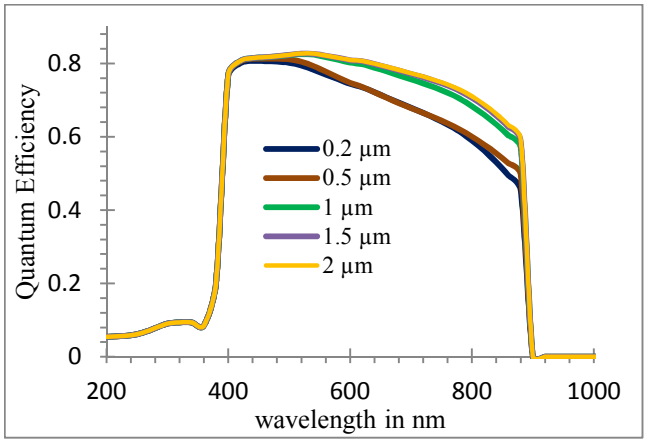


Fig.8. External Quantum Efficiency of CZTS solar cell for different thickness of CZTS

But no visible change has been observed in the quantum efficiency curve for CZTS thickness of more than 1.5µm due to the dominating bulk recombination than the photon generated carrier.

G. Thermal stability of the proposed cell

So as to investigate the stability of the optimized CZTS solar photovoltaic cell with increasing operating temperature, simulations were done in wxAMPS by varying the operating temperature of the cell from 25°C to 125°C and the result is shown in Fig.9.

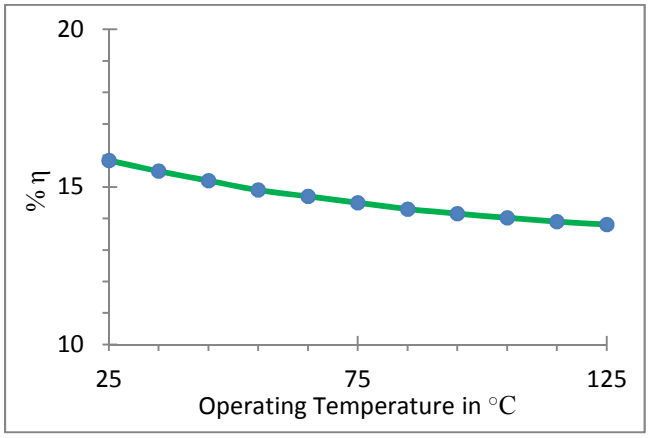


Fig.9. Thermal stability analysis of the optimized CZTS cell

The simulation result points that the optimized CZTS solar photovoltaic cell shows very superior stability with higher operating temperature. The temperature coefficient was calculated as 0.02%/°C from the numerical simulation.

H. Energy band structure of the simulated cell

Energy band structure of the simulated cell as illustrated in Fig.1. is shown in Fig.10.

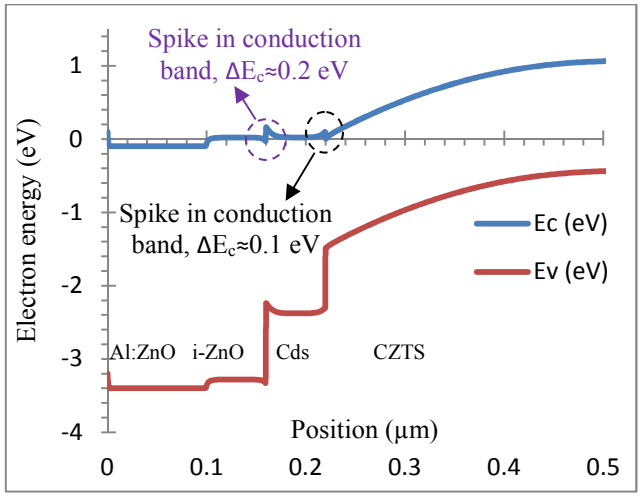


Fig.10. Energy band structure of CZTS solar cell as shown in Fig.1

Positive conduction band discontinuity ($\Delta E_c = \chi_1 - \chi_2$, where χ_1 and χ_2 are electron affinities of CZTS and CdS layer correspondingly) creates a spike in the conduction band in the CdS-CZTS heterointerface and also in CdS-i:ZnO heterointerface as shown in Fig.10. A cliff in the conduction band near absorber layer enhances interface recombination in chalcopyrite based solar photovoltaic cell consists of wide bandgap material [14]. As ΔE_c is 0.1 eV and 0.2eV here and the spike energy more than 0.3 eV is not good for carrier transport, so it doesn't create any obstracle in carrier transport and additionally it creates an electric field which further accelerates carrier transport in the heterointerface.

IV. Conclusion

Basic structure of thin-film CZTS solar cell with CdS buffer layer was modeled and simulated by varying different layer properties. The impact of different layer thickness on conversion efficiency and othe performance parameters were investigated. The dependancy of doping concentration and defect dencity of CZTS absorber layer were also investigated and analyzed. Simulation shows that 15.84 % (*FF*=72.86%, J_{sc}=27.98 mA/cm^2, *and* V_{oc}=0.78 Volt) conversion efficiency has been achieved for the optimized cell. The simulated cell with optimized layer properties presents a better conversion efficiency of 15.84% as compared to the counterparts [4] and the layers as well as the device can be fabricated by different popular deposition techniques like sputtering and chemical bath deposition.

Acknowledgment

The authors might want to recognize Department of EEE and Renewable Energy Laboratory (REL) of Chittagong University of Engineering and Technology (CUET), Chattogram-4349, Bangladesh for supporting this work.

References

[1] Alessio Bosio and Alessandro Romeo, "Thin Film Solar Cells: Current Status and Future Trends",Nova Science Publishers, Inc. New York, ISBN:978-1-53611-363-1.

[2] Hongxia Wang, "Progress in Thin Film Solar Cells Based on Cu_2ZnSnS_4", International Journal of Photoenergy, Volume 2011, Article ID 801292, doi:10.1155/2011/801292

[3] H. Katagiri, K. Saitoh, T. Washio, H. Shinohara, T. Kurumadani, and S. Miyajima, "Development of thin fim solar cell based on Cu_2ZnSnS_4 thin fims," Solar Energy Materials and Solar Cells, vol. 65, no. 1–4, pp. 141–148, 2001.

[4] Martin A. Green, Yoshihiro Hishikawa, Ewan D. Dunlop, Dean H. Levi, Jochen Hohl-Ebinger, Masahiro Yoshita, Anita W.Y. Ho-Baillie, "Solar cell efficiency tables (Version 53)", Prog Photovolt Res Appl. 2019;pp.27:3–12.

[5] William Shockley, Hans J. Queisser, "Detailed Balance Limit of Efficiency of p‐n Junction Solar Cells" Journal of Applied Physics, vol. 32, pp. 510-519,1961

[6] Sven Ruhle, "Tabulated values of the Shockley–Queisser limit for single junction solar cells," Solar Energy 130, pp. 139-147, 2016.

[7] Kentaro Ito and Tatsuo Nakazawa "Electrical and Optical Properties of Stannite-Type Quaternary Semiconductor Thin Films", Japanese Journal of Applied Physics, Volume 27, Part 1, Number 11

[8] Y Y. Liu, Y. Sun, and A. Rockett, "An improved algorithm for solving equations for intra-band tunneling current in heterojunction solar cells," *Thin Solid Films*, vol. 520, no. 15, pp. 4947–4950, 2012.

[9] Y. Liu, D. Heinzel and A. Rockett, "A new solar cell simulator: WxAMPS," Photovoltaic Specialists Conference (PVSC), 2011 37thIEEE, Seattle, WA, pp. 002753-002756,2011.

[10] M. Djinkwi Wanda, S. Ouédraogo, F. Tchoffo, F. Zougmoré and J. M. B. Ndjaka, "Numerical Investigations and Analysis of Cu_2ZnSnS_4 Based Solar Cells by SCAPS-1D," International Journal of Photoenergy Volume 2016, Article ID 2152018, 9 pages.

[11] M.T. Ferdaous, S.A. Shahahmadi, P. Chelvanathan, Md. Akhtaruzzaman, F.H. Alharbi, K. Sopian, S.K. Tiong, N. Amin, "Elucidating the role of interfacial MoS_2 layer in Cu_2ZnSnS_4 thin film solar cells by numerical analysis",Solar Energy 178 (2019), p.p.162–172

[12] W. Zhao, W. Zhou and X. Miao, "Numerical simulation of CZTS thin film solar cell," 2012 7th IEEE International Conference on Nano/Micro Engineered and Molecular Systems (NEMS), Kyoto, 2012, pp. 502-505. doi: 10.1109/NEMS.2012.6196826

[13] Xiaolei Liu, Xiaojing Hao, Shujuan Huang, Gavin Conibeer,"Numerical Modeling of CZTS Solar Cells", 39thIEEE Photovoltaic Specialists Conference, 2013.

[14] K.Wang, O. Gunawan, T. Todorov et al., "Thermally evaporatedCu_2ZnSnS_4 solar cells," *Applied Physics Letters*, vol. 97, no. 14, Article ID: 143508, 2010.

Modelling a Solar Charge Station for Electric Vehicle with Storage Backup

Muhammad Ali
Department of Electrical and Electronic Enineering
Faridpur Engineering College
Faridpur-7800, Bangladesh
muhammad.ali.sayem@gmail.com

Safayat Mohammad
Department of Electrical and Electronic Enineering
Faridpur Engineering College
Faridpur-7800, Bangladesh
safayatmohammad@gmail.com

Md. Minhajur Rahman
Department of Electrical and Electronic Enineering
Faridpur Engineering College
Faridpur-7800, Bangladesh
mimeee01@gmail.com

***Abstract*—The use of Electric Vehicles (EV) as transport is an eco-sustainable and plausible way to protect the reserved energy of developing countries such as Bangladesh. Being a tropical country, Bangladesh is remarkable for its abundance of solar radiation. By using this enormous amount of renewable energy, we can reduce the current power crisis. At present, EVs are running in Bangladesh are auto rickshaw and easy bike. Although these easy bikes and auto rickshaw seem eco - friendly, they are partly responsible for environmental pollution as they use grid power. Under this circumstance, the establishment of Solar Charge Station (SCS) for EVs will be a perfect solution. Because it reduces the additional pressure on the grid by charging directly the EVs and thus saves the environment is polluted.**

Keywords—Solar Charge Station, Electric Vehicles, Grid Power.

I. INTRODUCTION

In the most recent decade, two critical issues developed in the wording of vitality all through the world. The first one is running out of limited petroleum in the near future and the other one is carbon discharge result in a worldwide temperature alteration. Numerous nations at this time rely vigorously on oil, coal and aerosolized gasoline for its energy source. Petroleum derivatives are non-renewable. Non-renewable energy sources are non-sustainable, they draw on restricted assets which will, within the future, diminish, turning into too expensive or too earth harming, creating it not possible to recover. There are many petroleum consumption sectors. Transportation is one of them. As reported by The Energy Information Administration (EIA), around 29% of U.S. energy expenditure in 2017 was for transporting individuals and product from one place to another [1]. At present main transportation vehicles are running by a combustion engine. This combustion engine is 25-30% efficient [2]. In contrast, EV is more efficient up to 85% due to high motor efficiency[3]. According to The National Transportation Research Center (NTRC), World consumes 27706.5 Trillion Btu of energy for transportation and about 6.6 GT CO_2 is emitted [4] whereas EV is environmentally friendly and has a low noise level. Developing countries such as Bangladesh have faced power crises from the last twenty years [5]. Charging these electric vehicles causes an additional load on the national power grid and, as a result, an enormous load shadow experience for the people of the country. It is, therefore, necessary to recharge these EVs from an alternative energy source such as the country's renewable energy resources. Of all renewable energy sources in Bangladesh, solar energy has the highest potential. Therefore, Solar Charge Station (SCS) is a sound solution to this problem.

II. ELECTRIC VEHICLE

Electric vehicles use Electric motors to power their wheels. They drive part or all of the energy from large rechargeable batteries. Electric vehicles are a promising technology to reduce emissions of greenhouse gases and other environmental impacts due to road transport. Each vehicle consumes at best five watts of electricity during off-peak hours to recharge the batteries and 50,000 bikes consume at best 25 megawatts [6]. This causes an additional load on the national power grid. Totally driving an EV from the grid is certainly not a decent way that is followed in Bangladesh. That's why simply it is good thinking to implement solar charge station by using unused place and on the roof fueling station.

III. SITE SELECTION

The choice of the location for an EV charging station is an important factor. A good location can increase both the demand for charging stations and their popularity [7-8]. In this paper, a few criteria are resolved for choosing the area of EV Charge Station Factors of Location Selection [9-10]:

- Level of Charger
- Type of Destination
- Clientele
- Cost of Installation and Operation
- Local Infrastructure and Future Plans
- Engineering Feasibility
- Service Availability
- Geographical Environment

The site meets the above criteria; it will be perfect for solar

charge station. One place can't fulfill above all conditions. Therefore, optimization on the basis of the available location is necessary. For the Charge Station, our selected location is 'Faridpur Engineering College, Faridpur'. Field data were collected from this site and this data may differ slightly from other sites in Bangladesh.

A. Potential of Solar Energy

Bangladesh is located between 20°34′ to 26°38′ north latitude & 88°01′ to 92°41′ east longitude and has a total area of 1.4757×10^8 m^2. On an average of 5 kWh /m^2 solar radiation falls on this country over 300 days per annum. Solar Energy is a pollution free source of renewable energy and it is available everywhere. At present condition, usage of energy in Bangladesh is about 0.15 watt/sq. meter land area, whereas the opportunity is on top of 208 watts/sq. meter [11]. Therefore, this abundant solar energy has a large potential to be used in a various sector in Bangladesh reducing the traditional fossil fuel-based power consumption and ensuring a green environment for the future generation. According to NASA Surface, Meteorology and Energy solar monthly averaged radiation incident (kWh/m^2/day) is shown below throughout the year (January 1984-December 2013) in Faridpur is shown in Table I below [12].

TABLE I. SOLAR RADIATION THROUGHOUT THE YEAR IN FARIDPUR CITY

Month	Average Solar Radiation (kWh/m^2/day)	Month	Average Solar Radiation (kWh/m^2/day)
January	6.18	July	2.58
February	6.15	August	2.77
March	6.08	September	2.93
April	5.58	October	4.42
May	4.49	November	5.74
June	3.03	December	6.47

B. Potential of Wind Energy

Wind turbine systems give a supply of renewable energy. they're most adapted to gusty rural areas. At optimum action, the efficiency of a aerogenerator in converting wind energy to electricity is about 45% [13]. Studies have found that wind energy can play a vital role as an assistance resource for solar energy to generate power for this charge station [14].

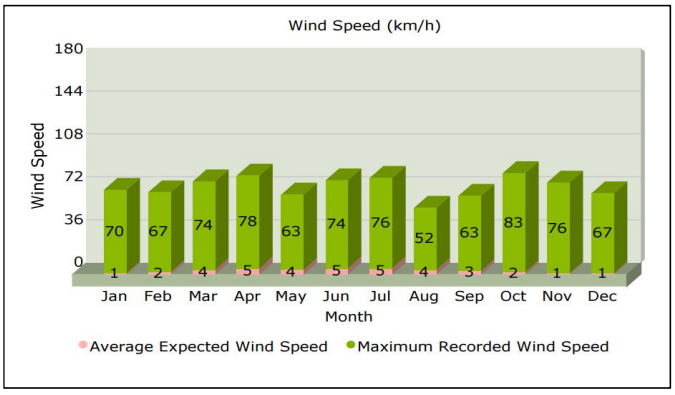


Fig. 1. The Wind Speed (km/h) Data in Faridpur

The potential of wind energy in the coastal areas of Bangladesh is very high but experimentally measured wind speed data are not available over Bangladesh. Wind energy potential in our proposed location is the moderate level. The Wind Speed (km/h) data for various months throughout the year in Faridpur is shown in Fig.1 [15].

IV. METHODOLOGY

PV Panel is the main element of the Solar charge Station. It is connected Series and parallel to equivalent the system voltage by array connection box. To maintain the DC battery charging system and DC load, the MPPT charge controller is used. A 3-phase pure sine wave inverter inverts DC to AC for operating the AC load from main distribution to the sub-distribution board.

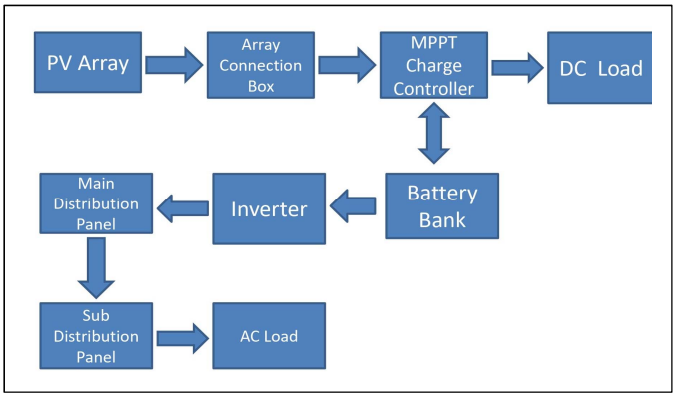


Fig. 2. The Basic Operational Block Diagram

V. ENERGY CONSUMPTION OF ELECTRIC VEHICLE

At present EVs are used in Bangladesh are easy bikes and auto rickshaws. This paper apprizes the approximate energy consumption of an easy bike used in Bangladesh.

Number of Battery used in an easy bike = 5 pcs

The rated voltage of battery = 12 V

The current capacity of battery = 140 Amp-hour

So, the energy stored in each battery

$= \{(140 \times 12) \div 1000\}$ kWh
$= 1.68$ kWh

Now, the total energy consumption of an easy bike

$= (5 \times 1.68)$ kWh
$= 8.4$ kWh

Let the efficiency of solar charge controller is about 90%.

So, charge controller loss = $\{(10 \div 100) \times 8.4\}$ kWh = 0.84 kWh

Therefore, overall energy consumption of an easy bike

$= (8.4 + 0.84)$ kWh
$= 9.24$ kWh

The proposed solar charge station can able to charge a maximum of twenty vehicles at a time. So, the maximum demand of energy consumption is 184.8 kWh. By considering some extra demand, the capacity for reliable operation of the SCS is 200 kWh.

VI. Modelling Of Pv Module

We take a solar PV module, model of SM280-18V with rated power (Pp) of 280W. The average sun hour (ASH) in the selected location is 8 hours. The peak watt Pp is given under standard cell temperature condition (Tc = 25°C), but solar cells are usually 10-30°C hotter than the day time ambient temperature due to solar heating. If the ambient air is 30°C, the cell temperature, Tc could be expected to rise to 40-60 °C [16]. So temperature correction is required. The number of required modules considering the cell temperature (Tc) effect and overall efficiency (η), is then given by [17]

$$N = \frac{Daily\ load\ energy\ demand\ in\ kWh}{Panel\ rated\ power \times AverageSun\ Suning\ Hour}$$

$$= \frac{load\ demand\ in\ kWh}{Pp \times ASH \times [1-(Tc-25)\ 0.5\%] \times \eta}$$

$$= \frac{load\ demand\ in\ kWh}{2.20kWh \times [1-(Tc-25)\ 0.5\%] \times \eta}$$

$$= \frac{200\ kWh}{2.20kWh \times [1-(60-25)\ 0.5\%] \times 0.8}$$

$$= 135.85 \approx 136$$

For the system voltage is 60 V, the required charging voltage is about 70 V. The rated output voltage of the panel is 18 V. The series of 4 solar PV module give 72 V. Hence the number of parallel string = (136 ÷ 4) = 34. So, the total number of solar panel required = 136.

VII. Modelling Of Battery Bank

The number of the battery (N_b) required for a PV system is calculated by considering daily load energy demand, Rated battery voltage (V_b), Battery ampere hour rating (B_{ah}) & overall efficiency of the battery (η_b = 90%) prescribed.

$$N_b = \frac{Daily\ Load\ Enegy\ Demad}{Vb \times Bah \times \eta b} = \frac{200kWh}{12 \times 130 \times 0.9} = 142.15 \approx 145.$$

Then the number of battery required = 145. To maintaining charging system voltage, 5 batteries are connected in series by 29 rows.

VIII. Cost Analysis Of Charging Station

The cost of various components with respect to quantity are given in Table II. The installation, civil work, wiring and maintenance cost is considered approximately. The cost of those components may be varied according to its brand, quality and place.

TABLE II. Components Cost Of SCS System

Components	Model	Quantity	Price (BDT)
Solar Panel	SM280W-18V	136	1713600
Battery	PT-130	145	1740000
Battery Stand	Hand made	5	100000
Solar panel Stand	Hand made	10	100000
Charge Controller	Hand made	8	15000
Inverter	Local Made	1	60000
Wiring	Hand made	-	15000
Civil Work	Approximate	-	100000
Technical Cost	Approximate	-	30000
Capital Cost			3873600

The one-year electricity generation data is calculated by daily and monthly in Table III.

TABLE III. Daily And Monthly Electricity Generation

Month	Sun Daylight (hr.)	Daily Energy Generation (kWh)	Monthly Energy Generation (kWh)
January	8.87	337.77	10470.87
February	9.4	357.95	10022.60
March	10.02	381.56	11828.36
April	10.02	381.56	11446.80
May	11.27	429.16	13303.96
June	11.55	439.82	13194.72
July	10.43	397.18	12312.58
August	10.95	416.98	12926.38
September	10.3	392.22	11766.60
October	9.63	366.48	11360.88
November	9.02	343.48	10304.40
December	8.87	337.77	10470.87
Total Generation (kWh)			139409.02

The annual depreciation value of the charging station is calculated for various components in Table IV.

TABLE IV. Depreciation Cost On SCS Component

Equipment's Name	Investment (BDT)	Life Time (Years)	Depreciation Charge (A/S)	Annual Depreciation (BDT)
Solar Panel	1713600	25	0.55%	94248.8
Charge Controller	15000	5	15 %	22580
Battery	1740000	5	15 %	261000
Civil work	50000	25	0.55%	275
Inverter	60000	5	15 %	9000
Total annual depreciation on equipment				281949.8

The total annual operating cost of charging station is determined in Table V.

TABLE V. Annual Operating Cost Of SCS

Total Cost List	Rate of Quantity	Amount in BDT
Annual labor cost	15000 x 12	180000
Maintenance cost at 1.5% of total capital	0.015×3873600	58104
Annual depreciation	-	281949.8
14% Interest of total capital	0.14×3873600	542304
Total fixed charges		1062357.8

The per unit generation cost of annual operation is determined in Table VI.

TABLE VI. Cost Of Per Unit Generation

Total Cost List	Amount of Parameters
(a) Total annual fixed charges	1062357.8 BDT
(b) Energy generation per year	139409.02kWh
Cost of generation (a/b)	7.62 BDT/Unit

IX. Payback Time Calculation

After a fixed time, Cost will be recovered from this project. So, the payback time is very important for any cost implementation project. The payback time is calculated to the capital cost of a system divided by annual cash outcome.

$$\text{Payback Time} = \frac{Capital\ Cost\ of\ Project}{Annual\ Cash\ outcome}$$

If we consider the five-taka unit as electricity price for this charging system-

Then the total annual income = (5 × 39409.02) BDT
= 697045.1 BDT

The system total capital cost = 3873600 BDT

Payback Time = (3873600 ÷ 697045.1) = 5.5 Years

X. Recent Built Solar Charging sation In Bangladesh

With the monetary help of IDCOL and World Bank, Solar E-Technology has started on a tryout project at Chuadanga and organized an exhaustive R & D workout on a solar powered battery charging station as another choice to charge batteries for easy bikes [18].



Fig. 3. Solar Powered Battery Charging Station R & D site at Chuadanga

XI. Conclusion

Day by day, the number of EVs increased rapidly in Bangladesh. It is good news for EVs future but it creates additional demand on the grid & partially liable for power crisis. Although these EVs seem eco - friendly, they are partly responsible for environmental pollution as they use grid power. Because this grid power is generated by natural gas, oil, coal etc. At this moment, if the proposed SCS can full fill the current demand of EVs, it will get its popularity. Then it will be easy to develop a project system model and reduce generation cost.

Acknowledgment

This thesis has been assisted by the Department of Electrical and Electronic Engineering (EEE) and Power & Solar Energy Laboratory, Faridpur Engineering College (FEC), Faridpur-7800, Bangladesh.

References

[1] "Use of Energy in the United Sates Explained: Energy uses for transportation," Online Available: "https://www.eia.gov/energyexplained /? page=us_energy_transportation" (4th September 2018)

[2] Shyam K. Agrawal "Internal Combustion Engines", Published by New Age International, 2006

[3] Jack Rosenblatt, M. Harold Friedman "Direct and Alternating Current Machinery", illustrated 2nd Ed., Published by C.E. Merrill Publishing Company, 1984

[4] "Transportation Energy Data Book," Online Available: 'http://cta.ornl.gov /data/chapter2.shtml" (10th September 2018)

[5] "Electric Rickshaws Run out of Steam,'' The Daily Star. Online Available: "https://www.thedailystar.net/news-detail-187825" (10th September 2018)

[6] "Key Statistics, Electric vehicle power consumption as on June, 2015, Bangladesh Power Development Board," Available: "http://www.bpdb.gov.bd" (26th August 2015).

[7] Kobayashi, Y, Kiyama, N, Aoshima, H.; Kashiyama, M. "A route search method for electric vehicles in consideration of range and locations of charging stations", Intelligent Vehicles Symposium (IV), 2011 IEEE, 5-9 June 2011, DOI: 10.1109/IVS.2011.5940556

[8] Md. SohelRana, Fahim Hossain, ShuvangkorShusmoy Roy, Suman Kumar Mitra "Energy consumption by battery operated Auto rickshaws" Daffodil international university journal of science and technology, volume 8, issue 2, July 2013.

[9] "How to Choose the Best Locations for Public EV Charging Stations?" Online Available: "https://www.fleetcarma.com/ev-charging-stations-choosing-best-locations/". (26th August 2015)

[10] Yunna Wu, Meng Yang, Haobo Zhang, Kaifeng Chen and Yang Wang "Optimal Site Selection of Electric Vehicle Charging Stations Based on a Cloud Model and the PROMETHEE Method" Energies 2016, 9(3), 157; https://doi.org/10.3390/en9030157

[11] Eusuf, M. 1997. "Prospect and problem of Solar Energy in Bangladesh: Implementation stage of solar systems", Bangladesh Centre for Advanced Studies, Dhanmondi, Dhaka - 1209, Bangladesh

[12] "NASA Surface meteorology and Solar Energy Database," Online Available: "https://power.larc.nasa.gov/" (27th August 2018)

[13] Vaughn Nelson "Wind Energy: Renewable Energy and the Environment" Second Edition (illustrated, revised), Published by CRC Press, 2013

[14] "Wind turbine systems and Renewable energy," Online Available: "http://www.level.org.nz/" (20th August 2018).

[15] "Climate History for Faridpur _ Local _ Bangladesh," Online Available: "http://www.myweather2.com/" (22th August 2018).

[16] Tomas Markvart, K. Bogus, "Solar Electricity" Illustrated Edition. Published by John Wiley & Sons, 2000

[17] Nicholas Jenkins, Janaka Ekanayake "Renewable Energy Engineering" Illustrated Edition, Published by Cambridge University Press, 2017, ISBN- 9781107028487.

[18] "Solar powered battery charging station: A pilot R & D site at Chuadanga sponsored by IDCOL & World bank" Solar E Technology, June'2016

Comparison of Different Hybrid Renewable Energy Systems With Optimized PV Configuration to Realize the Effects of Multiple Schemes

Faishal Islam Bappy
Department of Electrical & Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
faishalbappy61@gmail.com

Md Jahirul Islam
Department of Electrical & Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
jahirul.kuet@gmail.com

Amit Kumer Podder
Department of Electrical & Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
amitpodder09@yahoo.com

Debopriya Roy Dipta
Department of Electrical & Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
roydebopriya54@gmail.com

Hossain Mansur Resalat Faruque
Department of Electrical & Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
resalat_f@yahoo.com

Eklas Hossain
Department of Electrical Engineering and Renewable Energy
Oregon Institute of Technology
Oregon, USA
eklas.hossain@oit.edu

Abstract— **Hybrid renewable energy systems (HRESs) have a great potentiality to provide a more reliable power supply, when compared to a system based on a standalone source. The hybrid system supports the modular design that can mitigate problems associated with the use of individual sources by combining them all in a systematic procedure. In this paper, comparative study for analyzing the effects of four different systems namely: photovoltaic (PV) stand-alone with storage unit, PV-battery-grid system, PV-diesel-battery system, and PV-diesel-grid system are carried out using the HOMER software. Furthermore, an optimal cost analysis is also conducted based on net present cost (NPC). In addition, using real satellite footage with the help of HELIOSCOPE software, an effective configuration of PV array is designed based on real solar irradiance data. The cost optimization of the HRES is developed on the basis of the sets of real data collected from Khulna University of Engineering and Technology (KUET), Khulna, Bangladesh. The study reveals that the approach with PV-diesel-battery system has comparatively lower NPC, less fuel cost, and a limited carbon footprint along with excess electricity to contribute to the national grid.**

Keywords— HOMER, NPC, HELIOSCOPE, Micro-grid, Hybrid Renewable Energy System.

I. INTRODUCTION

The limited availability of fuel and a gradual increase in its cost has made it necessary to utilize renewable energy for electrical power production [1]. Being inexhaustible in nature and also for overcoming environmental constraints, renewable energies are getting much more interest. Renewable energy system (RES) also provides flexibility of being complimentary, as the integration between them shows tremendous promises [2]. Despite the availability of solar photovoltaic and wind energy, a PV or a wind generator system acting as stand-alone, cannot meet-up the 24 hours demands [3]. This is the reason why research on HRES has come in vigilant.

A HRES usually consists of multiple renewable energy sources carried out together to provide sustainable energy production as well as economic efficiency [4-7]. Additionally, it can also offer extra power to the national grid after meeting up self- demands. But for designing an effective HRES, a lot of factors need to be considered based on load analysis, geographical location as well as weather conditions [8]. Many researchers have already investigated the HRES system model based on different circumstances[9-10]. Economic cost analysis of PV array and wind generator based HRES system for remotely located village has been proposed earlier [11]. But, for areas or institutions such as KUET, which has more sophisticated load demand and also would like to approach toward self-sufficiency in regard to electricity; the implementation of HRES has not yet been studied.

The main purpose of this study is to design, simulate and optimize a system to meet the electrical load demand of KUET premises. In terms of system reliability, economy and environmental issues related to decentralized electrification, evaluations are made to overcome the existing constraints and finding out the best-suited hybrid system configuration. By considering the geographical position of KUET and the existing climate conditions, solar energy has been selected as a promising form of source for power production. Along with it, battery storage system and diesel generator have also been considered for designing the HRES system. Four possible case studies have been analyzed to determining the most effective combination using HOMER software. Additionally, using HELIOSCOPE—another advanced solar design software—

different configurations of PV array on different buildings of KUET have been simulated and analyzed by feeding real solar Irradiance data for that location. By considering the effectiveness and cost-efficiency, the more suitable PV lay-out has been proposed. The protocol that has been proposed here can also be applicable for other institutions and also in industrial premises having the similar kind of load profiles.

The rest part of the paper has been organized as follows: in section II, system modeling along with case study analysis has been provided. Detailed PV array configuration with necessary mathematical illustration has been demonstrated in section III. The later section explains the economic cost analysis for required optimization. Simulation results of the proposed system have been presented in section V. This section also describes the numerical comparison of different cases based on the system parameters. Finally, in section VI, the key features along with brief details have been described as a conclusion.

II. System Modeling and Case Study Analysis

In this paper, systematic model design of a renewable hybrid energy system for KUET campus is configured and optimized using the HOMER software. Alternative system configurations based on different other well-known systems in conjunction with the solar system are also modeled and compared. In general, four different configurations are proposed entitled with case number and thereafter evaluation and comparison are made with some fundamental criteria such as energy efficiency, financial viability, and environmental impact. In Fig. 1, the grid monthly input data has been shown which have been collected from KUET substation. The average demand of KUET in kWh/d is 4907.6 and the peak demand is 967 kW. To analyze and find out the optimal solution using NPC in HOMER software these grid monthly input data has been used. The other input parameters that have been considered for sizing cost and other parameters are standardized from National Renewable Energy Laboratory (NREL) [12]. In Table I, the standard capital costs, replacement costs, and operational and maintenance costs for different units are listed [13]. For all the systematic analysis same PV and diesel parameters are used in all four different cases whenever required.

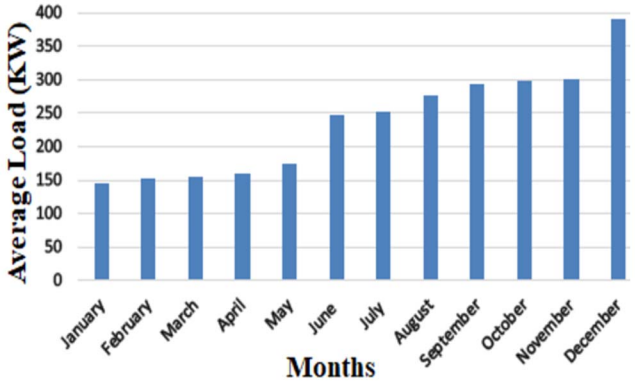


Fig. 1. Monthly load profile of KUET

TABLE I. Input Data for Cost Analysis

Units	Capital Cost	Replacement Cost	Operational & Maintenance Cost	Expected Life time
Diesel Generator	$125/kW	$125/kW	50.15/Hour	9.60y
Battery	$75/Battery	$75/Battery	$2/Battery/Year	510kWh
PV	$7500/kW	$7500/kW	0	25y
Converter	$1000/kW	$1000/kW	$100/Year	15y

In case 1, to meet up the electrical demand of KUET, two solar micro-grids rated at 630 kW and 680 kW along with associated converters and storage units have been considered. As a storage unit 1108 kW battery and 1000 kW converter has been adopted. The hours of operation considered for the converter, having a capacity factor of 23.2%, is 8760 h/year, whereas for PV array it is 4371 h/year. Individual contribution of the two solar micro-grid is to produce 160,612 kWh/year, nevertheless case 1 faces a capacity shortage of around 1448 kWh/year, making it less compatible as a PV stand-alone system. On the other hand, in Case 2, the configuration is same as of Case 1 with an addition of grid connection. In this case the energy is exchanged over the grid i.e. energy is taken from grid when required and also sale to the grid when surplus energy is produced. The purchase and sellback price for this case has been considered as $.10 /kWh and $.05/kWh respectively. The system is designed to feed electricity from the grid during night when the PV is off. The two PV micro-grids produce 7.65% of power separately per year, whereas the rest 84.7% power is purchased from grid.

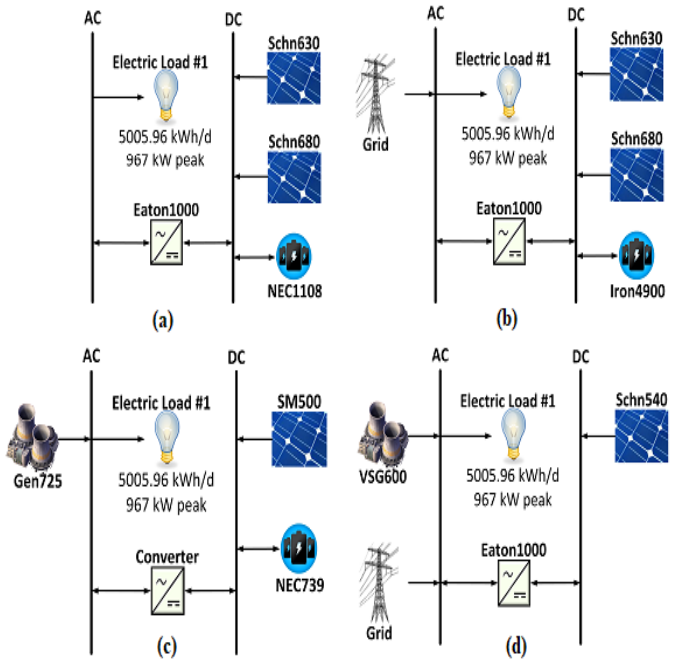


Fig. 2. Proposed hybrid system configurations. (a) PV only with storage system, (b) PV with grid and storage system, (c) PV, diesel generator with storage system, and (d) PV, diesel and grid system

For Case 3, a 725 kW diesel generator along with 500 kW solar micro-grid, 739kW battery storage system, and 1000kW converter are used to mitigate the same load profile. When the solar system is off, the load demand is supplied from diesel generator and storage units. The PV array generates 46.4% of power annually, while the remaining 53.6% production is met by the diesel generator. Lastly, a 600 kW diesel generator, a 540 kW solar micro-grid, the grid connection, and DC/AC converter have been used for Case 4 in the same load profile of the existing electric system of KUET. In this model, 72.1% power per year is required to purchase from grid, the rest are generated by PV cells (8.75%) and diesel generator (19.2%). The model configurations (i.e. cases) are labeled as (a) to (d) and are shown in Fig. 2, by using generic HOMER components. However, it is noteworthy that for all the cases, a 6 percent annual interest rate and built-in constraints have been considered.

III. PV Array Configuration

The solar resource input data for HOMER is made up of monthly averaged daily insolation incidents on a horizontal surface (kWh/m^2/day) from the NASA Surface Meteorology and Solar Energy website [14]. It is noted that NASA provides monthly averaged values from 22 years of data. The standard solar radiation data for the location of KUET ($22°54.0'N, 89°30.1'E$) is shown in Fig. 3. As shown in the figure, the solar radiation is dependent on the day time and pattern of cloud in the geographical area like Bangladesh. HOMER evaluates PV array power for the year on an hourly basis, and uses latitude value to calculate the average daily radiation from the clearness index and vice versa. The annual averaged daily solar insolation in this area was found to be 4.56 kWh/m^2/day.

To find out a sustainable and inoffensive appearance to the beauty of KUET green campus, the proposed solar panels must be arranged thoughtfully. Therefore, careful designs were made to utilize the rooftops of different building of KUET campus. For the 1st and 2nd cases, a 1.3 MW PV array design configurations, and for the 3rd and 4th cases, another 655.2 kW PV array design configuration is required. A specialized software 'HELIOSCOPE' has been employed for this purpose. The HELIOSCOPE design of both the 1.3 MW and 655.2 kW PV array configurations has been illustrated in Fig. 4 and Fig. 5 respectively.

The calculation of output power provided by installed PV array has been determined by utilizing the following equations [15]:

$$T_c = T_a + g(t) \times \left(\frac{T_N - 20}{0.8}\right) \quad (1)$$

$$I = g(t) \times \left(I_{sc} + K_c \times (T_c - 25)\right) \quad (2)$$

$$V = (V_{oc} + K_v \times T_c) \quad (3)$$

$$FF = \frac{V_p \times I_p}{V_{oc} \times I_{sc}} \quad (4)$$

$$P_{pv}(g(t)) = n_{cell} \times FF \times V \times I \quad (5)$$

The above-mentioned parameters are briefly summarized on Table II.

TABLE II. SUMMARY OF DIFFERENT PARAMETERS.

Parameter	Meaning
T_c	cell temperature [°C] of the PV panel
T_a	ambient temperature [°C]
$g(t)$	random irradiance
T_N	nominal temperature for the operation of the PV cell [°C]
V_{oc}	open-circuit voltage [V]
I_{sc}	short-circuit current [A]
K_c	current temperature coefficient [mA/°C]
K_v	voltage temperature coefficient [mV/°C]
V_p	voltage at peak power [V]
I_p	current at peak power [A]
FF	Fill factor
n_{cell}	number of PV cells
V	terminal voltage [V]
I	PV cell current [A]
P_{pv}	PV power output [W]

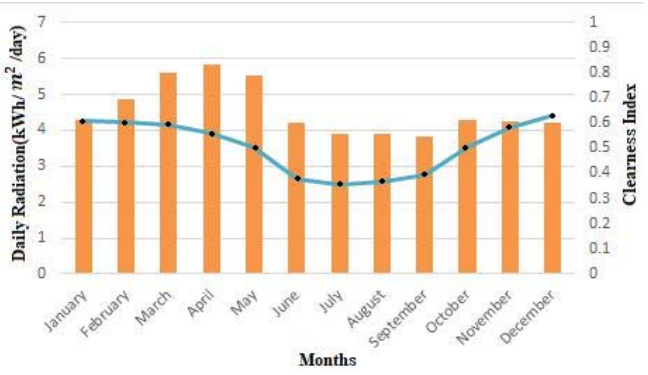


Fig. 3. Yearly solar irradiance profile at KUET.

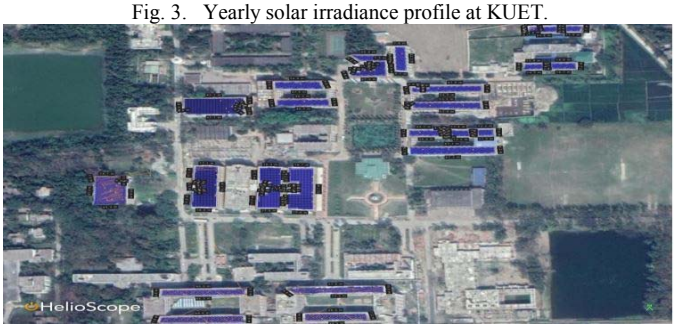


Fig. 4. PV panel installation for cases 1 and 2.

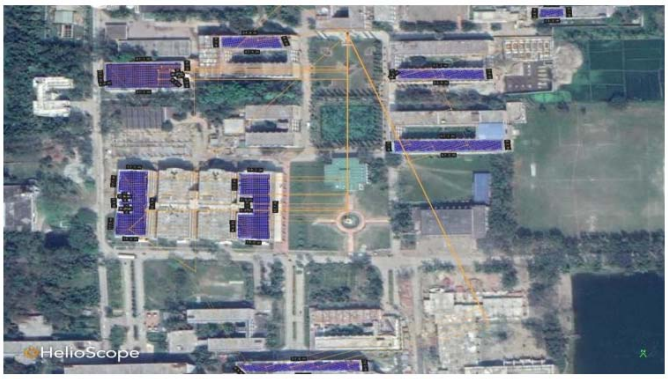


Fig. 5. PV panel installation for cases 3 and 4.

IV. Economic Modeling With Optimization

For optimization, one of the important parameters—low NPC—has been considered for attaining economic feasibility. The value of NPC is determined from all the costs and revenues of the life-cycle of the system. The calculations also accounted for the initial capital cost, maintenance cost, component replacement and fuel costs using the discount rate to calculate the future cash-flows.

The NPC is calculated using the following mathematical relation,

$$NPC = \sum_{t=0}^{T} C_{ct} + C_{om,t} + C_{r,t} + C_{f,t} - R_{s,t} \qquad (6)$$

Here, T is the lifetime of the project, C_{ct} represents the capital cost for the system in year t, $C_{om,t}$ represents the cost of operation and maintenance in year t, the replacement cost at year t is $C_{r,t}$. The other parameters, $C_{f,t}$ is the cost of fuels in year t, and $R_{s,t}$ represents the salvage price in same time period *t*. HOMER performs time series simulation on hourly basis for every possible system configurations, and thereafter finds the optimal solution where the system meets the total energy demand and achieves the least possible NPC. The optimization also provides a list of configurations sorted by the Cost of Electricity (COE) for each of the optimized systems and offers the opportunity to prioritize the system on COE for evaluating the financial feasibility [17]. The COE represents the average production cost of each kWh useful electrical energy produced by the system. COE is basically a ratio of annualized cost of producing electricity to the total useful electric energy production. The relationship of these terms are given as,

$$COE = \frac{C_{tot}}{E_{PAC}+E_{PDC}+E_d+E_g} \qquad (7)$$

Where C_{tot} is the aggregative amount of the cost of investing is, installing, operating, and maintaining the energy system, E_{PAC} represents the AC electricity demand of the primary load, E_{PDC} is the DC electricity demand of the primary load, E_d denotes the electricity demand of the deferrable load, and E_g represents the electricity sold to the grid.

Another aspect of optimization is to consider the environmental effect for the four cases. As for case 1 and case 2, the produced power mainly depend on solar power, therefore, no CO_2 emission occurs in the process. But for case 3 and case 4 diesel generators are mainly responsible for producing CO_2 emissions. For determining the rate of emission of CO_2, the fundamental step is to evaluate the specific fuel consumption rate for two different cases. The specific fuel consumption (F_{sfc}) can be determined in the following manner [13]:

$$F_{sfc} = \frac{F_T}{E_g} \qquad (8)$$

Where, F_T is the total annual fuel consumption [L/year] and E_g represents the annual electrical production [kWh/year], of the generator. Finally, the emission of CO_2 due to fuel consumption can be calculated by taking benefit of the value of specific fuel consumption as shown below [16]:

$$E_{CO_2} = F_{sfc} \times EF \qquad (9)$$

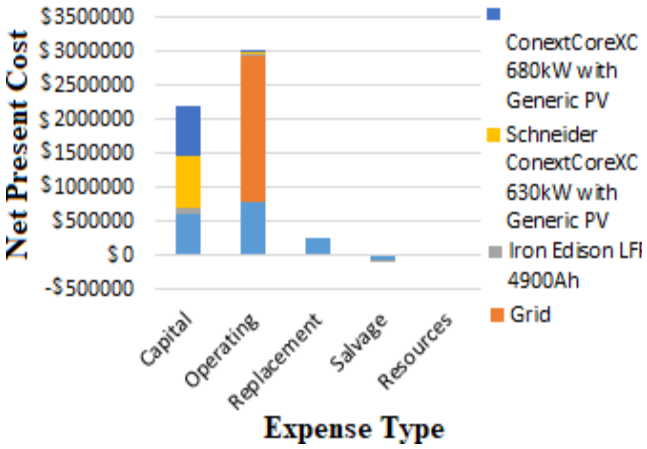
Where, E_{CO_2}: represents the emission of CO_2 [kg/kWh],
EF: Emission factor [kg/L],
F_{sfc}: Specific fuel consumption [kg/kWh].
It is notable to mention that, the emission factor for diesel fuel has been considered as 2.66 kg/L in this paper for calculation.

V. Simulation Results and Discussions

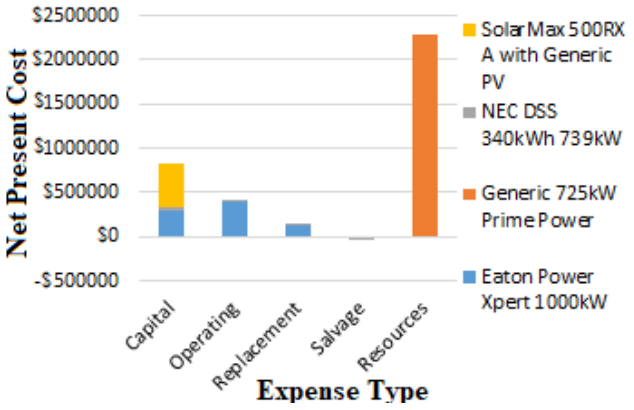
To find out a comprehensive comparison among the proposed cases systematic investigations were carried out. The analysis includes the basic parameters such as NPC, CO_2 emission, and other parameters related to the cost are calculated using HOMER. On the other hand, the system loss due to the fact such as shading, inverter, converter, reflection etc. has been simulated and calculated using HELIOSCOPE. A detailed analysis of the NPC profile is shown in Fig. 6 separately for all the four cases. As it is shown in the Fig. 6(a), the capital cost is much higher than other expenses for case 1, where only two solar micro-grids were used. On contrast, operating costs for 4 is very high due to the running cost of diesel engine. However, most resources with the lowest capital cost are required for case 3 which combined solar, diesel and storage systems.



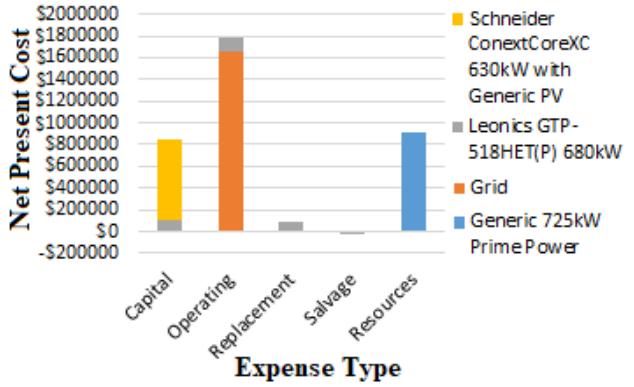
(a)



(b)



(c)



(d)

Fig. 6. Simulation results of net present cost (NPC) for different expense types: (a) case 1, (b) case 2, (c) case 3, and (4) case 4.

A detail comparison of the costs for the four individual cases is summarized in Table III. The simulation reveals that the case 3 involves an initial cost of $0.83M which is the lowest initial one amongst the four cases. Since the system totally depends on renewable resources in Case 1, the involvement of large number of storage unit i.e. battery led to a higher initial cost. As smaller solar micro-grid integration is considered in Case 3, the initial cost has not been increased significantly. Among the four cases, the fuel cost is involved with the systems with diesel generators. When the renewable fraction increases the fuel cost reduces. Neither the fuel cost nor CO_2 emission is associated with the solar PV only micro-grid configurations. One of the noteworthy features of the analysis is that, despite of having no fuel cost and negligible maintenance cost for the solar PV only micro-grid, it has the highest NPC compared to the other cases mainly due to the high initial cost. On the contrary, the hybrid configurations in case 3 and case 4 have much lower NPC than the other two. The PV with diesel micro-grid configuration (i.e. case 3) offers least initial investment along with moderate CO_2 emission of 1395.70 lb/MWh. According to final "Carbon Pollution Standard for New Power Plants", Environmental Protection Agency (EPA) has confirmed a limit of the emission of CO_2 gas for renewable and fuel-based power plants [18]. For renewable power plants the limit is no more than 1000 pounds (lbs) of CO_2 emission per MWh of electricity produced, whereas for fuel-based power plants, the restriction is within 1400 lbs/MWh [16]. By considering this fact, case 3 ensures CO_2 emission to be within this limit. Though case 4 offers the least cost of energy ($0.157) among the four cases, its CO_2 emission profile (1542.31 lb. /MWh) is quite above the aforementioned boundary. Furthermore, in terms of excess electricity case 4 has also been proved to be lagged behind comparatively, making it a less suitable system to prefer. On the contrary, case 3 produces the second highest excess electricity of 4142 KWh/year when compared to the other cases and has a higher PV penetration rate of 44.9%. Although case 2 ensures the highest excess electricity among these four, its high NPC, initial cost as well as cost of energy has made it less appropriate.

TABLE III. SUMMARIZED RESULTS OBTAINED BY HOMER

Item	Case 1	Case 2	Case 3	Case 4
Initial Cost($)	3.30M	2.21M	825625	850125
Cost of Energy ($)	0.216	0.203	0.157	0.153
NPC($)	5.11M	5.39M	3.62M	3.63M
CO_2 Emission (lb. / MWh)	0	0	1395.70	1542.31
Excess Electricity (kWh/year)	0	5888	4142	13.30
PV Penetration	8.79%	8.79%	44.9%	8.99%

For the optimization of the system, different types of system losses must be included in the analysis. A number of sources such as reflection, shading, soiling, irradiance, inverter, clipping etc. are mainly responsible for the losses. Fig. 7 clearly depict the losses in a pie chart. For all the cases, temperature rise results in most of the system losses. The other sources like AC system, reflection, and converter account for moderate losses. As it is shown in the figures, for case 1 and case 2, the collective loss is 31% [see Fig. 7(a)]. On the other hand, for case 3 and case 4, these losses account for 30.1% [see Fig. 7(b)].
By considering all above-mentioned fact, it can be clearly depicted that PV-diesel based micro-grid system (Case 3) offers the most balanced profile out of this four and has been proposed in this paper as the most suitable system to be adopted.

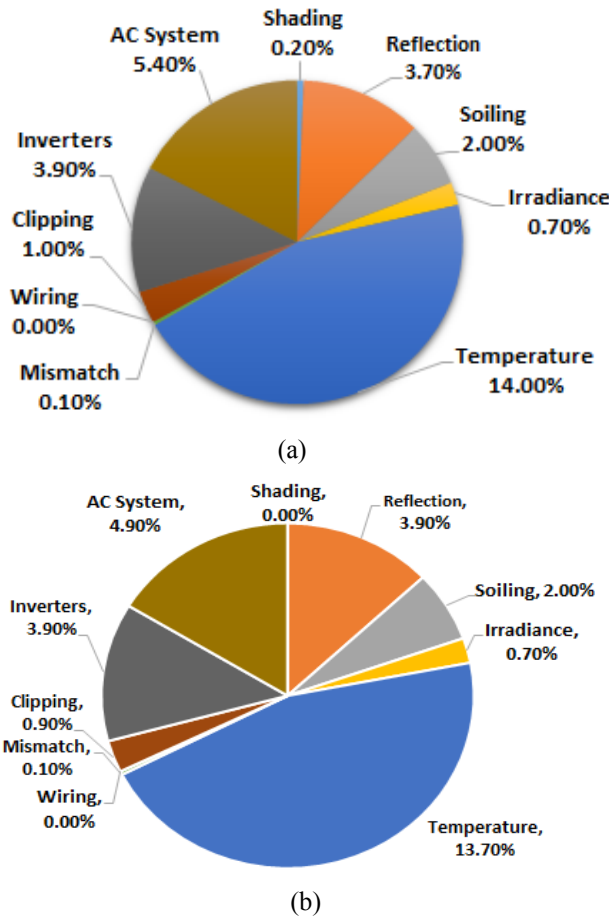


Fig. 7. Sources of different system losses. (a) For case 1 and case 2, (b) for case 3 and case 4.

VI. Conclusion

In this paper, we have systematically designed four combinations of HRESs for successfully meeting the electrical energy demand of a sophisticated load profiling area such as KUET. The combinations have been designed using the existing grid supply, solar micro-grid units, diesel-electric conversion system (i.e. generator), and storage battery system to optimize and find out a suitable hybrid system for KUET. The combinations have been entitled with case numbers and these include solar PV only with storage system (case 1); PV with storage and grid system (case 2); PV with diesel generator system (case 3); and PV with grid and diesel generator system (case 4). Simulations have been carried out for the aforementioned cases to find out the suitable and optimized combination using HOMER and HELIOSCOPE software. Parameters such as NPC, per unit cost of energy, CO_2 emission, initial cost etc. have been considered for the choice of suitable HRES. It was found that PV-diesel based micro-grid system i.e. case 3, proved to bet the most suitable choice having comparatively lower NPC ($3.62M), lower initial cost ($0.83M), lower system losses, moderate CO_2 emission along with better excess electricity profile (4142kWh/year). Additionally, the real solar irradiance data has been taken into account and most effective and cost-efficient PV array-lay out have been proposed. The protocol proposed by this study can also be applicable to areas or institutions having similar kind of load profile.

References

[1] M. Hossain, S. Mekhilef, and L. Olatomiwa, "Performance evaluation of a stand-alone PV-wind-diesel-battery hybrid system feasible for a large resort center in South China Sea, Malaysia," Sustain. Cities and Soc., vol 28, pp. 358-366, January 2017.

[2] M. A. M. Ramli, A. Hiendro, and Y. A. Al-Turki, "Techno-economic energy analysis of wind/solar hybrid system: Case study for western coastal area of Saudi Arabia," Renewable Energy, vol. 91, pp. 374-385, June 2016.

[3] J. Hazelton, A. Bruce, and I. MacGill, "A review of the potential benefits and risks of photovoltaic hybrid mini-grid systems," Renewable Energy, vol 67, pp. 222-229, July 2014.

[4] M. K. Deshmukh, and S. S. Deshmukh, "Modeling of hybrid renewable energy systems," Renewable and Sustain. Ener. Rev., vol 12, pp. 235-249, January 2008.

[5] M. Fadaee, and M. A. M. Radzi, "Multi-objective optimization of a stand-alone hybrid renewable energy system by using evolutionary algorithms: A review," Renewable and Sustain. Ener. Rev., vol 16, pp. 3364-3369, June 2012.

[6] A. K. Podder, M. R. Hasan, N. K. Roy, and M. M. R. Komol, "Economic Analysis of a Grid Connected PV Systems: A Case Study in Khulna," European Journal of Engineering Research and Science, vol. 3, pp. 16-21, 2018.

[7] G. Bekele, and B. Palm, "Feasibility study for a standalone solar-wind-based hybrid energy system for application in Ethiopia," App. Ener., vol. 87, pp. 487-495, February 2010.

[8] G. Bekele, and B. Palm, "Feasibility study for a standalone solar-wind-based hybrid energy system for application in Ethiopia," App. Ener., vol. 87, pp. 487-495, February 2010.

[9] E. Kabalci, R. Bayindir and E. Hossain, "Hybrid microgrid testbed involving wind/solar/fuel cell plants: A desing and analysis testbed," 2014 International Conference on Renewable Energy Research and Application (ICRERA), Milwaukee, WI, 2014, pp. 880-885.

[10] E. Kabalci, E. Hossain and R. Bayindir, "Microgrid test-bed design with renewable energy sources," 2014 16th International Power Electronics and Motion Control Conference and Exposition, Antalya, 2014, pp. 907-911.

[11] U. Sureshkumar, P. S. Manoharan, and A. P. S. Ramalakshmi, "Economic cost analysis of hybrid renewable energy system using HOMER," in Proc. Of the IEEE-International Conference on Advances in Engineering, Science and Management (ICAESM -2012), 30-31 March 2012, Nagapattinam, Tamil Nadu, India.

[12] Energy Systems Analysis Tools. [Online]. Available: https://www.nrel.gov/analysis/energy-systems-tools.html. [Accessed Dec. 26, 2019].

[13] Energy Systems Analysis Tools. [Online]. Available: www.nrel.gov/homer/. [Accessed Dec. 29, 2019].

[14] NASA open data portal. [Online]. Available: https://data.nasa.gov/Earth-Science/Surface-Meteorology-and-Solar-Energy/wn3p-qsan. [Accessed Oct. 11, 2019].

[15] T. Adefarati, R. C. Bansal and J. John Justo, "Techno-economic analysis of a PV–wind–battery–diesel standalone power system in a remote area," in The Journal of Engineering, vol. 2017, no. 13, pp. 740-744, 2017.

[16] TWDC Revised 2006 GHG Inventory Report. [online]. Available: http://www.greenproductionguide.com/wp-content/uploads/2014/11/Production-Emission-Calculator-Methodology-and-ReferencesRev.pdf. [Accessed Jan 26, 2019].

[17] D. B. Nelson, M. H. Nehrir, and C. Wang, "Unit sizing and cost analysis of stand-alone hybrid wind/PV/fuel cell power generation systems," in Proc. of the IEEE Power Engineering Society General Meeting, 16-16 June 2005, San Francisco, CA, USA.

[18] Regulating Power Sector Carbon Emissions. [Online]. Available: https://www.c2es.org/content/regulating-power-sector-carbon-emissions/. [Accessed Feb. 10, 2019].

Thickness Dependency of Zinc Selenide (ZnSe) Thin Film Deposited By Vacuum Evaporation Method

Md. Abu Sayeed[1], Hasan Khaled Rouf[1], Mohd. Rezaul Hasan[1], K.M.A. Hussain[2]
[1]*Department of Electrical & Electronic Engineering, University of Chittagong, Chittagong-4331, Bangladesh*
[2]*Experimental Physics Division, Atomic Energy Center, Dhaka-1000, Bangladesh*
Email: [1]mdasrasel29@gmail.com, [1]hasan.rouf@cu.ac.bd, [1]rezaulhasan1369@gmail.com, [2]kazi326@yahoo.com

***Abstract*—Fabrication of Zinc Selenide (ZnSe) thin film having different thickness has been performed on glass substrate using thermal vacuum evaporation technique by evaporating ZnSe powder and is further characterized by X-ray diffractometer (XRD) and UV-VIS-NIR spectrophotometer. XRD patterns revealed that ZnSe thin film has face centered cubic (fcc) structure with variable miller indices at maximum X-ray intensity as the thickness changes. Substrate temperature is fixed at 290°C and all films are annealed for 60 minutes at 300°C. We have got peak value of intensity in XRD patterns which drops as the thickness of ZnSe thin film reduces. Optical properties like transmittance, refractive index, energy bandgap and reflectance have also been extracted in both simulation and experiment and found promising results. Transmittance in near-infrared region for all the films are greater than 63 % in simulation and fewer less in experiment, variable reflectance of (7-31) %, optical bandgap from 2.44 eV to 2.64 eV and fluctuation in refractive index from 2.31 to 2.70 as the film thickness changes. Optical bandgap goes wider for higher thickness films and gets lower as the film thickness decreases.**

Keywords—ZnSe thin film, thermal vacuum evaporation, X-ray diffraction, UV-VIS-NIR spectrophotometer, transmittance

I. INTRODUCTION

Thin film refers to a layer having infinitesimal physical thickness and which is formed by depositing a simple or compound materials. This thin film technology has wider applications in optoelectronic devices like LED, transistor, solar cell and storage [1].

As a compound semiconductor, Zinc Selenide (ZnSe) has intrinsic and wide bandgap properties which is comprised of group IIB element Zinc (Zn) and group VIA element Selenium (Se). Energy-momentum (E-k) diagram of Zinc Selenide (ZnSe) shows that maximum of valence band aligns with the minimum of conduction which makes ZnSe as a direct bandgap semiconductor [2].

Researchers had performed many researches on thin film due to its potential applications particularly in the field of thin film solar cells, large area electronic devices like detection of radiator, LED, field effect transistor, non-linear integrated optical devices and LASER heterostructure to improve the performance of these devices.

These works are: J. Wang and J.B. Neaton studied the effect of thickness dependent magnetism to the polarization of its lattice parameters preformed on heterostructure $BiFeO_3$ multiferroic epitaxial thin film [3]. Another investigation was done by U. Helmersson and PO Kall towards the optical behavior of TiO_2 thin films spin coated on Si substrate which were annealed in air at different temperatures and they found a homogenous and carbon free TiO_2 films with high refractive index profile [4]. Again M.W.J. Prins and R.M. Wolf discovered the consequence of transparent oxidic thin films to the operation of field effect transistor and obtained an intrinsic memory function because of these films of ferroelectric insulator [5].

Like others, H.H. Gullu and M. Parlak had scrutinized the effect of post annealing process of Zinc Selenide (ZnSe) thin film under Nitrogen atmosphere and got the deposited films nearly stoichiometric with increase in Zn and decrease in Se contents [6]. Partial similarity of our proposed work had been performed by SJ Ikhmayies and RN Ahmed on fabrication and characterization of CdTe thin film deposited by vacuum evaporation at ambient temperature and they found CdTe thin film possessing polycrystalline Zinc Blend structure also with uniform lattice structure [7]. These shortly described works were on different thin films and their properties at different external phenomena like annealing temperature, fabrication techniques and magnetic effects.

Again, compound semiconductor such as Cadmium Sulfide (CdS) [8], Cadmium Selenide (CdSe) [9] have lower transmittance and also because of their surface roughness, these two materials are not so much effective in several applications. In our research, we have thoroughly studied the response of ZnSe thin film at several thicknesses. We have found relatively higher transmittance, small bandgap and less reflectance for the ZnSe thin films having lower thickness and vice-versa.

II. EXPERIMENTAL DETAILS

Zinc Selenide (ZnSe) thin film has been fabricated by thermal vacuum evaporation technique evaporating 99.99 % pure ZnSe powder. In this evaporator, Molybdenum (Mo) boat is used as evaporation source and a glass substrate placed directly above this source. This glass substrate is polished by acetone solution and dried in sample drier.

Evaporating ZnSe powder in the vacuum chamber beyond its boiling point and fixing chamber pressure at 2×10^(-5) torr, a thin layer of ZnSe is deposited on the glass substrate. Thickness is monitored by a crystal sensor which is already connected to a digital monitoring system. During this fabrication process, substrate temperature is fixed at 290°C and after this fabrication, ZnSe films have been annealed for 60 minutes at 300°C. Using thickness monitor, we have fabricated five types of ZnSe thin films having various thickness (90 nm, 70 nm, 50 nm, 40 nm and 30 nm) by this vacuum evaporation technique and this whole fabrication process is depicted below in “Fig. 1,”.

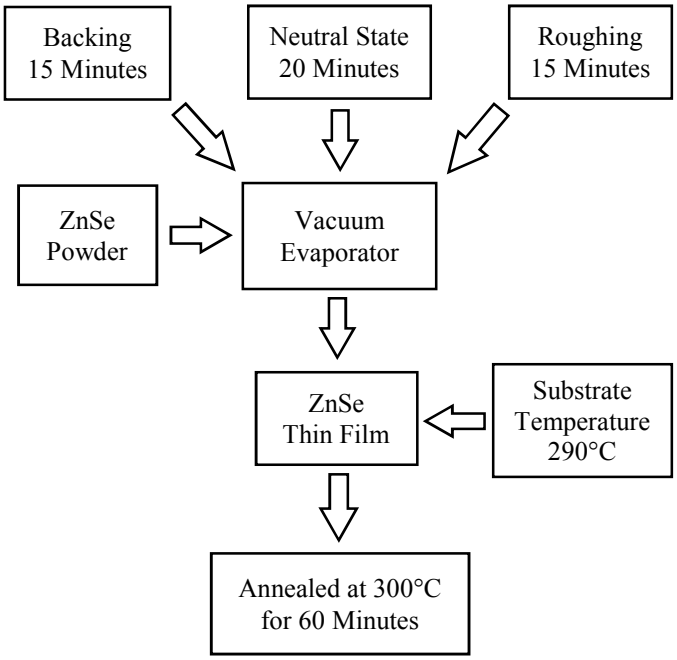


Fig. 1. Block diagram of ZnSe thin film deposition by vacuum evaporation technique.

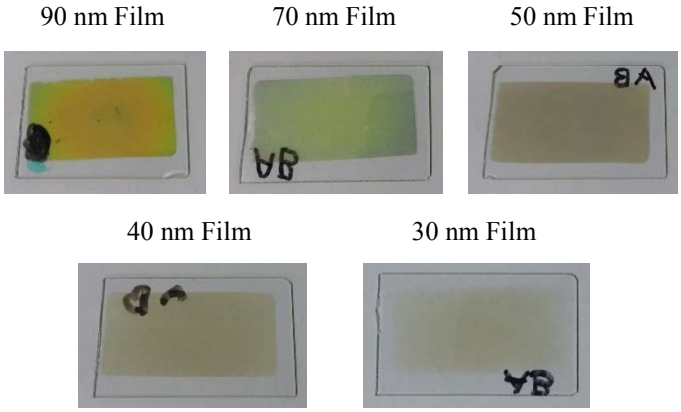


Fig. 2. Deposited ZnSe thin films on glass substrate at different thickness.

Deposited ZnSe thin films are shown in "Fig. 2", from which we can observe that as the film thickness reduces, corresponding visual intensity of film on the glass substrate decreases. That's for 90 nm film, visual intensity on glass substrate is maximum and for 30 nm film it is minimum.

III. RESULTS & DISCUSSION

After deposition, we have studied the structural and optical characteristics of five types of ZnSe thin films. Structural characterization is performed by X-ray diffractometer and UV-VIS-NIR spectrophotometer is utilized for optical characterization.

A. Structural Characterization

On the basis of X-ray diffraction principle, diffractometer operates for observing materials lattice structures. Here we have used this X-ray diffractometer to perform structural characterization of our developed films and to realize the effect

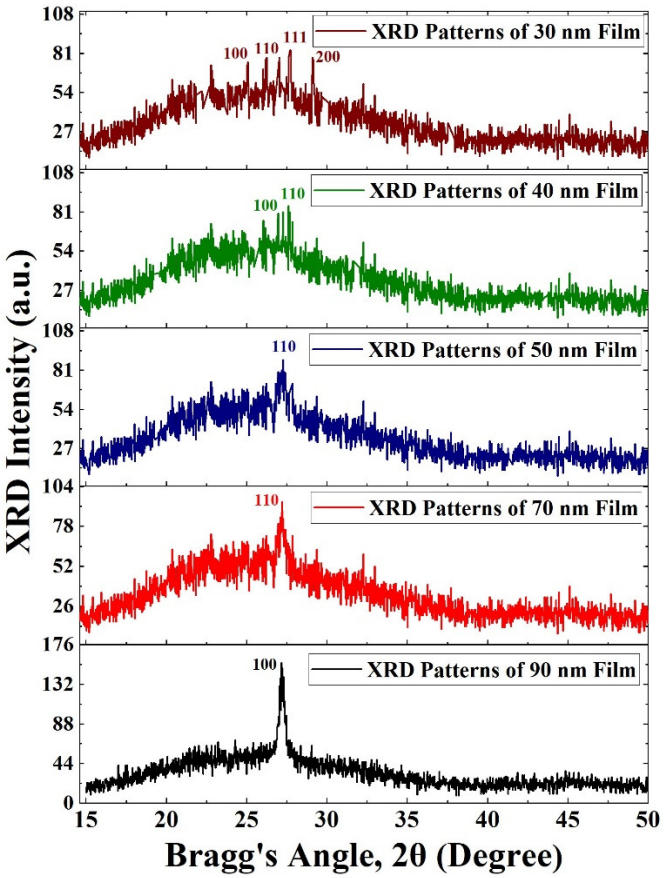


Fig. 3. X-ray diffraction patterns ZnSe thin films.

of thickness on ZnSe thin films. "Fig. 3", shows XRD pattern for all the five types of ZnSe thin films from which we can notice that ZnSe films with 90 nm thickness has maximum XRD intensity. As the film thickness decreases, maximum XRD intensity gets lower with different Bragg's angles.

90 nm ZnSe thin film has maximum XRD intensity of 156 a.u. (Arbitrary Units) at Bragg's angle 2Θ = 27.17° for 100 plane, for 70 nm film it is 94 a.u. at 2Θ = 27.49° for 110 plane. Again, 50 nm ZnSe film has maximum XRD intensity of 88 a.u. at 2Θ = 27.61° for 110 plane and which is 81 a.u. at 27.60° Bragg's angle for 40 nm film with two separate planes 100 and 110. At last, 30 nm ZnSe film shows maximum XRD intensity of 77 a.u. at Bragg's angle 2Θ = 27.68° consisting of four types of planes: 100, 110, 111 and 200.

These XRD data provide us the atomic configuration of these 5 films that ZnSe thin films having higher thickness possesses more stacked or immobile electrons in lattice because X-ray photons only interact with stacked electrons and reflected back to detector placed in the diffractometer. Using these XRD patterns we have calculated plane spacing, $d_{spacing}$ = 3.29 Å (Angstrom) which is almost same for all the five films and lattice constant, a = 5.6113 Å. For convenience, a unit cell structure of our developed ZnSe thin film is presented in "Fig. 4". From this unit cell structure, we can clearly identify that our fabricated ZnSe thin film has face centered cubic (fcc) structure.

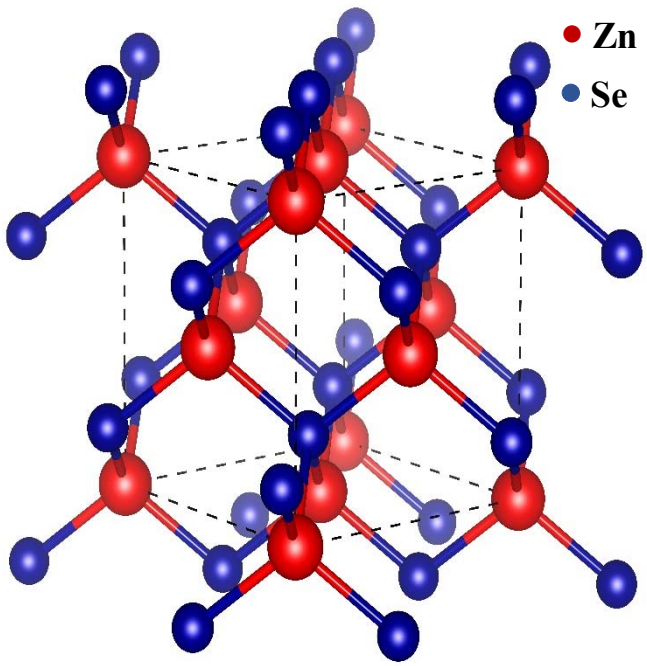


Fig. 4. Unit cell structure of ZnSe thin film.

B. *Optical Characterization*

Variable thickness of ZnSe thin film influences it's optical parameters like transmittance, bandgap, refractive index and reflectance. These optical properties are practically measured by UV-VIS-Near-infrared spectrophotometer and also simulated by Essential Macleod. First of all, lets discuss about both simulated and experimental transmittance of ZnSe which is graphically shown in "Fig. 5", and "Fig. 6", respectively. From "Fig. 5", we can notice that all the films have almost linear upward tendency of transmittance in near-infrared region and small fluctuation in visible region for higher thickness films.

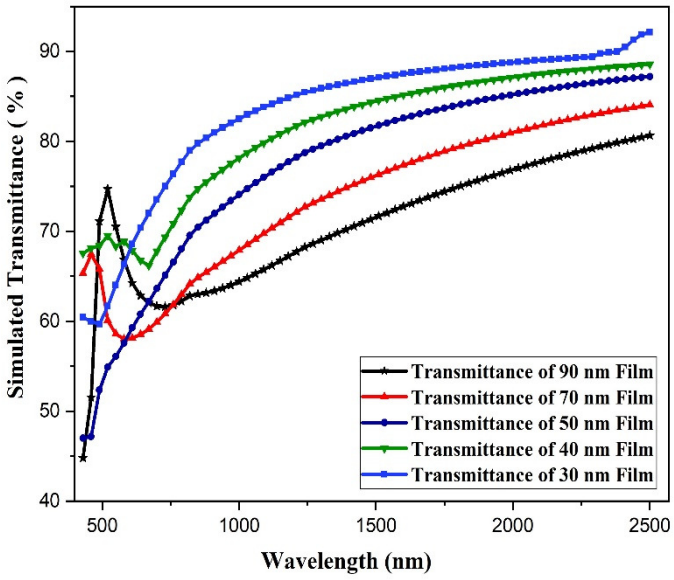


Fig. 5. Simulated transmittance of ZnSe thin films.

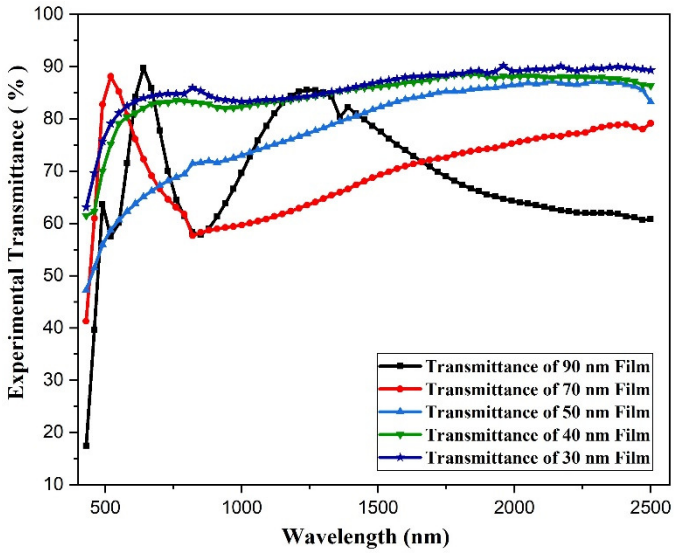


Fig. 6. Experimental transmittance of ZnSe thin films.

Another important factor we can notice from "Fig. 5", that as the film thickness decreases, corresponding transmittance increases which is maximum of 93% for 30 nm film and minimum for 90 nm film at a certain wavelength. "Fig. 6", shows practical transmittance of ZnSe thin film where 90 nm and 70 nm films have small variations in near-infrared region.

Higher thickness films have larger XRD intensity i.e. greater number of stacked electrons than free electrons. That's why these films have lower electrons in conduction band than valence band to conduct electricity.

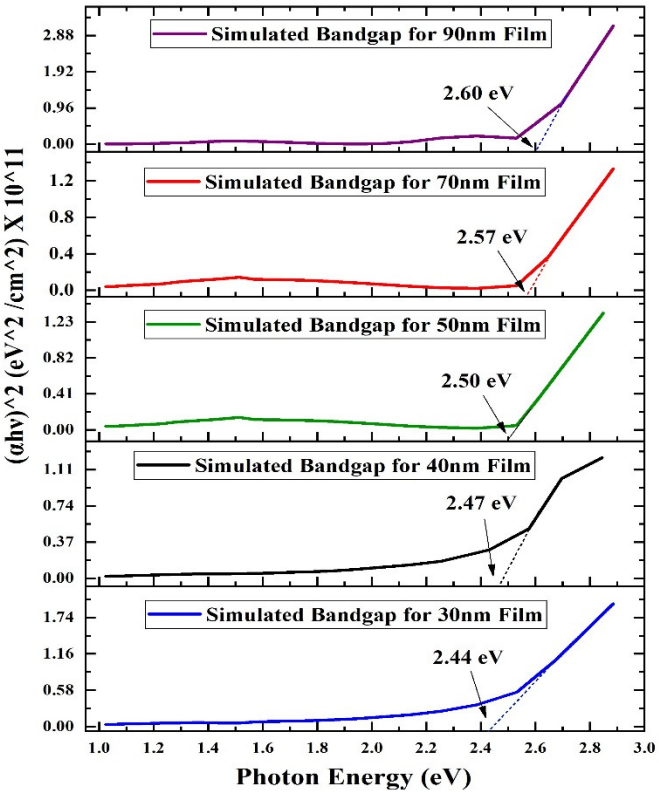


Fig. 7. Bandgap comparison of ZnSe thin films by simulation.

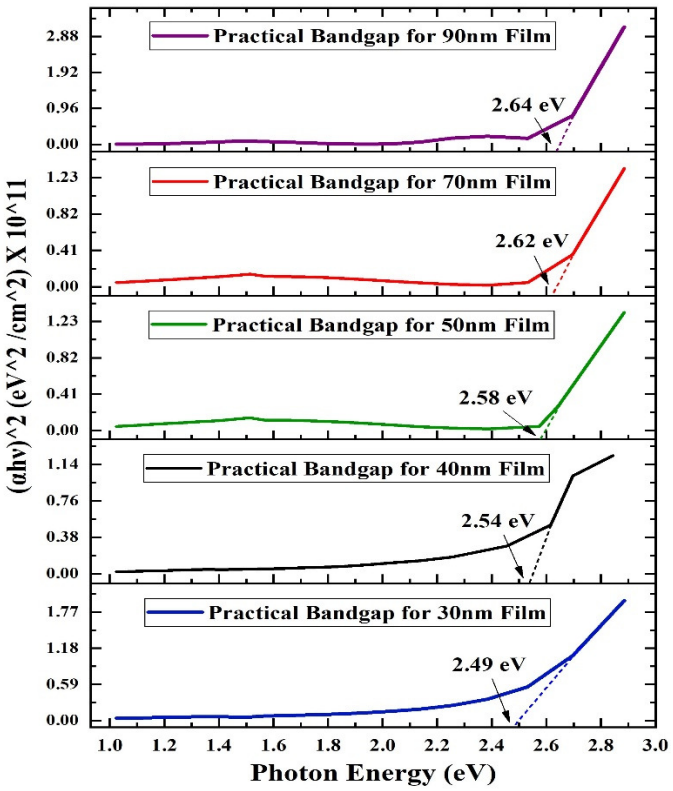


Fig. 8. Experimentally measured bandgap of ZnSe thin films.

This shortage of electrons in conduction band makes its optical bandgap wider. So ZnSe thin films with higher thickness have wider optical bandgap than films with lower thickness. Bandgap comparison for five films by both simulation and experiment is shown in "Fig. 7", and "Fig. 8", respectively and also in tabular form in "TABLE I".

Thickness also influences reflectance in such a way that for higher thickness films it is larger and for lower thickness films it gets lesser which is just opposite of transmittance. Due to the majority of stacked electrons in 90 nm film, most of the incoming photons in spectrophotometer is reflected by these stacked electrons. For this, 90 nm film has relatively the highest reflectance and 30 nm film has the lowest reflectance in the whole near-infrared region. Simulation as well as practical results of reflectance among the five ZnSe thin films are graphically shown in "Fig. 9", and "Fig. 10", respectively. Though simulation data provides smooth behavior, practical reflectance shows fluctuation for higher thickness films due to some experimental leakages.

TABLE I. OPTICAL BANDGAP FOR ALL THE FILMS

ZnSe Thin Film	Simulation Bandgap	Practical Bandgap
90 nm	2.60 eV	2.64 eV
70 nm	2.57 eV	2.62 eV
50 nm	2.50 eV	2.58 eV
40 nm	2.47 eV	2.54 eV
30 nm	2.44 eV	2.49 eV

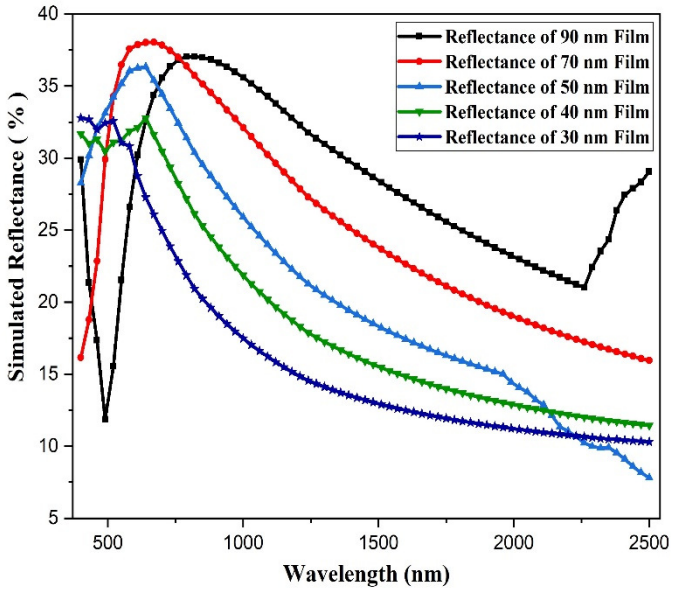


Fig. 9. Reflectance by simulation for all the five ZnSe thin films.

Photon energy of incoming light in UV-VIS-NIR spectrophotometer is related by the following equation with wavelength:

$$\text{Photon energy} = (4.135\times300) / \lambda \qquad (1)$$

Here λ is the wavelength of light in UV-VIS-NIR spectrophotometer. Refractive index (n) has been obtained from reflectance (R) and a new parameter extinction co-efficient (k). "Equation (2)" has been utilized to calculate both practical and simulation refractive index [10]:

$$\text{Refractive Index, } n = \frac{1+R}{1-R} + \sqrt{\{\frac{4R}{(1-R)^2} - k^2\}} \qquad (2)$$

For convenience, refractive index profile is shown in "Fig. 11", for both simulation and experiment. Like reflectance, refractive index increases as the film thickness increases in simulation as well as experiment.

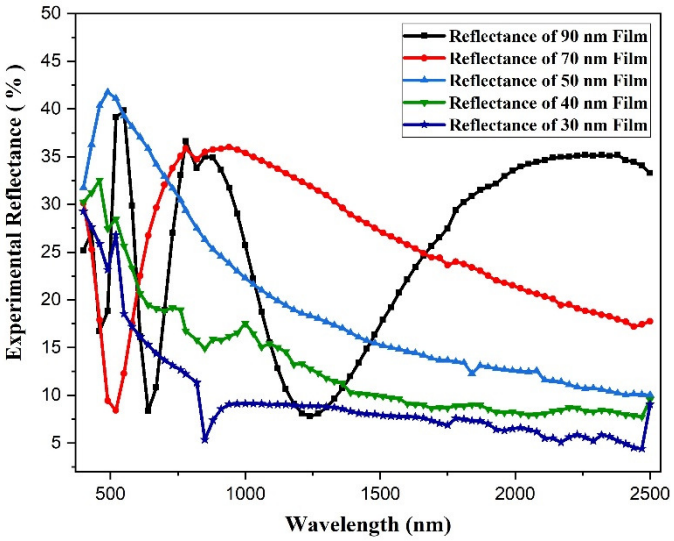


Fig. 10. Experimental reflectance for all the five ZnSe thin films.

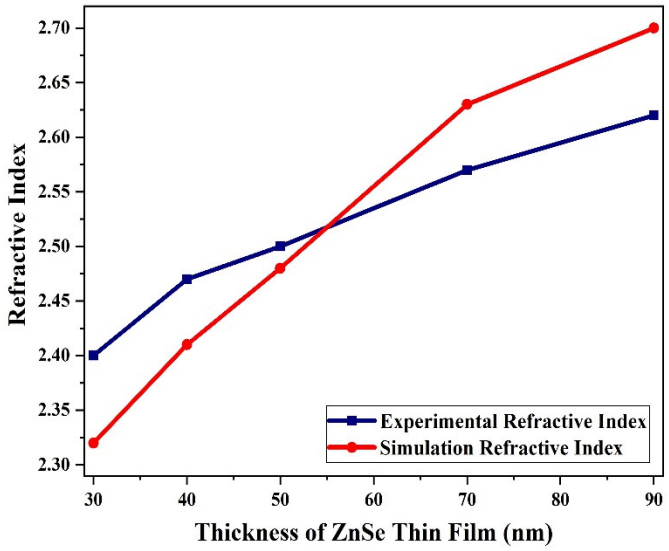


Fig. 11. Refractive index profile for all the five ZnSe thin films.

"Equation (3)" shows a way to calculate absolute dielectric constant which is composed of two parts: real and imaginary part [10]. Refractive index (n), extinction coefficient (k) and using real and imaginary part of dielectric constant we have calculated absolute dielectric constant by this complex number formula:

$$\text{Dielectric Constant (Real)} = n^2 - k^2$$

$$\text{Dielectric Constant (Imaginary)} = 2nk$$

$$\text{Absolute Dielectric Constant, } dk = \sqrt{\{(n^2 - k^2)^2 + (2nk)^2\}} \quad (3)$$

"Fig. 12", shows a graphical comparison of absolute dielectric constant in both simulation and experiment for the five ZnSe thin films. From which we have got a decreasing tendency of dielectric constant as the film thickness reduces.

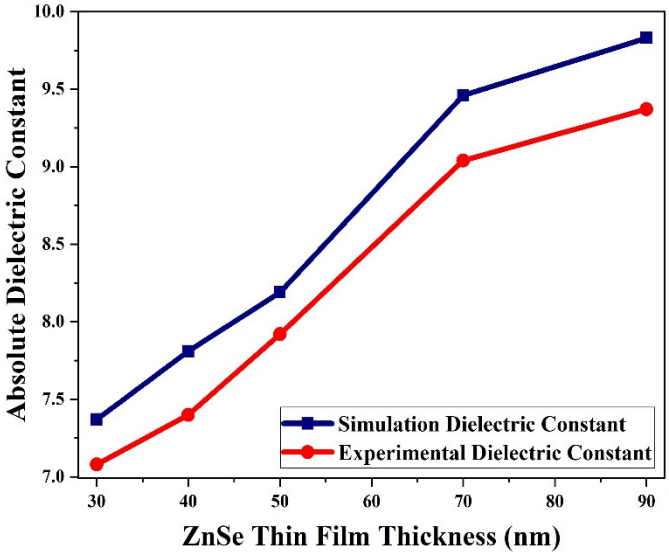


Fig. 12. Absolute dielectric constant for all the five ZnSe thin films.

IV. Conclusion

For investigating the effect of thickness in ZnSe thin film, we have performed this research and have found something promising. As the thickness of ZnSe thin film increases, optical bandgap increases creating an opportunity to apply it in wider bandgap applications such as LED, transducer and also in high electron mobility transistor (HEMT). XRD data revealed that lower thickness reduces the number of stacked electrons and turns ZnSe lattice into porous form. Again, XRD analysis and unit cell structure have proved ZnSe thin film has polycrystalline as well as face centered cubic structure. Because of this porous cubic polycrystalline structure, ZnSe thin films with lower thickness have high transmission coefficient, low reflection coefficient and reduced optical bandgap. These make lower thickness films more suitable than higher thickness films to use as a transmitting layer in several optoelectronic applications.

References

[1] Smith, Donald Leonard. Thin-film deposition: principles and practice. Vol. 108. New York etc: McGraw-hill, 1995.

[2] Muthumari, S., G. Devi, P. Revathi, R. Vijayalakshmi, and C. Sanjeeviraja. "Structural Investigation of Zinc Selenide Thin Films." Journal of Applied Sciences 12, no. 16 (2012): 1722-1725.

[3] Wang, J. B. N. J., J. B. Neaton, H. Zheng, V. Nagarajan, S. B. Ogale, B. Liu, D. Viehland et al. "Epitaxial BiFeO3 multiferroic thin film heterostructures." science 299, no. 5613 (2003): 1719-1722.

[4] Wang, Zhongchun, Ulf Helmersson, and Per-Olov Käll. "Optical properties of anatase TiO2 thin films prepared by aqueous sol–gel process at low temperature." Thin Solid Films 405, no. 1-2 (2002): 50-54.

[5] Prins, M. W. J., K‐O. Grosse‐Holz, G. Müller, J. F. M. Cillessen, J. B. Giesbers, R. P. Weening, and R. M. Wolf. "A ferroelectric transparent thin‐film transistor." Applied physics letters 68, no. 25 (1996): 3650-3652.

[6] Güllü, H. H., E. Coşkun, and M. Parlak. "Investigation of optical parameters of thermally evaporated ZnSe thin films." physica status solidi (c) 12, no. 9‐11 (2015): 1224-1228.

[7] Ikhmayies, Shadia J., and Riyad N. Ahmad-Bitar. "Characterization of vacuum evaporated CdTe thin films prepared at ambient temperature." Materials Science in Semiconductor Processing 16, no. 1 (2013): 118-125.

[8] Kim, Maeng Jun, Hong Tak Kim, Jin Kyu Kang, Dae Hwan Kim, Dong Ha Lee, Sung Ho Lee, and Sang Ho Sohn. "Effects of the surface roughness on optical properties of CdS thin films." Molecular Crystals and Liquid Crystals 532, no. 1 (2010): 21-437.

[9] Baban, C., G. I. Rusu, and P. Prepelita. "On the optical properties of polycrystalline CdSe thin films." Journal of Optoelectronics and Advanced materials 7, no. 2 (2005): 817-821.

[10] Amrane, Noureddine, Maamar Benkraouda, and Fathalla Hamed. "Ab Initio Calculations of Optical Properties of TlBr and TlCl Radiation Detectors." World Journal of Condensed Matter Physics 1, no. 03 (2011): 63.

Feasibility Analysis and Economic Impact Assessment of Microgrid Renewable Energy Systems in Burao, Somaliland

Abdiaziz Mouse Abdi Miad
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Kumira, Chittagong-4318, Bangladesh
E-mail: abdulasiis@yahoo.com

S. M. G. Mostafa
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Kumira, Chittagong-4318, Bangladesh
E-mail: mostafa_93eee_iiuc@yahoo.com

Mohammad Azim uddin
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Kumira, Chittagong-4318, Bangladesh
E-mail: azimctg50@gmail.com

Yasir Arafat
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Chittagong-4203, Bangladesh
E-mail: yaeeeiiuc@gmail.com

***Abstract* - The political and socioeconomic challenges in Somaliland have caused a lack of scaled and interconnected power infrastructure which resulted in a power shortage with power loss in transmission. The fluctuating cost of diesel fights to move away from the recovery of the post-conflict situation to a relatively rapid economic development. The reliance of the majority of the population on biomass to meet their urban and rural energy demands has depleted the forest and used crop residues, however, the use of biomass has recently decreased in urban cities. This paper presents a feasibility study of Solar Photovoltaic, wind turbine, and generators based on hybrid electric supply system to Burao Somaliland urban center (9.5259° N, 45.5346° E). This study has been carried out to show the supply of electricity to a small residential load and the target to reduce the cost of energy. Besides, a comparative study was carried out with renewable resource-based microgrid hybrid energy system compared to a current diesel-only micro-grid, hence in this case study, the best cost-effective solution is considered by using different renewable energy sources. Solar and wind potentials have been estimated from NASA surface meteorology. HOMER has been used for optimization and sensitivity analysis for the hybrid system.**

Keywords—Burao, Micro-Grid system, Renewable Energy, Optimization, HOMER, Wind Energy, PV, Biomass.

I. INTRODUCTION

The consumption of total fossil energy around the globe is 80% which contains fossil fuel energies, i.e., oil, coal, and natural gas [1]. However, the fossil energy is limited in supply, and it reduces the fossil fuel resources; thus, alternative energy to fossil energy is a crucial point of discussion as the fossil energy sources will be depleted one day compared with consumption rate [2]. Also, the fossil fuel burning is the main cause for growing the total of Greenhouse Gases (GHGs) in the atmosphere and another motive to think of alternative solution. The hybrid system and renewable energy can be of great aid in this condition for its environmental nature and eco-friendliness.

Recently, renewable energy becoming popular around all over the world to confirm the excellent development of getting or greeting the resource of the environment, e.g. sun or wind etc. through as electricity. The introduction of new technology which provides renewable energy is allowing both isolated and urban communities to increase their quality of life by finding renewable and reliable energy systems. Increasingly, Communities require to investing in off-grid Renewable energy facility systems. Burao is located in the region of Somaliland (09°31′27″N 45°32′13″E) and has a total population of 288,211 residents, and it is the second-largest city in Somaliland. The average daily solar global radiation for Burao ranges from 4.99 to 6.41 kWh/m2/day, and monthly average wind speed varies from 4.2 to 8.2 m/s [3]. The cost of generating energy from the above hybrid system was different COE US$/kWh Based on the cases). But for the diesel, the only option in the existed arrangement leveled cost of energy for Burao $1.0/kWh and sometimes less than 0.8 $ /liter. The problem of high-cost diesel has affected the production and business activities which requires affordable electricity to continue their day-to-day productions and works, in some cases their outsource operations to other countries with a lot of reasonable power facilities. Two hundred eighty-eight thousand two hundred eleven residents, it is the second-largest city in Somaliland.

The people need electricity at a cheap rate and also keep environment pollution free. The diesel based generators increase carbon dioxide in atmosphere catastrophically. But using the resources of a renewable energy does not affect environment or there is almost zero carbon dioxide emission. This paper focuses on the feasibility d analysis and economic impact assessment of microgrid renewable energy systems in Burao, Somaliland whereas the cost of electricity is high from national grid also the private grids.

The cost of energy is the lowest (USD 0.2522/kWh CEOs) of the hybrid micro-grid system. The optimum system configurations varied based on assumption of diesel pricing, but hybrid energy systems are the most economically viable solution comparing with diesel driven arrangements. Moreover, the majority of people living in rural and urban areas have no access to electricity, although the power is only available in city center and in practically totally generated by diesel imported from foreign countries [4]. The people live in the city, or urban areas are paying some of the highest prices in all over the world for a limited ($0.80 to $1.50) [5] per kilowatt hour, and sometimes the services are not reliable and there are electricity outages and its unreliable service on productivity in the developing world mostly it leads to substantial aggregate effects of economics.

II. Electrical Load Profile

In this project, the electrical load consists of residential demand and some business stores. The average load of each residence is considered to be 50 kWh/day. The characteristics of the load are estimated according to TABLE I.

TABLE I. LOAD MEASUREMENT FOR DIFFERENT FAMILIES

No. of residences	100
Peak demand	532.75[kW]
Duration of peak demand	1800-2200
Duration of low demand	0000-0004
Annual Average Load	2588[kW]

The yearly daily load profiles of 100 residences are shown in Fig. 1. The horizontal axis represents hours, and a vertical axis corresponds to the kW of consumption.

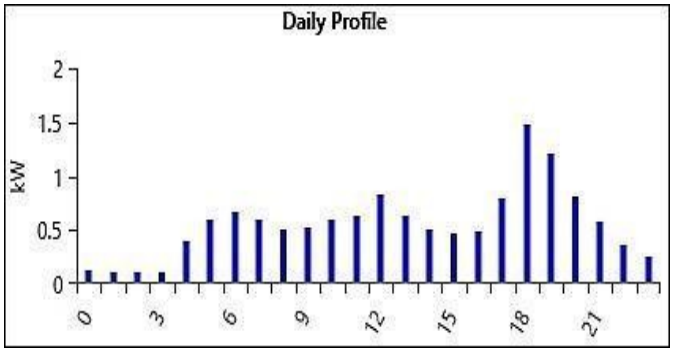


Fig. 1 The daily load profile of 100 residences.

During night the annual power consumption is high.

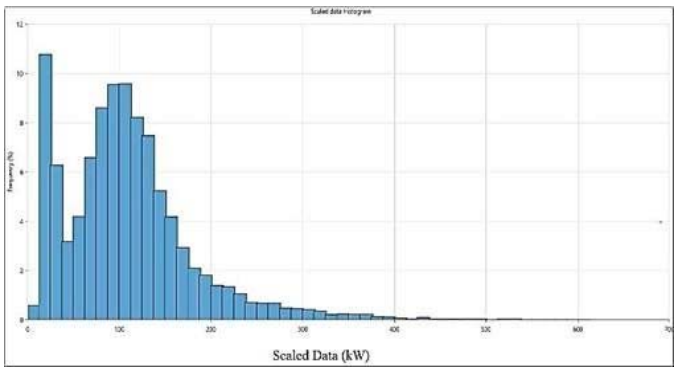


Fig. 2 Frequency histogram of annual power consumption of the 100 Residences.

III. Renewable Resources

The annual horizontal level of radiation in the area is 5.960 kWh/m^2/day [6]. In Fig 3. shows the monthly level of solar irradiation. The monthly variations of solar radiation and Clearness Index (CI) are given in TABLE II.

TABLE II. MONTHLY SOLAR RESOURCES IN BURAO

Months	Clearness Index(CI)	Solar Radiation (kWh/m^2/day)
January	0.589	5.270
February	0.631	6.060
March	0.625	6.410
April	0.587	6.170
May	0.604	6.280
June	0.616	6.300
July	0.605	6.210
August	0.611	6.350
September	0.619	6.360
October	0.580	5.660
November	0.602	5.460
December	0.574	4.990

HOMER needed numerous data of solar irradiation to calculate the solar power generated using the panel. The monthly mean solar radiation data has been taken from NASA surface meteorological department. It has monthly averaged global horizontal radiation for periodic of 20 years.

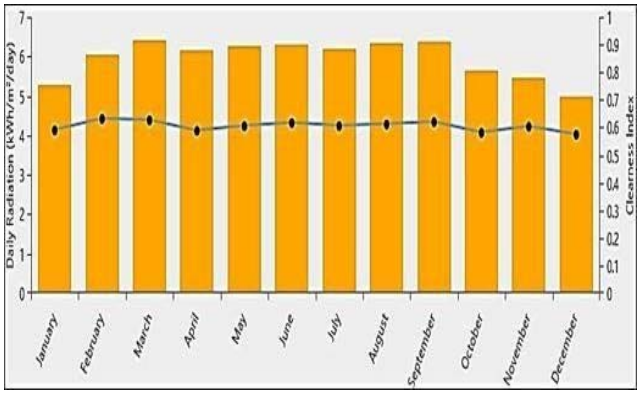


Fig. 3 Graphical presentation of solar radiation per month.

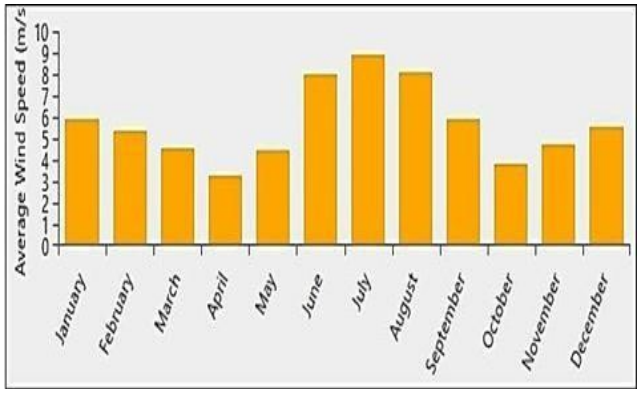


Fig. 4 Graphical presentation average wind speed per month.

The average wind velocity has considered to be 5.710 m/s annually [7]. It is observed from Fig. 4 that the wind velocity remains in the range of 5-6 m/s throughout the year. This range is very appropriate for the implementation of wind power project.

In Fig.5. demonstrates a combined hybrid system designed to fulfill electrical power demand in Burao. That is the Diesel Generator, Bio-gas generator, Wind Turbine, PV and Battery. The designed configuration has a total Net Present Cost (NPC) of $3,079,776.00 and the Cost Of Energy (COE) of $0.2522/kWh [8].

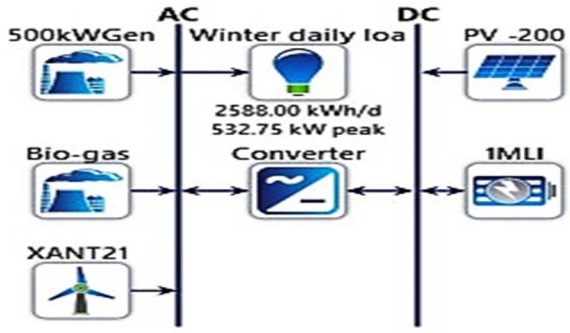


Fig. 5 Designing of hybrid energy system in Burao.

A. Photovoltaic System

To produce electrical power from solar energy a Generic flat plate Photovoltaic (PV) has incorporated in the proposed designing. 200 kW PV arrays has used that have 20.1% capacity factor that could supply around 40.2 kW. Fig.6 depicts the output power of the arrays that used in proposed model. The cost of installation for the PV array is around 1,035,000.00$ with a lifetime of 20 years. To ensure the best use of day lights the sun tracker attached with PV panels also moves on horizontal axis every day. In each year the PV modules produce around 351,778 kWh by the optimized model that is almost 32 percent of the total electrical power generation by the proposed system.

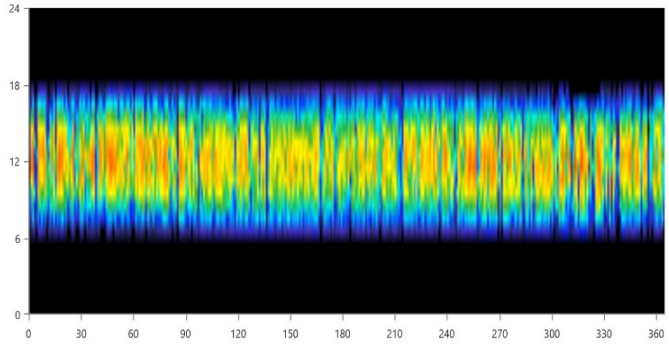


Fig. 6 Power output of generic flat plate PV

B. Wind Turbine

In this proposed hybrid model a XANT21 wind turbine has used to convert available wind energy to electrical power. Every wind turbine 200 kW rated that can supply about 44.6 kW along with 22.3% capacity factor. Twenty-five wind turbines (WT) had incorporated in proposed hybrid model considering 25 years' project lifetime. The initial installation cost is around 53,333$, and O & M cost is about 215$ per year has considered. In this optimized model, the twenty-five WT generate around 390,336 kWh annually via 6967 hours'/year operation.

C. Diesel Based Generator

To generate the electricity from the diesel generator, we have used Generic 500kW Fixed Capacity Genset with a fuel curve of intercept 7.00L/Hr. and the fuel curve slope is 0.244L/hr/kW. The capital cost of a diesel generator is $81,468.00. And the emission of a carbon dioxide is (g/L fuel) 13.566 and the lifetime hours is 15,000hrs.

D. Li-Ion Battery as a Storage

However, the renewable energy sources provide average quantity energy for an extended period, hence it must need energy storage devices to supply uninterrupted power to the customers or buyers when the resources furnish low amount of power. We have used three batteries that is a Generic 1MWh Li-Ion unit to store the energy of PV panel which at a 12V nominal voltage. Two groups separate panels have used in every string to assure 24V at each bus. The capital cost of a battery is $233,333.33, and cost of replacement for each cell is $98,997.23.

E. Converter Arrangement

The converter can be employed to convert both DC to AC or AC to DC which implies the converter serve as either rectifier nor inverter. In the design model, it is needed to convert the DC voltage supplied by the PV modules and batteries as the energy is used by Alternating Current loads. So, in this research 100kW converter has used for the conversation purpose. The capital cost or the amount required is $55,904.26 and the replacement cost for the converter is $23,718.71. This converter has considered 90% efficient for 15 years of a lifetime.

F. Biogas Generator

To produce electrical power from household and other wastage a generic biogas Genset has considered with minimum 50% load ratio. The first investment cost is 3,000$ per kW where 5$/hour Operation and Maintenance (O&M) cost has taken in consideration for 20 thousand hours' life span. Also, 5.5 MJ/ kg lower heating value of the biogas considered that contain 5.5% carbon content. The people of Somaliland depend on live stocks like Camels, Cattle, sheep, and Goats; also the second most important resource is crop farming, i.e., Sorghum and Maize which play the role of economic activities even its considerable underdeveloped potential especially the cover plants annual average produce 0.710 Tons, so it is possible to create an electric potential by using a Biogas Generator.

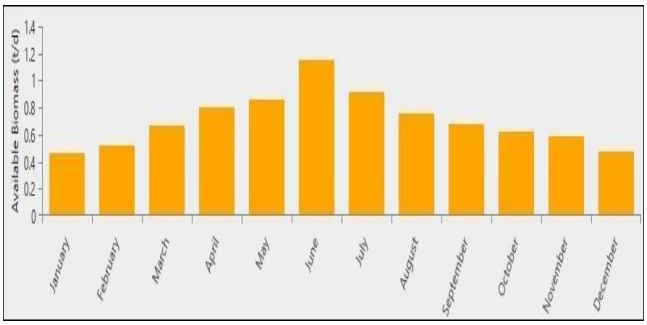


Fig. 7 Variation of Average Biomass Resources in Burao throughout the year.

The monthly biomass resource production is shown in fig.7. The highest output is found in June. At that time there are many waste plants and animals like the cover of sorghum and maize and the lowest month is January.

IV. RESULT ANALYSIS

The optimal configuration obtained from the simulation is considered as follows:

- ✓ Diesel power project 500 kW
- ✓ Solar energy project 200 kW
- ✓ Wind power generator 200 kW
- ✓ Biogas Generator 100 kW

In this project, the income obtained from electricity consumption of 100 residences gives a monthly amount of USD 0.2522$. Therefore, the system is going to collect USD $25.22 and during the 20 years USD 6052.8$. This cost is enough to cover the cost due to maintenance and operation. The summary of the values of the initial capital cost is given by Fig 10. Although the significant cost to implement the proposed hybrid model is the capital cost and it is around 1.77 million dollars. Though the installation cost seems high, during the 20 years, the total cost for O & M services are $238,842.76.

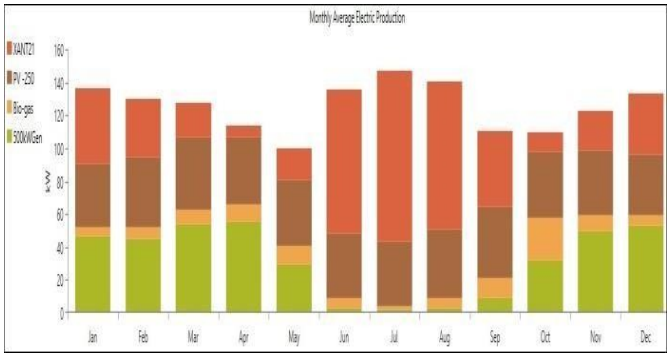


Fig. 8 Monthly average electricity production from different types of Resources.

The NPC of the project is $3,079,776.49 which is less than the rural electrification budget USD $29,368,000.00 million [9, 10]. Hence it is economically feasible. Fig. 8 depicts the average generation of electric power per month of numerous types of resources.

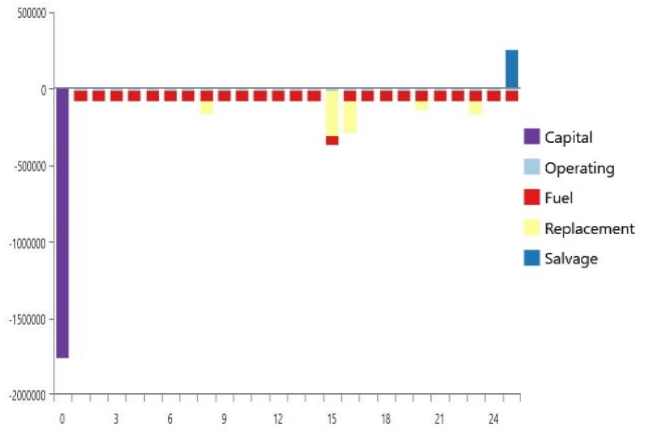


Fig. 9 Cash Flow.

V. ENVIRONMENTAL AND SOCIAL ASSESMENT

In this project environmental assessment of micro-grid is also discussed. Somaliland is a developing country that affects climate change. This is mainly due to the high reliance on fossil fuel.

A. Environmental Assessment

The use of power from renewable energy will decrease or reduce the number of gaseous pollutants in the atmosphere. Diesel fuel contains 80,658L of pure carbon. A liter of an Atomic carbon forms 3.78541gallon of carbon dioxide when oxidized with O2 considering 99% of the fuel is oxidized the amount of carbon dioxide per gallon becomes 305,322.793litre [11]. Table V the amount of carbon dioxide produced by 500kW generator if operated alone and when integrated with renewable resources. It is noticeable that the amount of carbon dioxide and carbon monoxide is reduced when the diesel generated energy is combined with Renewable Energy. Nowadays the electricity generation in Somaliland is generally done through diesel generators. However, it has one of the highest electricity rates in the northwestern Somalia; however, the source of renewable energy is the energy that is collected from renewable resources, and it has a high in solar radiation all the months or the year, and also three months are the most top resources of wind speed (June, July, and August) as we have seen in Fig. 4 [12], and the other months also it's available the resource of renewable energy. Using the diesel generator, it has a lot of effects or allots of carbon dioxide or emission, but renewable energy sources have no impact of the environment problem, the comparison of electrical production in renewable energy sources and also diesel generator are the following in TABLE III.

TABLE III. ELECTRICAL POWER GENERATION

Production	kWh/yr	%
Generic Flat Plate PV	351,778	32.0
Generator 500kW	274,309	24.9
Wind turbine	390,336	35.5
Biogas Generator	84,080	7.64
Total	1,100,503	100

There is a slight difference between the production of renewable energy and the diesel generator, although we are using a considerable difference capacity between the diesel generator and solar or wind [13]. Here we understood the benefits of renewable energy are higher than the diesel generator. TABLE VI shows that the integrated renewable energy systems will reduce CO_2, CO, SO_2, and NO_2 more than four times and five times respectively than existing diesel-based generator systems.

TABLE VI. COMPARISON OF EMISSIONS: DIESEL VS INTEGRATED RENEWABLE ENERGY

Production	Alone	Integrated
Gal/Hr	227.124	227.124
Hrs/Yr	3100	2124
CO_2	967,333kg/yr.	211,551 kg/yr.
CO	5,004 kg/yr.	1095kg/yr.
Unburned Hydrocarbons	266kg/yr.	58.1kg/yr.
Particulate Matter	42.8kg/yr.	9.36kg/yr.
Sulfur dioxide	2,365kg/yr.	517kg/yr.
Nitrogen Oxides	959 kg/yr.	210kg/yr.

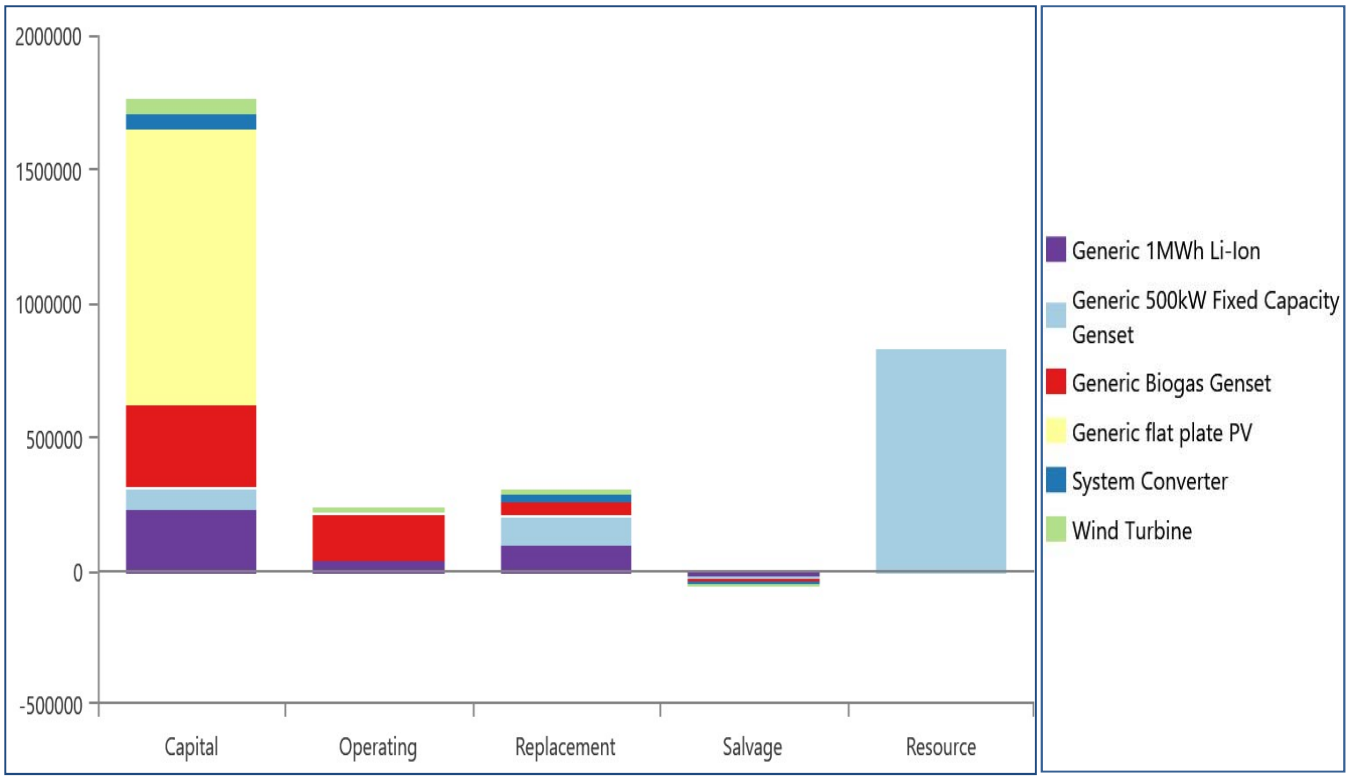


Fig. 10 Summary of cost of the proposed designed model for different components.

TABLE V: COST SUMMARY DESIGNED SYSTEM FOR DIFFERENT COMPONENT

Name	Capital	Operating	Replacement	Salvage	Resource	Total
Generic 1MWh Li-Ion	$233,333	$43,092	$98,997	-$18,632	$0.00	$356,790
Generic 500kW Fixed Capacity Genset	$81,468	$5,070	$108,541	-$14,279	$834,170	$1.01M
Generic Biogas Genset	$300,000	$164,826	$50,994	-$12,165	$0.00	$503,655
Generic flat plate PV	$1.04M	$0.00	$0.00	$0.00	$0.00	$1.04M
System Converter	$55,904	$0.00	$23,719	-$4,464	$0.00	$75,159
Wind Turbine	$60,000	$25,855	$19,128	-$10,780	$0.00	$94,203
System	$1.77M	$238,843	$301,379	-$60,321	$834,170	$3.08M

In table V, it denotes the summary of cost for the optimized hybrid model by various components. The main cost to implement the proposed model is capital cost which is about $1.77 million. But the amount of a total system is 3.08 million USD.

B. Social Assessment

When microgrid installed in rural areas, new job opportunities will be created. Job openings will not only be in the design and installation phase, but also experts will be required for the operation and maintenance of microgrid. It will also help to develop the livelihood of the local lower-middle-class families by creating jobs for them [14]. Reliable and affordable supply of electricity will help to boost up the local economy. In our time, electric infrastructure is the backbone of the economy. Initially, the business and industries will not operate due to the lack of supply and diesel generator is not economically efficient but by installing microgrid the business hours are also increased and hence produced capacity increases.

VI. CONCLUSION

In this study, we discussed the impact of the microgrid on the environmental and socioeconomics of Burao Somaliland. The result of the simulation obtained from HOMER shows that it is feasible to deploy microgrid by integrating conventional diesel generator with renewable resources. The introduction of the microgrid in rural areas will help to boost the local economy and also it will have a positive impact on the environment. The simulation shows that the cost of energy is meager and affordable by poor people. The amount of pollutants released in the atmosphere is reduced by encouraging renewable energy resources. The proposed system is very eco-friendly by reducing GHG emission at a remarkable level.

REFERENCES

[1] M. Shahadat, H. Lipu, S. Uddin, M. Ahad, and R. Miah, "A feasibility study of a solar-wind-diesel hybrid system in rural and remote areas of Bangladesh," International Journal of Renewable Energy Research, vol. 3, no. 4, pp. 892-900, 2013.

[2] Noam Lior,, Energy resources and use: The present situation and possible paths to the future,‟ Energy, vol. 33, pp. 842- 857, 2008.

[3] NASA surface meteorology and solar energy [Online]Availablehttps://eosweb.larc.nasa.gov/cgibin/sse/grid.cgi?uid=8584.

[4] World Bank. Doing business in Hargeisa: comparing regulation for domestic firms in 183 economies. Doing business sub-national. Washington, DC; 2012.

[5] Owuor, VO. Wind Energy, Shuraako.org, 2012, viewed 3rd June 2014.

[6] Bekele, G. and G. Tadesse. "Feasibility study of small Hydro/PV/Wind hybrid system for off-grid rural electrification in Ethiopia." Applied Energy, vol. 97, pp. 5-15, 2012.

[7] http://shuraako.org/sites/shuraako.org/files/case-studywindenergy.pdf.

[8] ‘‘HOMER’’ a software Hybrid Optimization Model for Electric Renewables developed by US national renewable energy laboratory.

[9] Nuñez, Jami Nelson. "Powering Progress: The Potential of Renewable Energy." (2015).

[10] A. M Abdilahi, et al.," Feasibility study of renewable energy based microgrid system in Somaliland ′s urban centers," Renewable and Sustainable Energy Reviews, vol. 40, pp. 1048 1059, 2014.

[11] N. E. Benti, "Combining Wind and Solar Energy to Meet Demands in Somali Region of Ethiopia (A Case of Dembel District)," American Journal of Modern Energy, vol. 3, no. 4, pp. 73-83, 2017.

[12] F.Bin Kashem and M. S. Arefin, "Feasibility Analysis of Grid Connected Mini Hybrid Energy System in Coastal Region of Chittagong, Bangladesh," International Conference on Electrical Information and Communication Technology (EICT), Khulna Bangladesh, pp. 1-5, 2017.

[13] D. C, Poplack Kelsey Powering Progress II, Realizing the Potential of Renewable Energy in Somaliland, Hargeisa, Somaliland, 2016.

[14] G and B. Bekele Palm. "Feasibility study for a standalone solar–wind- based hybrid energy system for application in Ethiopia." Applied Energy vol. 87, no. 2: pp. 487-495, 2010.

Lightning Effect on a Large-Scale Solar Power Plant with Protection System

Mohammad Riajul Karim* and Md. Raju Ahmed
Department of Electrical & Electronic Engineering
Dhaka University of Engineering and Technology (DUET)
*Email: riaz4623@gmail.com

***Abstract*—While lighting is a potential natural threat to a solar power generation plants and their electrical equipment, conventional protection against it has still remained inadequate. A possible reason is that the effect of lighting is not completely realized with the requirements and design considerations of the protection system. Unlike the other installations and systems susceptible to lighting, the solar panels extended over the large and open area are usually more exposed to the lightning strike. Lightning creates a strong electromagnetic field and induces extremely high voltage for a moment that can damage the photovoltaic (PV) panels, DC lines, inverter, underline cables and other equipment. In support of safety-protection, in this paper, we have modeled a Lightning Protection System (LPS) and investigate the lightning effect on a large-scale solar power plant with the proposed LPS. Additionally, we have analyzed the variations in the electromagnetic field, induced voltage and current due to lightning in the plant with the LPS using Virtual Surge Test Lab (VSTL) simulation. The presented analysis and recommendations would help to determine fundamental requirements and design considerations of an appropriate LPS for a large scale solar power plant.**

***Index Terms*—Electromagnetic field; lightning Induced voltage; Protection system; PV panels; FDTD method; VSTL.**

I. Introduction

As a potential green energy source, the solar power plants are rapidly growing all over the world [1]–[3]. To support the global clean energy revolution and to mitigate the lack of fossil fuel, Bangladesh has also taken initiatives for establishing solar power plants in the government and private sectors [4]. Many mini-grid solar power plants have already been running in Bangladesh such as the plants at St Martin, Kutubdia Island, Sandhip, Vola, Bangladesh Bank Motijil and Manikgong. Several large scale solar power plants are also planned to establish towards the government's target of producing 10% of total power demand with renewable energy by 2020 [5]. With this rapid expansion of the plants, their safety against the increasing natural calamities like lightning has posed an additional challenge.

A Lightning Protection System (LPS) is usually considered adequate for the protection of a solar power plant against the lightning effect. However, unlike other installations and systems susceptible to lighting, the solar panels extended over the large and open area are usually more exposed to the lightning attack. So, the requirement of an LPS for a solar power plant should reasonably be different from that of the other installations and systems. This gap in design considerations of an LPS may potentially leave it to be inadequate for a solar power plant. One obvious reason for employing such an inadequate LPS is that the effect of lighting is not completely realized for a solar power plant.

To determine the requirements for developing an effective and adequate LPS for the solar power plants, study of the lightning effect on the solar power plant is essentially required. Conventional LPSs, which are usually considered for the mini-grid solar systems in Bangladesh, include only a few (*e.g.*, one or two) air-termination rods with direct earthing from inverter to ground. Such LPSs will presumably not protect the whole PV plant against lightning. This is because, the solar power plant can be affected by both the direct and indirect strike of lightning. Both the direct and indirect lightning creates a strong electromagnetic field and induces high voltage which can damage the photovoltaic (PV) panels, PV generator, DC lines, inverter, underline cables and other equipment of the solar power plants [6]. Those effects of lightning can be more critical for the safety of a large scale solar power plant.

Most of the lightning effect analyses in the literature capture the power lines [7], [8], microwave relay station [9], [10], wind turbine [11], [12], small scale solar panels [6], [13]–[16]. However, no or a little effort can be tracked on the analysis of medium and large scale solar power plant that occupies a relatively larger open area. Induced voltage on DC cable due to the lightning electromagnetic interference is studied in [14] to demonstrate that the induced voltage can also damage the underground DC lines. Using electromagnetic analysis, Rabbani *et al*. [15] showed that the induced voltage also depends on the depth of the pipeline, strike location, soil resistivity and permittivity. The above analyses although identified essential factors and their effect on the induced voltage, they mostly considered a small scale solar power plant.

Additionally, Charalambous *et al*. [16] studied the effect of lightning on a solar power plant with a protection system. Their proposed models were analyzed for the current flow and energy dissipation through the earth-terminal. Although an isolated and non-isolated protection systems were employed, induced voltage due to induction was not considered in their model. In our earlier work [17], we studied lightning effect on a large-scale solar power plant to analyze the electromagnetic field, induced voltage and current due to lightning in the plant using the Virtual Surge Test Lab (VSTL) simulation.

In this paper, we model an LPS and investigate the lightning effect on a large-scale solar power plant with the proposed

LPS. Unlike our earlier work in [17], the current in the DC lines of PV plants is calculated using time domain simulation process in presence of our proposed LPS. Additionally, we have analyzed the variations in the electromagnetic field, induced voltage and current due to lightning in the plant with the proposed LPS using VSTL. With simulation and analysis, we conclusively demonstrate the lightning effect on the electromagnetic induction, induced voltage and current in a large-scale solar power plant in presence of the proposed LPS.

The rest of this paper is structured as follows. Modeling of a large scale solar power plant and its LPS is presented in Sec. II. Simulation results and analysis with initial recommendations for an appropriate LPS of a solar power plant are presented in Sec. III. Concluding remarks are noted in Sec. IV.

II. Proposed Model for the LPS Analysis

In this section, we present the specifications of a large scale solar power plant, a lightning protection system (LPS) for it, and their models using Virtual Surge Test Lab (VSTL) to analyze the variation in lightning effect on the plant in presence of LPS. Before we describe those specifications and simulation models, we introduce below the working principle of VSTL.

There are many methods of analyzing the electromagnetic induction and induced voltage in a solar power plant. Virtual surge test lab (VSTL) is one of the most suitable methods to analyze the electromagnetic field, induced voltage and current, due to lightning on a large-scale solar power plant. VSTL is developed based on FDTD [18] method by the Central research institute of Electric power industry, is simulation software to analyze the lightning current.

Lightning stroke channel having propagation velocity of return stroke is one-third of the speed of light and represented by engineering model built in VSTL [19]. The current waveform of that model is evaluated in [20], and the analytical expression is given in Eq. (1).

$$i(0,t) = \frac{i_{\max}}{k} \times \frac{\left(t/T_1\right)^n}{1+\left(t/T_1\right)^n} \times e^{-t/T_2} \tag{1}$$

Here, $i_{\max}$ is the peak current, T_1 is the front time-constant, k is the correction factor for the peak current, T_2 is the tail time-constant, and n is an exponential factor. The value of n varies from 4 to 10, T_1 varies from 0.25us to 10 us, and T_2 varies from 100us to 350us. These values are considered with i_{max}=100KA, $k = 1$ $n = 8$, T_1= .25us/1us, and T_2= 100us. The lightning current shape used for our model is shown in Fig. 1.

A. Solar Power Plant Specifications

We considered 150*KW* solar power plant for our analysis. A general layout of PV panels is illustrated in Fig. 2. A standard PV panel has 250 *watt* output power. Therefore, for the generation of 150*KW* solar power, 600 panels are required as a minimum. Single solar panel occupies an area of $1.5m \times 1m$. Therefore, a $50m$ long and $4m$ width frame can hold 100 PVs in two sub-rows, which occupy $50 \times 4 = 200m^2$.

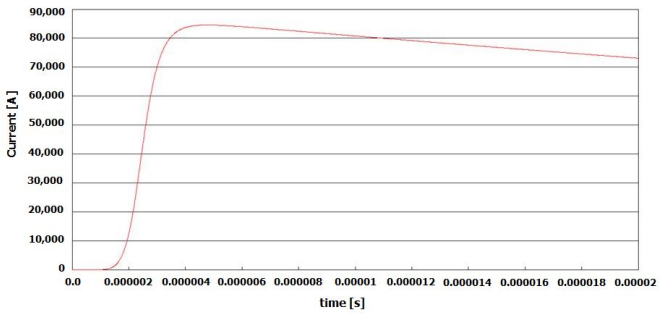


Fig. 1. The injected lightning current (*LC*) waveform used for the VSTL model

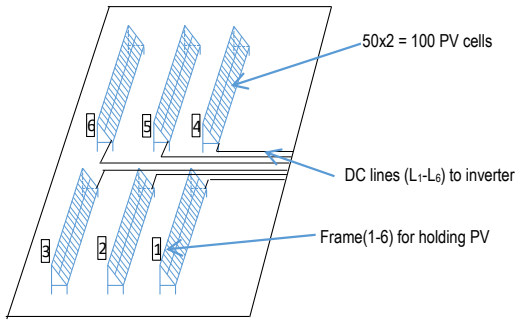


Fig. 2. The layout of the PV panels of 150*KW* solar power plant

The gap between the two rows is $6m$, and the gap between the two columns is $10m$. The required total area is thus $140m \times 45m = 6300m^2$. The PV panels are placed on the fixed inclination frame. The height at the front of a frame is $2m$ from the ground, and at the back side, the height of the frame is $4m$ from the ground. All the DC lines are placed $4m$ above the ground, and DC-lines of each group are combined and lead to the inverter.

B. Our Proposed LPS for the Solar Power Plant

The design of our proposed LPS for a large scale PV system mainly depends on the appropriate risk assessment, type of plant, location of the plant. In this paper, we have emphasized on external LPS, which is consist of air-termination system and earth-termination system. The layout of the proposed LPS is illustrated in Fig. 3. Air-termination rods of $10m$ height is located at the fixed $1m$ back of each PV Supporting frame. For reducing the cast, we can fix the air-termination rod with PV frame with *L*-type clump and insulator to avoid the direct contact with the frame. These air-termination rods will then be connected with the earth-termination system. In this proposed model, three air-termination rods are planned behind each PV frame, which is $50m$ long as such it will offer a necessary protection angle and protect the PV panels from the direct strike of lightning. All the air-termination rods are interconnected through earth-electrodes, which is made of copper wire and laid $1m$ under the ground. In the proposed LPS, the air-termination system intercepts between the lightning and the

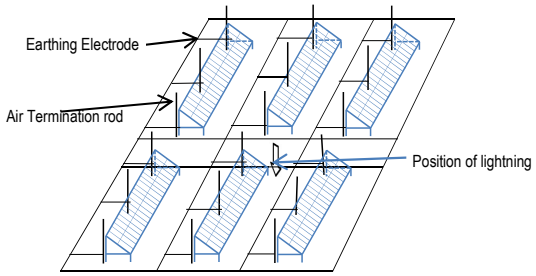


Fig. 3. Layout of the proposed LPS with air-termination rod and earth-termination system for 150kw solar power plant

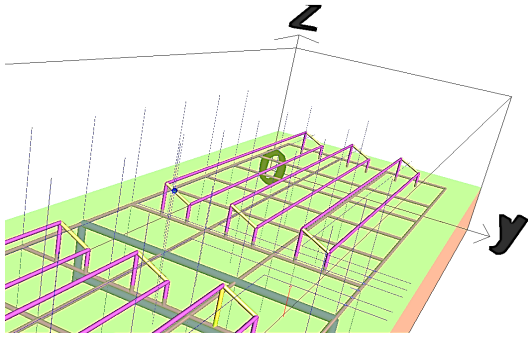


Fig. 4. Our VSTL simulation model of the 150kw solar power plant

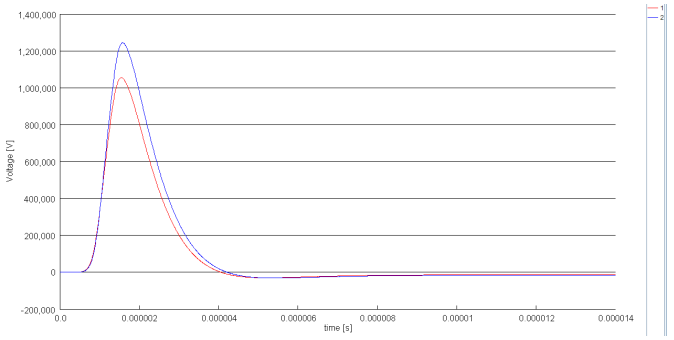


Fig. 5. The induced voltage in the DC lines without any LPS

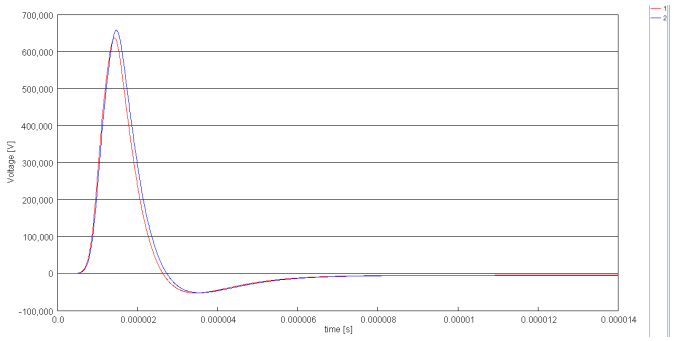


Fig. 6. The induced voltage in the DC lines with the proposed LPS

physical parts of the plant. Air-termination system embraces the down-conductors to discharge the lightning current in the general mass of earth through the earth-termination system.

C. VSTL Simulation Model

The VSTL simulation models have been developed for the $150KW$ solar power plant without any LPS and with the proposed LPS as showed in Fig. 2 and Fig. 3, respectively. The solar panels were considered as made of thin wire [21]. In the VSTL model dimension of the X, Y and Z are taken $140m$, $45m$ and $30m$, respectively. The earth of size $10m$ is considered in the Z axis. The ground resistance is considered to be 100Ω. The PV panels are considered to be placed on the $4m \times 50m$ metallic frame. Frames are plotted by thin wire as suggested in [21]. To accommodate 600 PV panels , six sets of frames are required. These 6 frames are placed in two columns, and each column holds 3-row frames. The front height of the frame is $2m$, and back height is $3m$. The pillar of the frame is $1m$ underground. The DC lines from each frame array combine together and lead towards the inverter as such the six pairs of wire from the six rows lead towards the inverter. The voltages are measured from the DC lines to ground. The lightning current source is placed in a different location to see the effect of the direct strike on the ground. Additionally, a set of air-termination rods and earthing electrodes have been plotted with thin wire as illustrated in Fig. 4.

III. Results and Analysis

The direct and indirect lightning strikes have been investigated to determine their effect on the Solar Power generation with and without the proposed LPS. For different locations of the lightning strike, the induced voltages are calculated. We observed that a strong magnetic field is being created in around the lightning strike. Large voltage induces in the PV frame and DC line leading to the inverter. The induced voltage depends on some parameters such as the location of the lightning, the rate of rising of lightning current, earth-resistance, *etc*. Due to the rise of the earth-resistance, the induced voltage increases, which can be an important criterion of site selection for a solar power plant. On the other hand, due to the induction of a lightning protection system, the induced voltage reduces. This is explained below with necessary illustrations.

A. Lightning Effect on DC Lines with and without LPS

We observed that, with an LPS, the induced voltages is much reduced. Fig. 5 shows that, without any earthing system, the maximum induced voltage is $1250KV$. But in case of employing an LPS, the voltage reduces to $650KV$ as illustrated in Fig. 6. We also observed that the induced voltage quickly dies to zero in case of LPS. High voltage for longer time produces heat, which can burn the DC lines. Similar decreasing trend is also evident in the induced current in presence of the LPS as illustrated in Fig. 7 and Fig. 8. So, our findings demonstrate how a proper LPS would help for the safety protection of a large scale solar power plant.

B. Lightning Effect on PV Frames with and without LPS

The metal frame which holds the PV panels is the main structure of a solar power plant. Due to lightning, a high voltage induces on the frames. This high voltage can damage

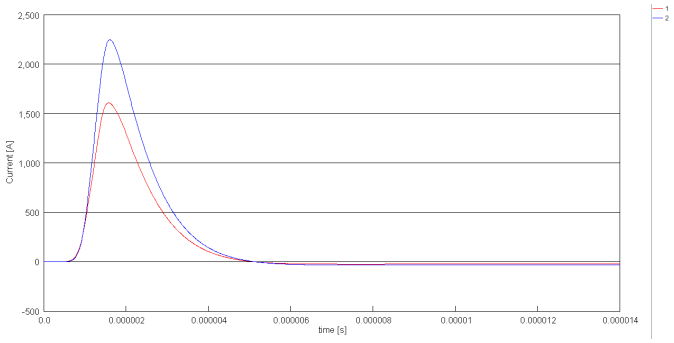


Fig. 7. The induced current in the DC lines without any LPS

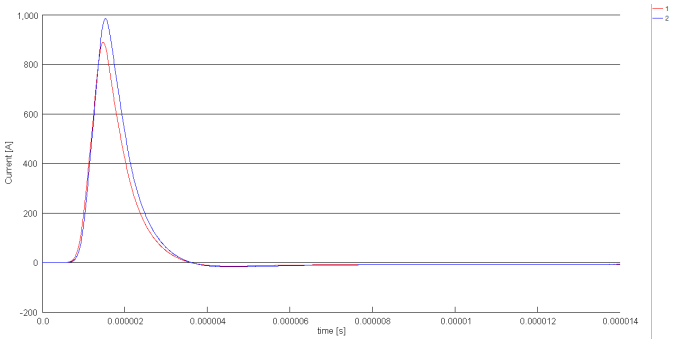


Fig. 8. The induced current in the DC lines with the proposed LPS

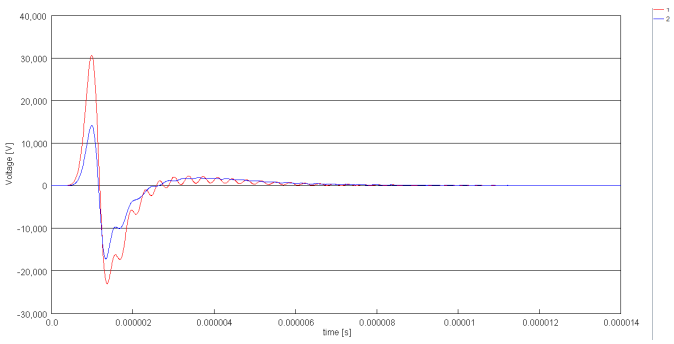


Fig. 10. The induced voltage in the frame with LPS

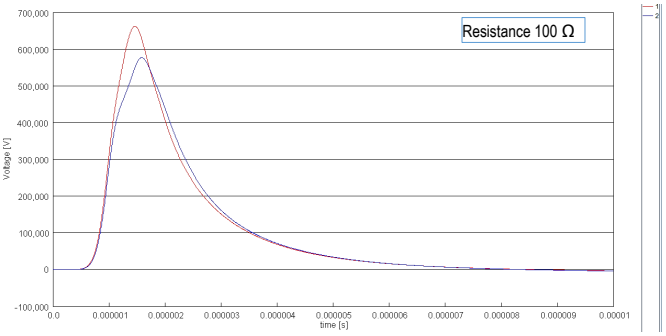


Fig. 11. The induced voltages with the earth-resistance of 100Ω

the PV panels and its wiring. In Fig. 9, it is observed that 74KV and 60Kv voltages induced in two frames. These induced voltages, however, reduced to 30KV and 12KV, when we introduced LPS as illustrated in Fig. 10. The induced voltages in the frames, however, can be further reduced with a more appropriate LPS.

C. Effect of Altering Soil Resistivity

The induced voltage due to the lightning effect also depends on the soil property. Fig. 11 and Fig. 12 show the induced voltage of DC lines leads to the inverter. As illustrated in Fig. 11, we observed that the peak induced voltage is 650KV, when the resistivity of the earth is 100Ω. In contrast, the Fig. 12 illustrates that the result of the induced voltage for the earth-resistance is 5000Ω. The induced voltage is as high as approximately 1000KV, which means that, with the increase of the soil resistivity, the induced voltage due to the lightning also increases. We also note that, at the low soil resistivity, the voltage rises to the peak and then within a few microseconds, it dies down to zero. On the other hand, for a higher soil resistivity, the voltage will not die down; instead, it fluctuates for some times. This dependence on the soil resistivity indicates that the soil resistance is an essential factor for the site selection of a large scale solar power plant. Wetland or swampy land is better than sandy or dry land. So solar power panels can be placed over the swampy land or at the island of a river or sea.

D. Recommendations for Lighting Protection

Considering the massive consequences, lighting has recently been announced as a natural disaster in Bangladesh. The safety protection of the large solar power plants in Bangladesh is therefore a more essential need than the other countries. To

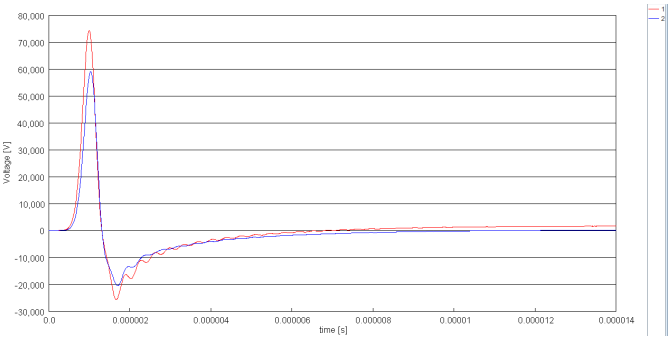


Fig. 9. The induced voltage in the frame without LPS

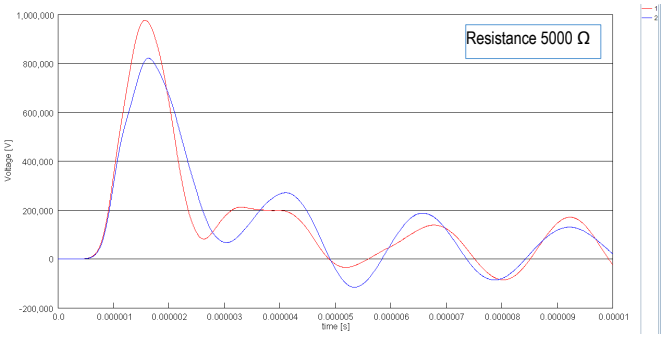


Fig. 12. The induced voltages with the earth-resistance of 5000Ω

design a large scale solar power plant and an appropriate LPS for it, the following recommendations can be made from our above findings and analysis:

1) A dedicated pre-analysis is to be carried out for establishing a large solar power plant to determine the requirements of an appropriate LPS for the plant. A conventional LPS used for the other installations or systems could be inadequate for the required safety protection.
2) An appropriate LPS system should consist of both the set of air-termination and earth-termination systems.
3) Since the low earth-resistance reduces the lightning effect, the solar power plant can be planned in the wet or swampy land.

IV. Conclusions

In this paper, modeling and analysis of a lighting protection system (LPS) is presented for a large scale solar power plant. Characteristics of induced voltage and current due to a lightning strike on the solar power plant has been studied using VSTL simulation software under different conditions, which provides knowledge for designing an appropriate LPS. The protection system for a large scale solar power plant is not yet standardized in many countries like Bangladesh. In this study, we observed that a properly designed LPS can significantly reduce the induced voltage in an acceptable range and thus can protect the PV panels, the frames and DC lines from both the direct and indirect lightning effect. We also found that the induced voltage due to lightning effect increases with the increase of soil resistivity. The location of large scale solar plant should therefore be chosen at the wet or swampy land, which offers low earth-resistance.

The analysis and findings presented in the paper are the outcome of our initial study towards designing an effective LPS system for a large scale solar power plant. Further study may be the choice and use of appropriate metal for air-termination rod and earthing electrodes to reduce the cost of LPS.

References

[1] D. Akinyele, R. Rayudu, and N. Nair, "Development of photovoltaic power plant for remote residential applications: The socio-technical and economic perspectives," *Applied Energy*, vol. 155, pp. 131–149, 2015.

[2] M. Hosenuzzaman, N. A. Rahim, J. Selvaraj, M. Hasanuzzaman, A. A. Malek, and A. Nahar, "Global prospects, progress, policies, and environmental impact of solar photovoltaic power generation," *Renewable and Sustainable Energy Reviews*, vol. 41, pp. 284–297, 2015.

[3] K. Neuhoff, "Large-scale deployment of renewables for electricity generation," *Oxford review of economic policy*, vol. 21, no. 1, pp. 88–110, 2005.

[4] "Energy Efficiency and Conservation Master Plan up to 2030," Ministry of Power, Energy and Mineral Resources, Government of the People's Republic of Bangladesh, Tech. Rep., 03 2015. [Online]. Available: https://goo.gl/83omfv

[5] "Renewable energy policy of bangladesh," Power Division, Ministry of Power, Energy and Mineral Resources, Government of the People's Republic of Bangladesh, Tech. Rep., 12 2008. [Online]. Available: https://goo.gl/w391ed

[6] N. Zaini, M. Ab-Kadir, M. Izadi, N. Ahmad, M. M. Radzi, N. Azis, and W. W. Hasan, "On the effect of lightning on a solar photovoltaic system," in *33rd International Conference on Lightning Protection (ICLP)*. IEEE, 2016, pp. 1–4.

[7] T. McDermott, T. Short, and J. Anderson, "Lightning protection of distribution lines," *IEEE Transactions on Power Delivery*, vol. 9, no. 1, pp. 138–152, 1994.

[8] A. Piantini, "Lightning protection of overhead power distribution lines," in *29th International Conference on Lightning Protection, Uppsala, Sweden*, 2008, pp. 1–29.

[9] N. Itamoto, H. Kawamura, K. Shinjo, H. Motoyama, and M. Ishii, "Accuracy of lightning surge analysis of tower surge response," in *International Conference on Power Systems Transients*, 2009.

[10] H. Kono, M. Fujino, M. Yokoyama, K. Yonezawa, Y. Takahashi, C. Isokawa, and A. Tatematsu, "Applying FDTD simulation to lightning surge route analysis in microwave relay stations," *Journal of International Council on Electrical Engineering*, vol. 1, no. 1, pp. 60–66, 2011.

[11] M. R. Ahmed and M. Ishii, "Effectiveness of interconnection of wind turbine grounding influenced by interconnection wire," in *2012 International Conference on Lightning Protection (ICLP)*. IEEE, 2012, pp. 1–6.

[12] ——, "Electromagnetic analysis of interconnected groundings at wind farm," *IEEJ Transactions on Power and Energy*, vol. 132, no. 7, pp. 684–689, 2012.

[13] J. C. Hernandez, P. G. Vidal, and F. Jurado, "Lightning and surge protection in photovoltaic installations," *IEEE Transactions on power delivery*, vol. 23, no. 4, pp. 1961–1971, 2008.

[14] M. A. Hossain and M. R. Ahmed, "Study and protection of lightning overvoltage on DC cables of solar power generation," in *Electrical and Computer Engineering (ICECE), 2014 International Conference on*. IEEE, 2014, pp. 325–328.

[15] M. Rabbani, A. Oo, M. Ahmed, and M. Kabir, "Analysis of lightning induced effects on buried pipeline due to direct strike on ground," in *Power Engineering Conference (AUPEC), 2016 Australasian Universities*. IEEE, 2016, pp. 1–5.

[16] C. A. Charalambous, N. D. Kokkinos, and N. Christofides, "External lightning protection and grounding in large-scale photovoltaic applications," *IEEE transactions on electromagnetic compatibility*, vol. 56, no. 2, pp. 427–434, 2014.

[17] M. R. Karim and M. R. Ahmed, "Analysis of electromagnetic induction due to lightning on a large-scale solar power generation," in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. IEEE, 2019, pp. 1–6.

[18] T. Noda and S. Yokoyama, "Development of surge simulation code based on finite-difference time-domain (FDTD) approximation of maxwell's equations," in *Proc. Int. Conf. on Power Systems Transients (IPST)*, 2001.

[19] V. P. Idone and R. E. Orville, "Lightning return stroke velocities in the thunderstorm research international program (TRIP)," *Journal of Geophysical Research: Oceans*, vol. 87, no. C7, pp. 4903–4916, 1982.

[20] F. Heidler, "Traveling current source model for LEMP calculation," in *Proc. 6th Int. Zurich Symp. Electromagnetic Compatibility, Zurich, Switzerland (1985-3)*, 1985.

[21] Y. Baba, N. Nagaoka, and A. Ametani, "Modeling of thin wires in a lossy medium for FDTD simulations," *IEEE Transactions on Electromagnetic Compatibility*, vol. 47, no. 1, pp. 54–60, 2005.

Analysis of NCFET with Si Doped HfO_2 as Ferro-Electric Material to Achieve Lower SS

Towfiqur Rahman
Dept. of EEE
BRAC University
Dhaka, Bangladesh
arttra.02@gmail.com

Antara Fairooz Khan
Dept. of EEE
BRAC University
Dhaka, Bangladesh
antarafairoozkhan@gmail.com

Nowshin Nawal
Dept. of EEE
BRAC University
Dhaka, Bangladesh
nowshin.nawal@gmail.com

***Abstract*— By using ferroelectric material as gate insulator in MOSFET (NCFET), Sub-Threshold Swing can be reduced below the fundamental limit of 60mV per decade. Achieving the negative capacitance depends on the ability to drive ferro-electricity away from its local minimum to non-equilibrium state which is very unstable. It can be stabilized by adding a series capacitance. Hence, a ferroelectric MOS capacitor can be modeled by a series connection of a ferroelectric capacitance and a channel capacitor to get a steeper I_d-V_{gs} curve than a regular MOSFET.**

Keywords—Negative Capacitance MOSFET, Ferroelectric, Hafnium Oxide.

I. INTRODUCTION

The rapid development of technology is commensurate with smaller and more efficient devices. In order to deal with the constant demand for smaller devices, we need smaller transistors. And CMOS being the dominant technology for integrated circuits, has served significantly to this scaling trend. CMOS can be scaled down by reducing its drain to source length (channel length) and the smaller the dimension the higher the performance achieved [1]. Consequently, today's CMOS occupies less than 1% of space compared to the initial transistors while being 20 times faster [2]. Ferro-electricity was first noticed in 1920 but the concept of negative capacitance is not very old. S. Salahuddin et al. first published that negative differential capacitance can be used as gate insulator in MOSFET that can lower the SS value below fundamental limit in 2008 [3]. Later in 2011 A. I. Khan with his fellow researchers demonstrated the effect of negative capacitance in a ferroelectric dielectric hetero-structure using a bilayer of ferroelectric $Pb(Zr_{0.2}Ti_{0.8})O_3$ and dielectric $SrTiO_3$ [4]. C. W. et al. achieved sub-30mV/decade SS in non-hysteretic NCFET using body profile engineering [5]. They also showed that NCFET works best in ultrathin body using TCAD simulation [6]. The first fabrication of NCFET was done in 2014 by C.W. Yeung and his colleagues [7]. Later in 2014 NCTFET was proposed by N. Chowdhury et al. [8] and a direct measurement of negative capacitance in an isolated ferroelectric capacitor was reported by A. I. Khan and his colleagues in 2015 [9]. In the years of 2016-2017 enormous amount of research has been conducted on NCFET. Among them simulation of first NC-FinFET [10], a spice compatible model for NC-FinFET [11],

Table I: Scaling of Parameters

SL.	Constant Field Scaling		
	Description	***Parameter***	***Scaling***
1	Device Dimension	L, W	1/S
2	Oxide thickness	T_{OX}	1/S
3	Channel Doping	N_{SUB}	S
4	Power Supply	V_{DD}	1/S
5	Junction Depth	X_J	1/S

negative capacitance in polycrystalline HfO_2 based thin film [12], a simulation model of NCFET with hysteresis edges less than 10mV/decade up to seven order of magnitude and ultra-low energy delay overhead (below 2.1%) [13], significant improvement in DIBL and NDC (Negative Output Differential Conductance effects) using an ultrathin quantum metallic layer [14], experimental demonstration of ferroelectric HfZrOx (FE-HZO) NCFET with high temperature annealing (hysteresis free) [15], NC MOSFET incorporating with ferroelectric Al doped HfO_2 [16] are some of the significant works.

However, the reduction of channel length generates some undesirable changes in the device characteristics. The most popular constant field scaling scheme (proposed by Robert Dennard and his team) to maintain qualitative performance of MOS requires the oxide thickness and depletion width to be decreased by the same factor as the supply voltage and all the electric field within the device remains constant [Table I]. To keep the power density constant, supply voltage and threshold voltage (V_T) need to be reduced proportionally. On the other hand, to keep the power consumption low, high threshold is required to limit the leakage current. Also the supply voltage must be significantly above the threshold

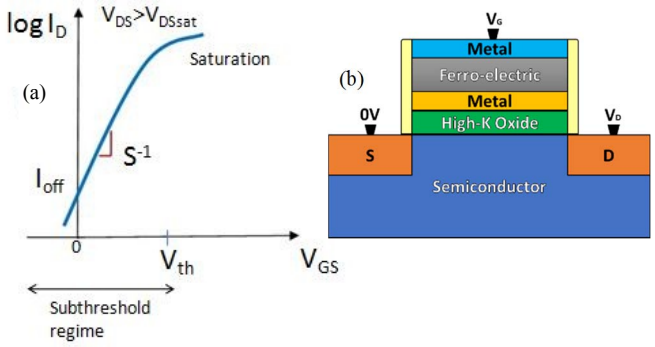


Fig. 1: (a) Here, showing the exponentially decreasing drain current, I_D in the sub-threshold regime known as sub-threshold current. (b) Basic structure of NCFET using ferroelectric material as gate insulator [9].

voltage which eventually limits the power supply to decrease enough to follow the constant field scaling method. Later it was observed that the smaller the device, the more the thermal noise and bit flipping increases in the conventional MOSFET structure. There are also two major types of power dissipation in CMOS - Static (leakage current in the sub-threshold region) and Dynamic (power consumption due to the charging and discharging of the output capacitive load) power dissipation [17].

For simple MOSFET it is assumed that the drain current, I_D is 0 when $V_{GS}<V_T$. However, in reality I_D decreases

exponentially to 0 when $V_{GS}<V_T$ resulting in off state or sub-threshold current. The slope of the sub-threshold current can be shown as $d(V_G)/d(\log I_D)$. This current depends on the depletion capacitance of the channel. Mathematically sub-threshold slope (SS) is defined as,

$$SS=\frac{d(V_G)}{d\log(I_D)}=\left(\frac{kT}{q}\ln 10\right)\left(1+\frac{C_D}{C_{OX}}\right)$$

C_D is the depletion capacitance in the channel. At room temp. (T=300K) substituting $k = 1.3806\times10^{-23} m^2 kg\ s^{-2}\ K^{-1}$, $q = 1.6\times10^{-19}$ C.

$$SS=\left(60\times10^{-3}\right)\left(1+\frac{C_D}{C_{OX}}\right)$$

This means at room temperature with the best case where $C_D \ll C_{OX}$, theoretically the lowest SS can be 60mV/decade. The less the I_{ON} and I_{OFF} ratio, the less the leakage current will be. Therefore, for a high performance this ratio needs to be as low as possible and the SS to be steeper. Since some of the fundamental parameters of MOSFET are reaching its limit, introduction of new structures are necessary. Different transport mechanisms such as impact ionization, tunneling and positive feedback can be applied to improve SS by amplifying the gate voltage [5].

II. Device Structure

With a view to studying the electrical performance of the negative capacitance MOSFET, a metal-FE-metal-insulator-semiconductor (MFMIS) structure is used to approximate the MFIS structure, as shown in Fig. 2. The suspended gate metal is assumed to have the same work-function as the top gate metal, and there are no net charges on the suspended gate. It is obvious that for $V_{ds} = 0$ V or $t_f = 0$ nm, the MFMIS structure is completely reduced to the MFIS device and there are no errors for this approximation. Even though there are some reasonable errors (<10%) depending on specific circumstances for large V_{ds} and t_f values, the errors can be ignored for small V_{ds} or t_f values. And, the latter is the typical case (low power supply and ultra-small dimension) of state of the art nano-scale devices being fabricated. Dynamic nonlocal path band to-band model, standard Shockley-Reed-Hall recombination and drift-diffusion model are used and Fermi-Dirac statistics is assumed in the MOSFET simulation. This hybrid 2D electrostatics for MOSFET and 1-D Landau equations are solved self-consistently.

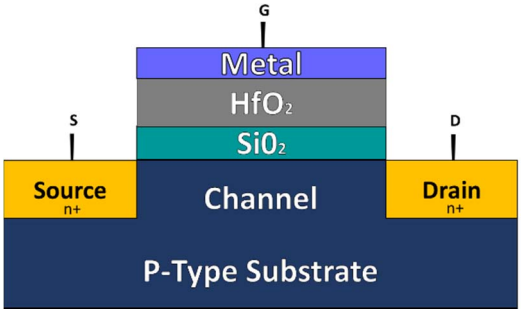


Fig. 2: Device structure of proposed NCFET showing different capacitances.

A. Working Principal

In the ferroelectric MOS capacitor, the ferroelectric polarization charge density and the channel charge density should match. Therefore, a static operation point is determined using the cross point of P-E curve and the channel charge load line. The negative slope of the P-E curve

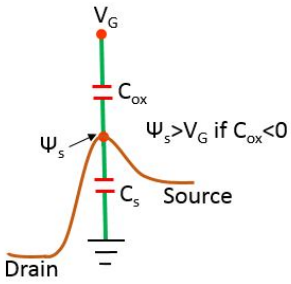


Fig. 3: Schematic diagram of an NCFET.

indicates negative capacitance region. If the two curves have a single cross point in negative capacitance region, the ferroelectric MOS transistor works as NCFET. If the thickness of the FE is too thick, the static operation point cannot be properly set and it shows hysteresis. However, it can be avoided by choosing suitable thickness [18]. For negative capacitance, $\frac{\partial V_g}{\partial \psi_s}=1+\frac{C_S}{C_{OX}}$ can be less than 1 as C_{ox} is negative. Therefore, NCFET seems to have a good potential to work below the SS limit [Fig. 3].

B. Fabrication Process

The fabrication technique of NCFET is a long process that starts with source and drain doping and activation. Then an SOI wafer is thinned down by using dry thermal oxidation and DHF etching. After that active area definition is done followed by the removal process of defects – caused by etching – using thermal oxide. Next steps are wafer patterning, source/drain ion implementation, activation of dopants (using rapid thermal annealing). Then blanket STO buffer is deposited using Molecular Beam Epitaxy (MBE) along with native oxide removal. The wafer is then cut into 10mm by 10mm dies and another blanket layer of ferroelectric is deposited. Finally, the gate metal is deposited by e-beam deposition and using a dry-wet etch process source/drain contact is etched.

C. Si doped HfO_2 as Ferroelectric Material

With enhance research in the nano-field to overthrow the classical limit of 60mV/decade, Salahuddin [19] came with the idea of negative capacitance field effect transistor which has the same structure of a normal MOSFET and the only difference in building criteria is to replace the gate insulator oxide with thin film of ferroelectric material. The nonlinear characteristic of ferroelectric material which is known as negative capacitance plays a vital role to exceed the value of $\partial\psi_s/\partial V_g$ more than 1. While in non-equilibrium, a ferroelectric material can show negative differential capacitance due to its energy barrier that actually forms during the phase transition and separates the two degenerate polarization states. Though the state of negative capacitance is unstable, by using a series die-electric capacitor the negative differential capacitance state can be stabilized.

Here it may be mentioned that negative capacitance has already been experimented on thin ferroelectric film [20]

which surely indicates a new way in the semiconductor industry. It is well known that the mechanism that works behind electron transport system in MOSFET is drift and diffusion and this very same technique also works as the working principle in negative capacitance MOSFET. NCFET can drive higher current than tunnel FET due to its higher charge density on the interface which works as an amplifier in the same off-current. Actually ferroelectric material works as a 'transformer' to boost up the input voltage. As a result, NCFET can successfully exceed the drive current of conventional MOSFET as well as it can surely overcome the limit of 60mV/decade. Moreover it is interesting to note that pass-gate CMOS circuits which actually flow the electricity in both ways between drain and source and standard pull-up and pull-down CMOS logic circuits both can be built by negative capacitance FET. So there is no reason to hugely modify the logic circuit architecture and IP macro design from the conventional MOSFET as it will be simply replaced by the new NCFET which got significant impact on semiconductor industry yet it has got some limitations too. The main barrier comes on the process of integration. In order to balance the large polarization charge density and FET charge density, the thickness of the insulator layer needs to be more than hundreds nanometer in case of traditional ferroelectric material such as Lead Zirconate Titanate and Barium Titanate. And here comes the real challenge as those are not compatible to advanced scale complementary MOS (CMOS) frond-end technologies. In addition to that, heavy metal contamination to the manufacturing line is also a big issue here in this case. Recently it has been found that by controlling crystalline phase of thin films of Hafnium Oxide, ferroelectricity has been observed in HfO_2.

With the 28 nanometer CMOS technology, thin films of Hafnium Oxide have already been implemented in Ferroelectric Random Access Memory (Fe-RAM) successfully [21]. Thereafter if the ferroelectric thin film of Hafnium Oxide is used in NCFET, the fabrication process will fully compatible with CMOS technology. Just by introducing ferroelectric HfO_2 gate insulator NCFET can be fabricated by any CMOS technology from the mature to most advanced technology in a cost efficient way.

III. Numerical Modeling

As discussed earlier, NCFETs are implemented by adding a ferroelectric material layer in the gate stack of MOSFETs. A lower SS is achieved in NCFETs with the help of "negative capacitance" which is contributed by the deposited ferroelectric layer. Ferroelectrics are materials with a high polarizability while ferroelectric capacitors tend to exhibit nonlinear capacitance, where the change in voltage produces a negative change in polarizability. The free energy of ferroelectric is represented by the Landau Theory and the electrostatic dynamics of this negative capacitance can also be well described by LK (Landau-Khalatnikov) equation:

$$\rho\frac{d\vec{P}}{dt}+\nabla_i U=0 \qquad (1)$$

Where, ρ is material dependent parameter that accounts for dissipative process during the ferro-electric switching, P is the polarization charge in per unit area and U is the Gibbs free energy. We can express it in terms of anisotropy energy:

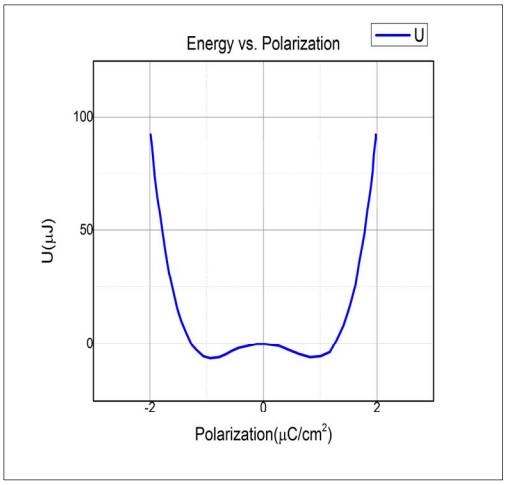


Fig. 4: Energy vs. Polarization.

$$U=\alpha P^2+\beta P^4+\gamma P^6-\vec{E}_{ext}\vec{P} \qquad (2)$$

Here the energy due to the external field is $\vec{E}_{ext}$ and α, β and γ are the material dependent parameters. In general, at zero bias the U vs P graph looks like a W like the one in Fig. 4. which tilts due to increase in V_g. From (1) and (2) we can write that:

$$E_{ext}=2\alpha P+4\beta P^3+6\gamma P^5+\rho\frac{dP}{dt} \qquad (3)$$

here the vector signs have been dropped assuming the spatial variations to be one-dimensional. It is easy to see that if we set, at steady state operation $dP/dt \approx 0$ and for FE material $P \approx Q$ while the external electric field is $E_{ext} = V/t_{ins}$ them taking all these into consideration and mingling (2) and (3) we get,

$$V=2\alpha t_{ins}Q+4\beta t_{ins}Q^3+6\gamma t_{ins}Q^5 \qquad (4)$$

where, t_{ins} is the ferroelectric insulator thickness and α, β and γ are the material dependent parameters. For, HfO_2 the values of these parameters are -1.5877 x 10^{11} m/F, 9.8007 x 10^{20} $m^5/F/Coul^2$ and 0 $m^9/F/Coul^4$ respectively. The plot of the steady-state polarization P versus E_{ext} and V vs P are shown in Fig. 5 and Fig. 6, were obtained from (3) and (4).

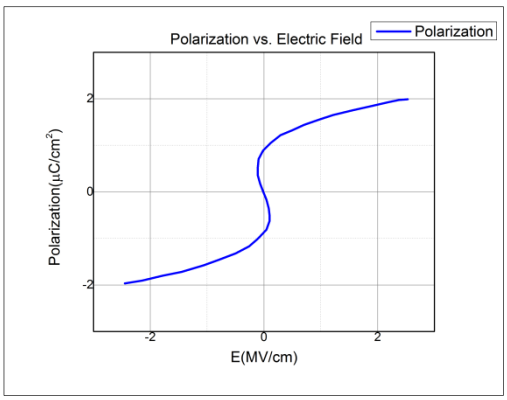


Fig. 5: Polarization vs. Electric Field.

We will consider that the FET has a ferroelectric insulator rather than an ordinary insulator at the gate. So, the gate circuit can be assumed to be a series combination of the ferroelectric capacitor and another capacitor that we will call C_s, representing semiconductor, channel-to-source and channel-to-drain capacitances. For this series combination, a voltage ψ_s appears across C_s, while the rest V_g-ψ_s appears across the ferroelectric such that both have the same charge Q and we can write:

$$\psi_s = Q/C_S \tag{5}$$

$$V_g - \psi_s \approx \alpha_0 Q + \beta_0 Q^3 + \gamma_0 Q^5 + \rho_0 \frac{dQ}{dt} \tag{6}$$

Considering $\alpha_0 = 2\alpha t_{ins}$, $\beta_0 = 4\beta t_{ins}$, $\gamma_0 = 6\gamma t_{ins}$

Reconsidering (3) and by combining these expressions, we can relate the applied voltage V_g to the potential ψ_s that appears inside the channel:

$$\tau \frac{d\psi_s}{dt} + (1+a_1)\psi_s + a_2\psi_s^3 + a_3\psi_s^5 = V_g \tag{7}$$

Where $\tau = \rho C_S t_{ins}$. By setting $d\psi_s/dt = 0$, the steady-state ψ_s versus V_g is obtained from (7) which is

$$(1+a_1)\psi_s + a_2\psi_s^3 + a_3\psi_s^5 = V_g \tag{8}$$

There is no doubt that the nature of the ψ_s vs V_g characteristics will very much depend on the three co-efficient a_1 and a_2 and a_3 which are scaled versions of the material parameters α, β and γ. If the parameter values are such that the nonlinear terms a_2 and a_3 are very small

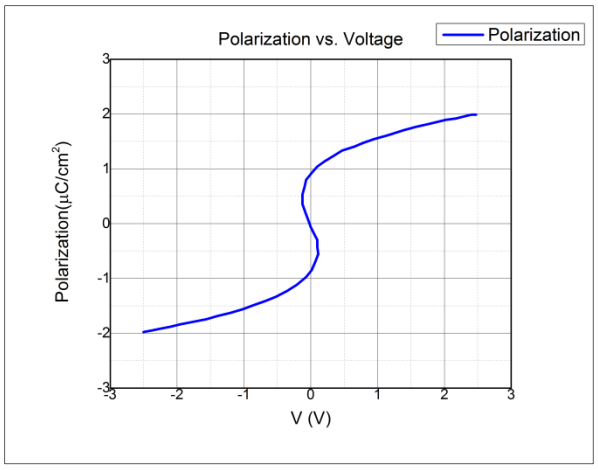


Fig. 6: Polarization vs. Oxide Voltage.

compared to a_1, which only ensures the relation between ψ_s and V_g is linear then

$$\partial V_g/(\partial \psi_s) \approx 1 + a_1 \tag{9}$$

This is same as $\frac{\partial V_g}{\partial \psi_s} = 1 - \frac{C_s}{C_o}$ $(\alpha_f C_o - 1)$ which represents the negative capacitance while it seems that we are neglecting the nonlinear terms. But the point to be noted that even though a_2 and a_3 are non-negligible, $(1+a_1)$ represents the slope close to the origin $\psi_s = 0$ & $V_g = 0$ and should be positive if we wish the origin to represent a stable operating point. With the help of a_1, it is easy to find that the condition $(1+a_1) > 0$ requires the thickness of the ferroelectric insulator to be less than a critical thickness defined as:

$$t_{ins} \leq \frac{1}{2|\alpha| C_s} \equiv t_c \tag{10}$$

The above condition confirms that the ferroelectric capacitor is large enough or the series capacitor C_s is small enough that the combination forms a stable positive capacitor, and there is no hysteresis at the origin unlike Fig. 7, which corresponds to a very thin ferroelectric. All results we present below assume $t_{ins} < t_c$, because a hysteresis at the origin seems undesirable for our device applications.

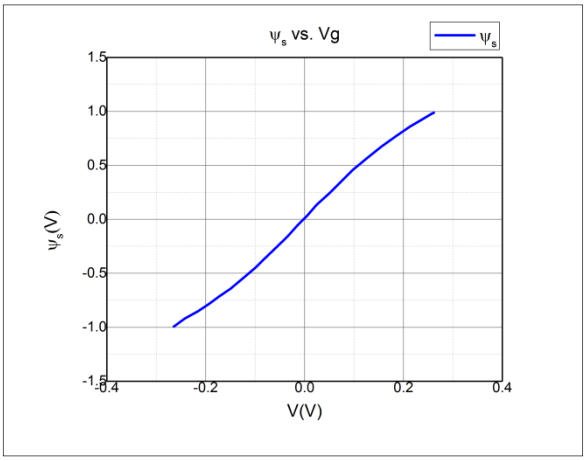


Fig. 7: ψ_s vs V_g

Fig. 7 shows ψ_s vs V_g for different values of t_{ins}, directly obtained from (8) using appropriate values of α, β and γ for HfO_2, in series with a linear capacitance $C_s = 25 \times 10^{-7}$ F/cm^2. The plots are obtained starting from an initial state with the ferroelectric capacitor fully negatively polarized, sweeping the voltage toward the positive maxi-mum, and then sweeping it back again to the negative maximum. Each point is obtained by solving the full-time dependent (1) and then finding ψ_s as the system reaches a steady state. Here also one thing may be noted that the insulator thicknesses we use are less than the critical thickness.

As the insulator thickness is increased, the ψ_s vs V_g plots become progressively steeper but at the same time, they start to show more and more hysteresis away from the origin but not at the origin due to the nonlinear terms a_2 and a_3. With more insulator thickness there is a hysteresis in ψ_s vs V_g away from the origin, but it is small. In general, a trade off will take place between steepness and hysteresis, unless materials can be engineered to minimize the nonlinear term β relative to the linear term α so as to avoid hysteresis even at large steepness.

It will be noted from a_1 and a_2 and a_3 that if C_s were smaller, the nonlinear terms a_2 and a_3 would become smaller in comparison to the linear term a_1 making the hysteresis negligible in our voltage range of interest. The value for C_s that we have used in this paper is more appropriate for ultra-small FETs [22] where the short

channel effects are significant. For present day commercial device, the value of C_s should be roughly a factor of 10 lower than that used in this paper and as such should show steep ψ_s vs V_g with negligible hysteresis. However, because the critical thickness is inversely proportional to C_s, the thickness of the ferroelectric insulator will have to be roughly a factor of 10 larger.

In terms of a free energy functional, the voltage amplification arises from an internal interaction commonly described by the Landau theory. Such interactions are quite general in metal oxides and commonly lead to the development of a spontaneous order parameter below a transition temperature. For example, the giant enhancement of longitudinal piezoelectric response recently modeled using the Landau-Ginzburg-Devonshire theory can be attributed to similar mechanisms, the precise correspondence between the effective field in the Landau approach in (3) and the effective voltage in a capacitor with positive feedback suggests that any microscopic mechanism that can provide the necessary positive feedback can be used to implement a step-up voltage transformer that would allow low-voltage/low-power operation of conventional FETs. For a ferroelectric capacitor, it is the dipole interaction that provides the positive feedback while avalanche breakdown, polaronic effect, etc. within the oxide can possibly provide the positive feedback needed for negative capacitance as well. The parameters α_o, β_o , γ_o and ρ_o will be determined by the specific mechanism involved.

A. Negative Capacitance Region

Capacitance C at a given charge Q_F is related to the potential energy U by the relation $C=[\frac{d^2U}{dQ_F^2}]^{-1}$. The parabolic graph is in Fig. 8. So, we can call the negative capacitance as negative differential capacitance. For a ferroelectric material, as shown in Fig. 8, the capacitance is negative only in the barrier region around P=0 [Considering $Q_F \propto P$]. Starting from an initial state,

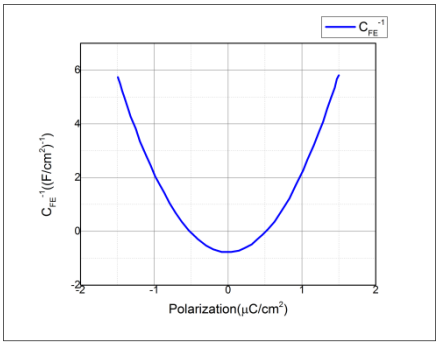


Fig. 8: Capaitance^{-1} vs Polarization

as voltage is applied across the ferroelectric capacitor, the energy landscape is tilted and the polarization will move to the nearest local minimum. Fig. 9 (a), Fig. 9 (b) are the graph for less than the coercive voltage whereas Fig. 9 (c), Fig. 9 (d) show this transition for a voltage that is greater than the coercive voltage V_c. If the voltage is larger than coercive voltage, one of the minima disappears and P moves to the remaining minimum of the energy landscape showed in the figure. Here may be noted that, as the polarization state descends, it passes through the region where $C=[\frac{d^2U}{dQ_F^2}]^{-1}<0$. Therefore, while switching from one stable polarization to the other, a ferroelectric material passes through a region where the differential capacitance is negative. And this is the region where the device we are working on should work.

In case of P-E curve where the slope is negative, we find the negative capacitance but it is unstable and not directly observed in experiments when it exhibits hysteric jumps in the polarization. Here is a solution for the unstable condition, if we place a ferroelectric capacitor series with a normal capacitor then the negative capacitance segment can be effectively stabilized. It is possible by making ψ_s on an internal node to change more than the voltage applied V_g externally working as a voltage amplifier also known as step-up voltage transformer.

The structure of the device may be a bulk Si MOSFET with high K gate oxide is referred to as the intrinsic MOSFET. FE oxide is assumed to be deposited on a commensurate metallic template grown on the oxide. The intermediate metallic layer between the FE and the high-K gate oxide is chosen following the experimental device. This layer averages out the non-uniform potential profile along the source-drain direction as well as any charge non-uniformity coming from domain formation in the ferroelectric. This experimental structure therefore justifies using a 1-D Landau model for the ferroelectric.

IV. Result

The intrinsic MOSFET is simulated using TCAD Sentaurus software taking into account the 2-D electrostatics. NCFET with an MFIS structure cannot be simulated directly even with state-of-the-art device simulators such as Sentaurus and Silvaco TCADs, because of the lack of a corresponding accurate FE polarization model. In order to study the electrical performance of the NCFET, a metal-FE-metal-insulator-semiconductor (MFMIS) structure is used to approximate the MFIS structure. The suspended gate metal is assumed to have the same work-function as the top gate metal, and there are no net charges on the suspended gate. It is obvious that for V_{ds} = 0 V or t_f = 0 nm, the MFMIS structure is completely reduced to the MFIS device and there are no errors for this approximation. Even though there are some reasonable errors (<10%) depending on specific circumstances for large V_{ds} and t_f values, the errors can be ignored for small V_{ds} or t_f values. And, the latter is the typical case (low power supply and ultra-small dimension) of state-of the-art nano-scale devices being fabricated. Dynamic nonlocal path band to band model, standard Shockley-Reed-Hall recombination and drift-diffusion model are used and Fermi-Dirac statistics is assumed in the MOSFET simulation. This hybrid 2D electrostatics for MOSFET and 1-D Landau equations are solved self-consistently. One significant point is that, negative capacitance regions are ordinarily unstable and not observed in experiments. But placing such a capacitor in

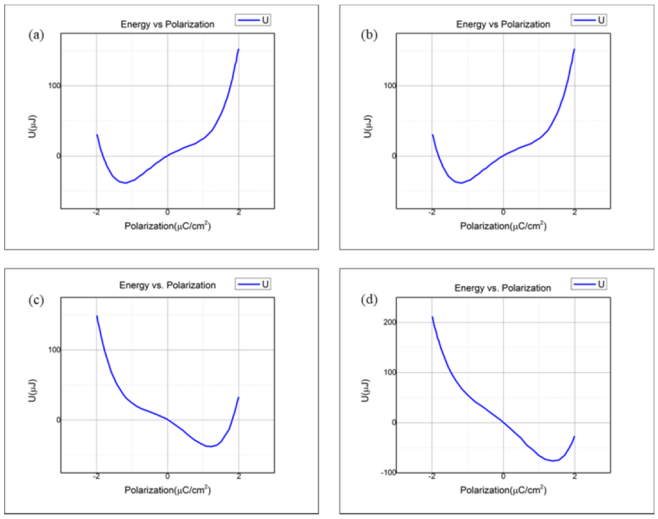


Fig. 9: Energy vs Polarization curve at (a) V_g= -6 V, (b) V_g = -3 V (c) V_g = 3 V & (d) V_g = 6 V.

series with a positive capacitor stabilizes it by making the effective α of the composite capacitor positive, provided its thickness is less than the critical thickness defined in (10). As R is proportional to (T-T_c) with T_c being the Curie Temperature, the series combination acts like a ferroelectric at a temperature above its T_c. After simulation, we get the I_D-V_{gs} curve in liner and in logarithmic scale as below which clearly shows that using Si doped HfO_2 as ferroelectric material in NCFET device provides much steeper curve of 30mV per decade than other regular MOSFET devices.

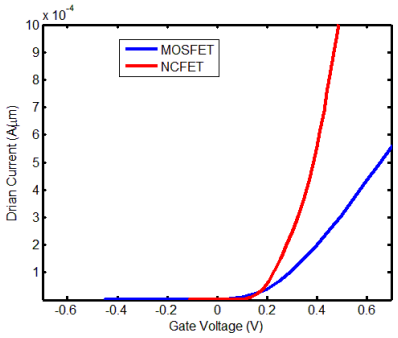


Fig. 10: Drain current in linear axis for MOSFET & NCFET. [V_{ds} = 0.5V]

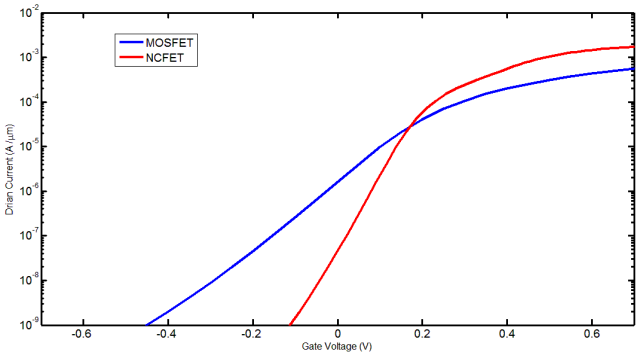


Fig. 11: Drain current in log. axis for MOSFET & NCFET. [V_{ds} = 0.5V]

V. CONCLUSION

HfO_2 as ferroelectric material has showed comparatively better performance than other materials in voltage amplification and low power dissipation. Besides, overall NCFET overcomes the limitations and boundaries faced by other FET devices. It may be said that HfO_2 based NCFET can lead to a better performance FET device in real in near future.

REFERENCES

[1] R. D. Isaac, "Reaching the limits of CMOS technology,"IEEE 7th Topical meeting on Electrical Performance of Electronic Packaging, 1998.

[2] S. Thompson et al.,"MOS Scaling: Transistor Challenges for the 21st Century;" Intel Technology Journal Q3'98.

[3] K. Gopalakrishnan et al., "I-MOS: A Novel Semiconductor Device with a Sub-threshold Slope Lower Than kT/q," IEDM Tech. Dig. 2002, pp. 289-290.

[4] A. I. Khan et al., "Negative Capacitance in a Ferroelectric Capacitor," Nature Materials, vol. 14, pp. 183, February 2015.

[5] G. McFardland and M. Flynn, "Limits of Scaling MOSFETs ", Technical Report CSL-TR-95-662, January 1995.

[6] S. Salahuddin and S. Datta, "Use of Negative Capacitance to provide voltage amplification for low power nanoscale devices," Nano letters 8, no-2, pp. 405-410, 2008.

[7] F. D'Agostino and D. Quercia, "Short Channel Effects in MOSFETs," December 2000

[8] A. I. Khan and S. Salahuddin, "Negative Capacitance in Ferroelectric Material and Implications for Steep Transistors," IEEE, 2015.

[9] K. Yamabe, "Time Dependent Dielectric Breakdown of Thin Thermally Grown SiO2 Films," IEEE Transaction on Electron Devices', 1985, pp.423.

[10] M. Kobayashi and T. Hiramoto, "On device design for steep-slope negative-capacitance field effect-transistor operating at sub-0.2 supply voltage with ferroelectric HfO2 thin film," AIP Advances, vol. 6, February 2016.

[11] R. A. McKee et al.,"Crystalline Oxides on Silicon: The First Five Monolayers," Phys. Rev. Lett. 81, 3014, 1998.

[12] A. I. Khan et al., "Experimental Evidence of Ferroelectric Negative Capacitance in Nanoscale Heterostructure, Applied Physics Letters, vo. 99, no. 11, 2011.

[13] A. I. Khan et al., " Ferroeletric Negative Capacitance MOSFET: Capacitance Tuning and Antiferroelectric Operation," IEEE International Electron Devices Meeting, 2011.

[14] C. W. Yeung et al., "Low Power Negative capacitance FETs for Future Quantum-well Body Technology," *VLSI-TSA, International Symposium*, 2013.

[15] N. Chowdhury et al., "Negative Capacitance Tunnel Field Effect Transistor: A Novel Device with Low Sub-threshold Swing and High ON Current," BUET,Dhaka, 2014

[16] C. W. Yeung, "Steep On/Off Transistors for Future Low Power Electronics," Engineering – Electrical Engineering and Computer Sciences, University of California, Berkeley, 2014

[17] R. H. Dennard et al.,"Design of Ion-Implanted MOSFET's with Very Small Physical Dimensions," IEEE Journal of solid state circuits, vol. 9, pp. 256-268, October 1974.

[18] C. Hu et al., "Green Transistor – A V_{DD} Scaling path for Future Low Power ICs," International Symposium on VLSI Technology, Systems and Applications, VLSI-TSA, 2008, pp. 14-15.

[19] S. Salahuddin and S. Datta, Tech. Dig. Int. Electron Devices Meeting (2008) pp. 693– 696.

[20] A. I. Khan, K. Chatterjee, B. Wang, S. Drapcho, L. You, C. Serrao, S. R. Bakaul, R. Ramesh, and S. Salahuddin, 14, 182-186 (2015).

[21] J. Müller, E. Yurchuk, T. Schlosser, J. Paul, R. Hoffmann, S. Müller, D. Martin, S.Slesazeck, P. Polakowski, J. Sundqvist, M. Czernohorsky, K. Seidel, P. Kücher, R. Boschke, M. Trentzsh, K. Gebauer, U. Schröder, and T. Mikolajick, in Tech. Dig.Symp. (VLSI Technology, 2012), pp. 25–26.

[22] Wakabayashi, H.; Tatsumi, T.; Ikarashi, N.; Oshida, M.; Kawamoto, H.; Ikezawa, N.; Ikezawa, T.; Yamamoto, T.; Hane, M.; Mochizuki, Y.; Mogami, T. Improved Sub-10-CMOS Devices with Elevated Source/Drain Extensionsby Tunneling Si-Selective-Epitaxial-Growth. Proc. IEDM 2005.

Large Magnetic Remanence Property Based a New Material Used in the Linear Generator for Wave Energy Extraction

Selim Molla and Omar Farrok
Department of Electrical and Electronic Engineering
Ahsanullah University of Science and Technology
Dhaka-1208, Bangladesh
eee.n_selim@yahoo.com and omarruet@gmail.com

***Abstract*—Almost all the conventional linear generators uses neodymium iron boron (NdFeB) permanent magnets (PMs) to generate magnetic field where neodymium is one of the rare earth elements (REEs). In this paper, a new alternative rare earth free (REF) PM is proposed to be used in the construction of a linear generator for wave energy extraction. The REF PM is a compound of iron nitride which is symbolized as $Fe_{16}N_2$. It is capable to produce strong magnetic field and has large magnetic remanence property. A linear generator is designed and simulated to compare the results for using both $Fe_{16}N_2$ compound and the conventional PM. The construction of the proposed linear generator, generated voltage, power output, armature current, and magnetic properties are shown in detail. Simulation results of the proposed linear generator are obtained by finite element analysis. ANSYS/Maxwell software is used to find excellent performances of the proposed linear generator.**

Keywords—cooling, energy conversion, generators, HVAC, permanent magnet machines, renewable energy sources, wave power.

NOMENCLATURE

i_a	Current of phase-A
i_b	Current of phase-B
i_c	Current of phase-C
i_d	Direct axis current
i_q	Quadrature axis current
I_m	Maximum value of armature current
ϖ	Angular frequency
α	Phase angle of i_a with respect to time origin
θ	Power factor angle

I. INTRODUCTION

Most of the electricity generation systems in the world are using fossil fuels which are well-known for carbon di oxide emissions [1]. This fact increases the atmospheric temperature for which the ice in the polar area is melting. On the other hand, the population is growing fast day by day. To overcome the challenges of meeting the electrical power demand, ensuring sustainable and green energy is very significant.

At present, many types of renewable energy are harnessed for the generation of electrical power for their environment friendly feature [2]. The renewable energies are wind, thermal, solar, and oceanic wave energy (OWE) [1]. Among these, the OWE has some advantages such as much availability, high power density, and long life span [3].

In order to achieve electric energy security, some countries have taken steps already for developing oceanic wave energy converter. China has established some projects to get 15% of energy from renewable sources [4] from which 21.79 GW electricity is expected to be generated from oceanic tidal and wave energy. US government looks into the affect of incentive and economic factors for the conversion of electrical energy from oceanic wave [5]. The European Union has targeted for the production of 100 GW electricity from sea wave and tidal energy within 2050 [6].

Recently, researcher's have presented extensive works in order to develop the oceanic wave energy converter (OWEC). A prototype of an overtopping type offshore OWEC is constructed in [7] which is consisted of automatic control panel and hydro turbine mechanism. A hydraulic system based OWEC is discussed in [8]. A direct drive linear generator (DDLG) is designed in [9] in which the electrical power output is enhanced using a high flux density magnetic core named Supermendur. Another DDLG is presented in [10] where N28H permanent magnet is used for withstanding in high temperature. A superconducting armature coil based linear generator is proposed in [11] which supplies more output power than that of using copper conductor. Another DDLG is presented in [12] where detent force is reduced by skewing the permanent magnets. A DDLG is proposed in [13] where its permanent magnets contain a rare earth element. Temperature effects of a DDLG during its operation are analysed in [14].

The rotational OWECs are presented in [7], [8] and DDLGs are presented in [9]–[14]. In general, DDLG is more popular topology than rotational OWEC owing to remove the complex mechanical structure [15]. Some drawbacks such as large weight and low power density due to low speed of sea wave are found from the conventional DDLG. The probable disadvantages of conventional DDLGs are solved in [16] by increasing the speed of movable magnetic field. Most of the conventional DDLG is consisted of neodymium iron boron (NdFeB) magnet [12]–[14] which is composed of a rare earth element and has limited stock. For this reason, it is necessary to search another type of magnet for the design and construction of linear generators. The compound $Fe_{16}N_2$ PM is found in [17] which has high magnetic energy product and can solve the drawbacks of NdFeB.

In this paper, $Fe_{16}N_2$ compound is proposed to construct a linear generator to avoid the use of rare earth elements (REEs) such as in NdFeB. $Fe_{16}N_2$ has the property of large magnetic remanence and energy product. The mathematical model, characteristics of the generator and performance are presented in detail. Simulation results show that, the proposed linear generator can generate adequate output power which is supplied to different loads to test its effectiveness.

II. Wave Energy Conversion System

The electricity generation system from sea wave is described in Fig. 1. It has some vital components such as linear generator, rectifier, inverter, transformer, and a floating buoy.

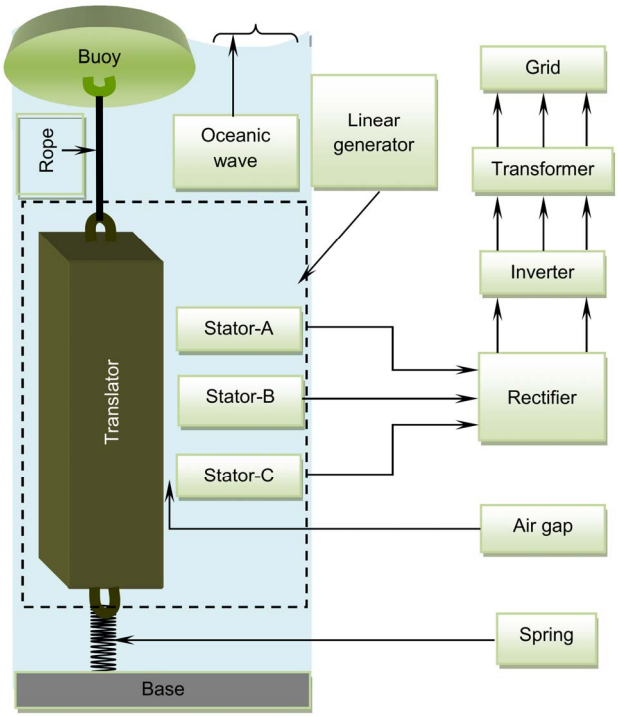


Fig. 1. Oceanic wave energy conversion system.

The effective elements of a linear generator are translator and stator. The translator moves with floating buoy as it is connected to a high strength rope. The translator speed is almost identical to the buoy speed whereas the stators are stationary. When the translator moves with respect to stators, emfs are induced in the stator windings. The frequency of generated voltage varies with the speed of translator/wave since most of the generators are synchronous. Therefore, the frequency of generated voltages are required to match the frequency of grid.

The generated voltages are supplied to the input of a rectifier so that, it can be converted into dc voltage. Further, the dc voltages are fed to the input of an inverter. The inverter is a power electronics device that converts dc voltage to ac voltage and capable to control both the voltage and frequency. A step up transformer is connected between the grid and inverter. The power electronic device such as phase locked loop takes signals from the grid and supplies required voltages with proper phase angle through the transformer by synchronizing it.

III. Comparisons of the Magnetic Characteristics

NdFeB permanent magnet consists neodymium which is a REE and it has much coercivity and remanence magnetism. Figs. 2 and 3 show the temperature dependencies of these two conventional NdFeB PMs, one is N28EH and the other is N35M. It is found that, both the PMs have high energy product at low temperature. With the increase in temperature, the coercivity becomes lower for both the PMs. Alnico PMs can withstand at higher temperature because it has high curie temperature. On the other hand, $Fe_{16}N_2$ compound has large remanence magnetism. Fig. 4 shows the hysteresis loop of $Fe_{16}N_2$ for different conditions and also its magnetic energy product.

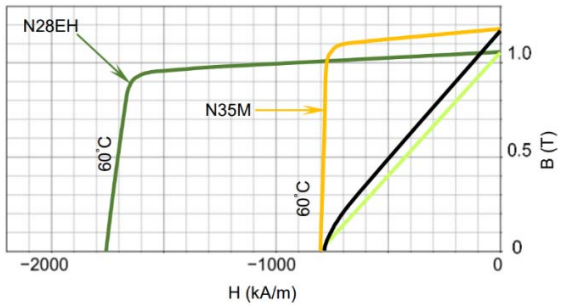


Fig. 2. Comparsons of two NdFeB magnets considering 60□ temperature.

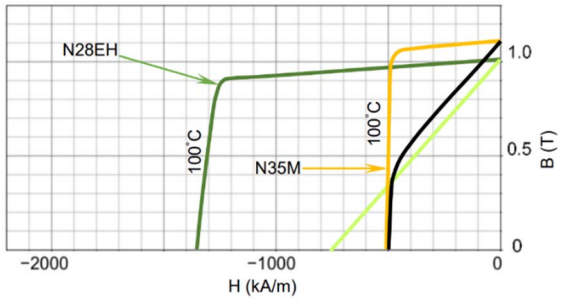


Fig. 3. Comparsons of two NdFeB magnets considering 100□ temperature.

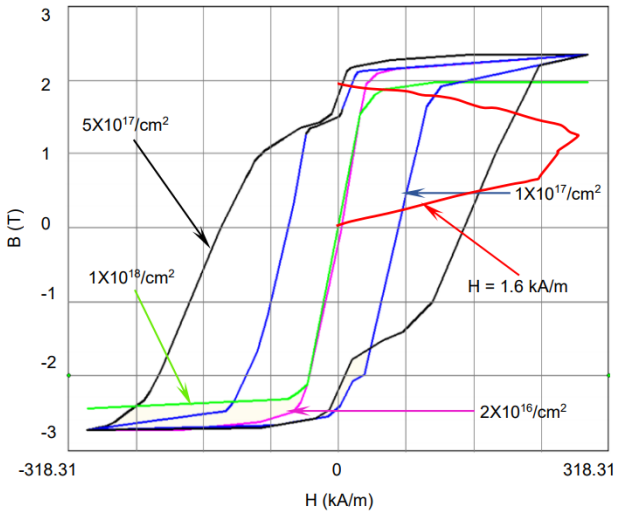


Fig. 4. Characteristics curves of proposed $Fe_{16}N_2$ PM.

The comparisons among different types of PMs are presented in Fig. 5 which includes Ferrite, ordinary and high grade Alnico, NdFeB with the proposed one. It is observed that, magnetic remanence of $Fe_{16}N_2$ is larger than any other existing PMs.

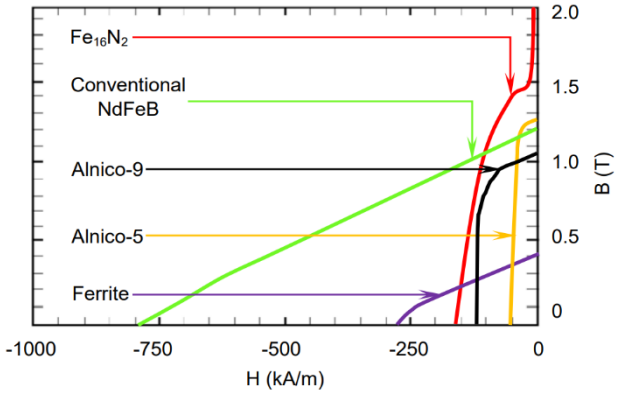


Fig. 5. Comparisons among different types of permanent magnets.

IV. Construction and Model of the Proposed Linear Generator

Fig. 6 illustrates the construction of the proposed linear generator where the translator is consisted of ferromagnetic core and PMs. The cutaway view of two major components namely translator and stator are shown. Windings are placed inside the stator cores.

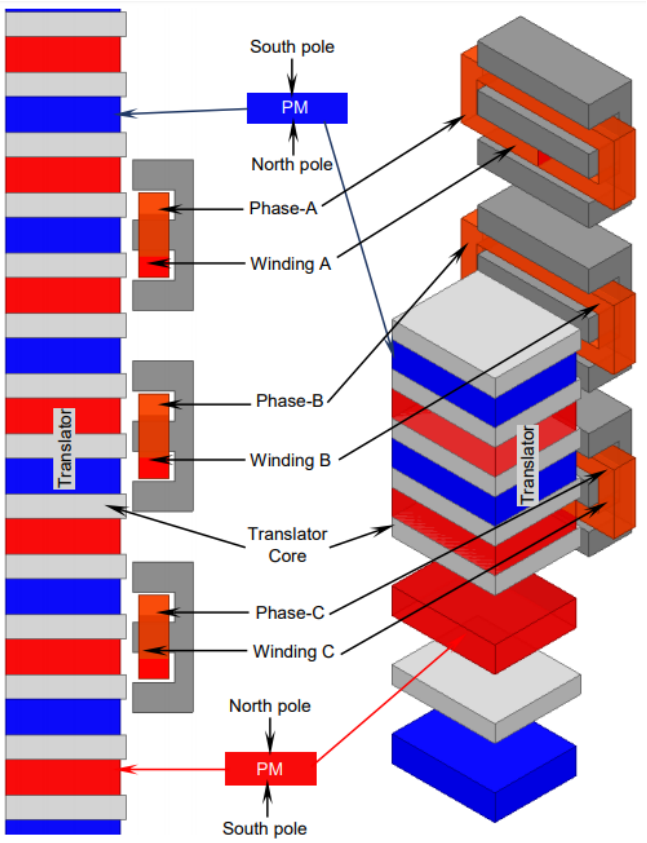


Fig. 6. Structure of proposed linear generator.

The width of proposed linear generator is 0.18m. The length of translator is designed 1.02m. The height of winding and magnetic core are 0.025m and 0.125m, respectively. On the other hand, the width of stator's core and winding are 0.05m and 0.07m, respectively.

The basic two pole machine of the proposed linear generator is presented in Fig. 7. The field winding is shown in direct axis (*d*-axis) and the armature winding A, B, and C are placed by 120°.

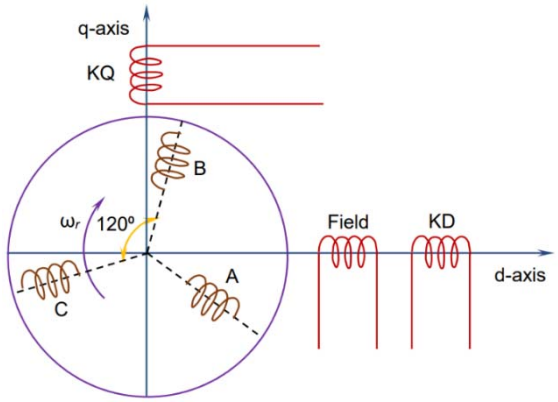


Fig. 7. Basic two pole representation of proposed linear generator.

Here, KD and KQ represent the *d*- and *q*-axis components of amortisseur winding, respectively. The proposed linear generator avoids the amortisseur. Three phase currents denoted as i_a, i_b, and i_c can be expressed in the following.

$$i_a = I_m \cos(\omega t + \alpha) \tag{1}$$

$$i_b = I_m \cos(\omega t + \alpha - \frac{2\pi}{3}) \tag{2}$$

$$i_c = I_m \cos(\omega t + \alpha - \frac{4\pi}{3}) \tag{3}$$

Hence, the d-axis current becomes,

$$i_d = \sqrt{\frac{2}{3}}\left[i_a \cos\theta + i_b \cos\left(\theta - \frac{2\pi}{3}\right) + i_c \cos\left(\theta - \frac{4\pi}{3}\right)\right] \tag{4}$$

where, θ is the equal to ωt and putting the value of i_a, i_b, and i_c in [4],

$$i_d = \sqrt{\frac{2}{3}} I_m \frac{1}{2}(3\cos\alpha)$$

$$= \sqrt{\frac{3}{2}} I_m \cos\alpha \tag{5}$$

Similarly,

$$i_q = \sqrt{\frac{3}{2}} I_m \cos\alpha \tag{6}$$

V. Simulation Results

In the simulation setup, the translator of the proposed linear generator is moved vertically with respect to stationary stator at 1 m/s speed. The air gap length between the stator and translator is selected 1 mm. In the parameter setup, 58 MS conductivity for copper and 7.9 g/m^3 steel core's specific gravity are considered. The generated voltage and armature currents are plotted in Fig. 8.

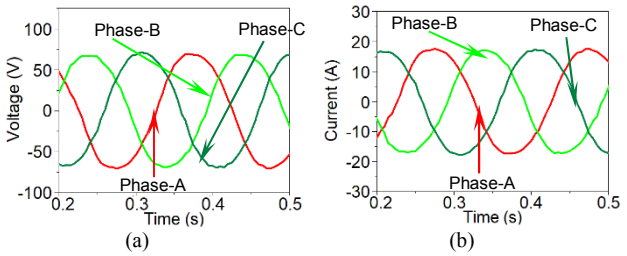


Fig. 8. Simulation results of proposed linear generator: (a) induced voltages and (b) load currents.

Here, the proposed linear generator produces 70V (phase voltage) approximately. The load current is found approximately 17A as shown in Fig. 8(b). The output power is plotted in Fig. 9(a) where approximately 660 W rms power is delivered from each phase. In Fig. 9(b) the flux linkages are shown where the peak value is around 2.5 Wb.

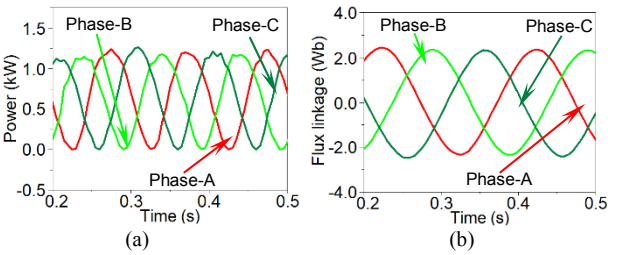


Fig. 9. (a) Output power and (b) magnetic flux linkage of the proposed linear generator.

The comparison of rms power output for using Alnico-9 and $Fe_{16}N_2$ in the linear generator is presented in Fig. 10(a). The generator made of proposed $Fe_{16}N_2$ produces more rms output power compared to that of using the others. Fig. 10(b) depicts the core loss profile of the linear generator using the proposed iron nitride component.

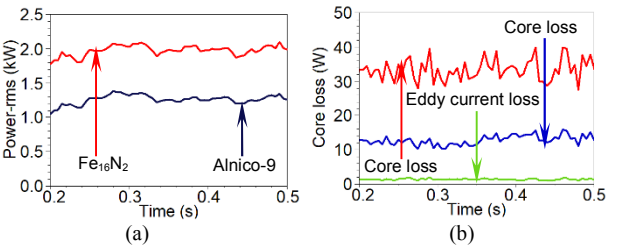


Fig. 10. (a) Output rms power for using the conventional and the proposed material and (b) core loss components of the linear generator using $Fe_{16}N_2$.

The magnetic field intensity and flux density of the proposed linear generator are shown in Figs. 11 and 12, respectively. Table I shows distinguished comparisons of the parameters using Alnico-9 and $Fe_{16}N_2$ in the linear generator. The output parameters of the proposed PM based generator are better than that of using Alnico-9 PM.

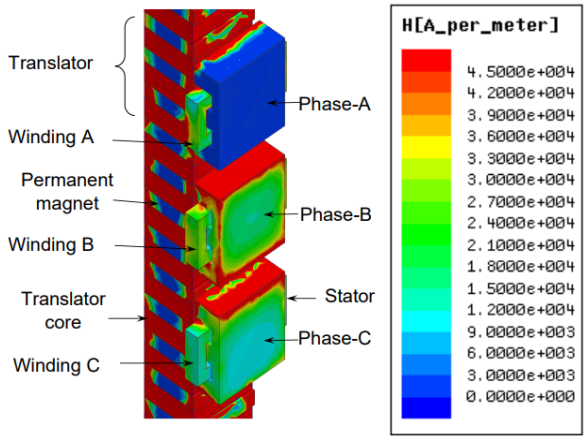


Fig. 11. Field intensity of proposed linear generator.

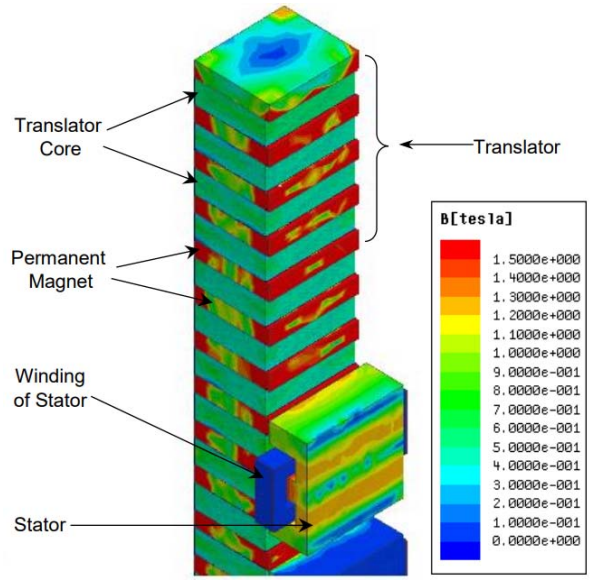


Fig. 12. Flux density of proposed linear generator.

TABLE I. Comparisons of the Parameters using Alnico and the Proposed Iron Nitride Compound

Name of the parameters	Using AlNiCo	Using $Fe_{16}N_2$
Peak generated voltage (V)	57.48	71.06
Peak load current (A)	−14.37	−17.76
Peak magnetic flux linkage (Wb)	1.51	1.78
Peak instantaneous power (W)	826.2	1262.64
Average output rms power (W)	1273.43	1974.83

VI. Conclusions

This paper proposes a linear generator in which nitrogen ion implanted $Fe_{16}N_2$ compound is used as permanent magnet to produce magnetic flux. The remanence and saturation magnetic flux density of the proposed $Fe_{16}N_2$ PMs are quite high. The characteristics curves of the existing PMs and the proposed $Fe_{16}N_2$ are presented and compared. The exciting feature of $Fe_{16}N_2$ compound is that, both Fe and N are REF materials and abundant in nature. Bulk production of PMs is

highly prospective for this compound. To produce the highest output power, proper selection of operating point from the characteristics curve is obvious which is shown in Fig. 4. From the analysis of magnetic properties it is found that there is a good flux linkage in the proposed linear generator.

REFERENCES

[1] S. Molla, O. Farrok, M. R. Islam, and K. M. Muttaqi, "Analysis and design of a high performance linear generator with high grade magnetic cores and high temperature superconducting coils for oceanic wave energy conversion," IEEE Trans. Appl. Supercond., vol. 29, no. 2, Art. ID 5201105, Mar. 2019.

[2] J. Faiz and A. Nematsaberi, "Linear electrical generator topologies for direct-drive marine wave energy conversion-an overview," IET Renewable Power Generation, vol. 11, no. 9, pp. 1163–1176, Sep. 2017.s

[3] B. Drew, A. R. Plummer, and M. N. Sahinkaya, "A review of wave energy converter technology," Proc. Inst. Mech. Eng. A, J. Power Energy, vol. 223, no. 8, pp. 887–902, Jun. 2009.

[4] X. Yang, Y. Song, G. Wang and W. Wang, "A comprehensive review on the development of sustainable energy strategy and implementation in China," IEEE Trans. Sustain. Energy, vol. 1, no. 2, pp. 57–65, Jul. 2010.

[5] J. G. Vining and A.Muetze, "Economic factors and incentives for ocean wave energy conversion," IEEE Trans. Ind. Appl., vol. 45, no. 2, pp. 547–554, Mar.-Apr. 2009.

[6] D. Magagna and A. Uihlein, "Ocean energy development in Europe: current status and future perspectives," Int. J. Marine Energy, vol. 11, pp. 84–104, Sep. 2015.

[7] J. P. Kofoed, P. Frigaard, E. Madsen, and H. Sørensen, "Prototype testing of the wave energy converter wave dragon," Renewable Energy, vol. 31, no. 2, pp. 181–189, Feb. 2006.

[8] R. Henderson, "Design, simulation, and testing of a novel hydraulic power take-off system for the Pelamis wave energy converter," Renew Energy, vol. 31, no. 2, pp. 271–283, Feb. 2006.

[9] S. Molla and O. Farrok, "Vitroperm 500F and supermendur ferromagnetic cores used in a linear generator for oceanic wave energy conversion," Int. Conf. Robotics, Elect. Signal Process. Techn., 10–12 Jan. 2019, Dhaka, Bangladesh, pp. 602–605.

[10] M. S. Bashir and O. Farrok, "Harvesting oceanic wave energy by a linear generator using high graded N28EH permanent magnets," Int. Conf. Elect. Eng. Inform. Commun. Technol., 13–15 Sept. 2018, Dhaka, Bangladesh, pp. 514–518.

[11] M. S. Bashir and O. Farrok, "Yttrium barium copper oxide superconductor used in a linear generator for high power generation from the oceanic wave," Int. Conf. Elect., Comput. Commun. *Eng.*, 7–9 Feb. 2019, Coxbazar, Bangladesh, pp. 1–5.

[12] M. Trapanese, *et al.*, "Minimization of detent force in a 1 kW linear permanent magnet generator for the conversion of sea waves energy: Numerical and experimental validation," IEEE Int. Magn. Conf. (INTERMAG), 11–15 May 2015, Beijing, China, pp. 1–1.

[13] W. N. L. W. Mahadi, S. R. Adi. and K. M. Nor, "Application of the rare earth permanent magnet in linear generator driven by an internal combustion engine," Proceedings. National Power Engineering Conference, 15–16 Dec. 2003, Bangi, Malaysia, pp. 256–261.

[14] W. N. L. W. Mahadi, S. R. Adi. and K. M. Nor, "Thermal Analysis of a Tubular Permanent Magnet Linear Generator using Multiphysics Solver," TENCON 2005 - 2005 IEEE Region 10 Conference, 21–24 Nov. 2005, Melbourne, Qld., Australia, pp. 1–6.

[15] R. Vermaak and M. J. Kamper, "Design aspects of a novel topology air-cored permanent magnet linear generator for direct drive wave energy converters," IEEE Trans. Ind. Electron., vol. 59, no. 5, pp. 2104–2115, May 2012.

[16] T. Xia, *et al.*, "Design and analysis of a field-modulated tubular linear permanent magnet generator for direct-drive wave energy conversion," IEEE Trans. Magn., vol. 53, no. 6, Art. ID 8103904, Jun. 2017.

[17] Y. Jiang, *et al.*, "Synthesis of $Fe_{16}N_2$ compound free-standing foils with 20 MGOe magnetic energy product by nitrogen ion-implantation," Nature, Art. ID 25436, 2016.

Performance Enhancement of Multigate FinFETs by Using High-k Stack Oxide

Emdadul Huq Minhaj
Department of Electrical and Electronic Engineering
Shahjalal University of Science and Technology
Sylhet-3114, Bangladesh
e.h.minhaj@gmail.com

Muhammad Abdur Razzak
Department of Electrical and Electronic Engineering
Shahjalal University of Science and Technology
Sylhet-3114, Bangladesh
abcrazzak20@gmail.com

Md. Majharul Islam
Department of Electrical and Electronic Engineering
Shahjalal University of Science and Technology
Sylhet-3114, Bangladesh
majharul30@gmail.com

Md. Mohsinur Rahman Adnan
Department of Electrical and Electronic Engineering
Shahjalal University of Science and Technology
Sylhet-3114, Bangladesh
adnanmohsin27@gmail.com

***Abstract*— FinFET has become extremely popular in recent times by virtue of its ability to suppress short channel effects (SCEs) in nanometer regime. In this work, we study and compare the electrical characteristics of multigate FinFETs (triple gate, Pi gate, Omega gate and Gate-All-Around (GAA)) and the effect of high-k dielectric oxide stack on the performance of multigate FinFETs. It is found that employing high-k dielectric gate stack reduces short channel effects (SCEs), subthreshold slope, leakage current and alleviates switching characteristics. Also, the superior ability of GAA FinFET to overcome SCEs is also noticed.**

Keywords*— *FinFET, multigate FET, Short channel effects, high-k dielectric, oxide stack

I. INTRODUCTION

In the past few decades, semiconductor industry has undergone a revolutionary change. This revolution is made possible by continuous scaling of transistors. With the reduction in transistor size, the number of transistor in a single chip has been increasing exponentially as predicted by Moore's law [1]. Downscaling of transistors improves device speed, functionality lowers power dissipation and reduces the overall cost of an IC[2][3]. However, in recent years MOSFET scaling has reached a limit owing to the Short Channel Effects (SCEs). In short channel MOSFETs, control of gate over the channel region is weakened, causing an increase in leakage current and SCEs. To continue the scaling, while keeping the SCEs at a optimum value multigate FETs are introduced. In multigate FETs the effective number of gates is increased, thus, improving the electrostatic control of the channel by the gate and suppressing SCEs [4].

FinFET is a variation of multigate FET where a rectangular Fin shaped channel is wrapped around by gate and the Fin is accommodated in a silicon-on-insulator (SOI) interface. FinFETs can have a tri gate structure or a slight modification of the tri gate structure (Pi gate and Omega gate)[5]. When the gate surrounds from all four sides, it is called Gate-All-Around (GAA) structure.

Scaling below 2nm of the gate oxide thickness is not a viable option, because it poses to tunneling current[5]. As the device is scaled down to nanometer regime, the gate oxide thickness (t_{ox}) reduces to a great extent such that SiO_2 can no longer be used as gate dielectric because of electron tunneling through the gate oxide [6]. As a consequence, gate leakage current increases exponentially [7]. To overcome this situation, high-k dielectric with higher relative permittivity (k) than SiO_2 is used instead of SiO_2. But the use of high-k dielectric is limited as a result of high fringing field. The fringing fields originating from the drain reduces the potential barrier and increases leakage current [6][8]. An alternate solution to this problem is using gate stack oxide where high-k dielectric is deposited on top of a thin silicon oxide layer instead of direct deposition.

In this paper, the impact of stack oxide on major electrical parameters of 8nm SOI multigate FinFETs is investigated and compared. The different FinFET structures discussed in this work are triple gate, Pi gate, Omega gate and GAA. The comparison is achieved by altering the high-k dielectrics in the oxide stack between the gate and the Silicon Fin. The dielectrics explored in this research are Silicon Oxide (SiO_2), Silicon Nitride (Si_3N_4), Hafnium Oxide (HfO_2) and Hafnium Silicate ($HfSiO_4$). The parameters considered in this work are Subthreshold Slope (SS), Drain Induced Barrier Lowering (DIBL), on current (I_{on}), off current (I_{off}) and I_{on}/I_{off} ratio. Silvaco Atlas 3D simulator is used for performing the simulation [9]. The simulation takes into account Shockley-Read-Hall model (SRH) and Lombardi mobility model (CVT) for physical modeling of the devices.

The paper is organized into four sections. Section I introduces the work and section II demonstrates the simulated device structure and the simulation methodology. Section III provides the obtained result and the analyses. The paper is concluded in section IV.

II. DEVICE STRUCTURE AND SIMULATION SETUP

In this work, we have studied the effect of stack oxide on four different FinFET structures. Fig. 1 illustrates the cross-sectional view of the FinFET structures. The first structure is a triple gate FinFET and the second structure is Pi gate FinFET. The two structures on the bottom are Omega and GAA FinFET. All the structures are SOI structures having the silicon fin on top a SiO_2 layer. The simulation has been done with 8nm gate length. The Silicon Fin has a height (H_{fin}) of 8nm and a width (W_{fin}) of 4nm. The oxide layer between gate metal and silicon fin has a thickness (t_{ox}) of 1nm. The work function (ϕ) of gate metal is 4.7 eV. In the source and the drain region, the doping concentration (N_d) is $9x10^{19}$ cm^{-3}, whereas the doping concentration (N_a) on the channel region is $9x10^{15}cm^{-3}$.

The simulation is performed using Silvaco TCAD software. At first, the structures are designed using DevEdit tool and then the electrical characteristics are analysed using Atlas device simulator. The models used in this simulation are Shockley-read-hall model (SRH)[10] and Lombardi mobility model (CVT). SRH model is used for considering the interface trap effect and the presence of high impurity atom in the channel region. The mobility model (CVT) takes into account both concentration and parallel field dependence.

The metal gate technology is efficacious when the device is scaled to smaller and smaller dimensions with high-k dielectrics [11]. The oxide layer is varied for single oxide and also for dual-k stack oxide.

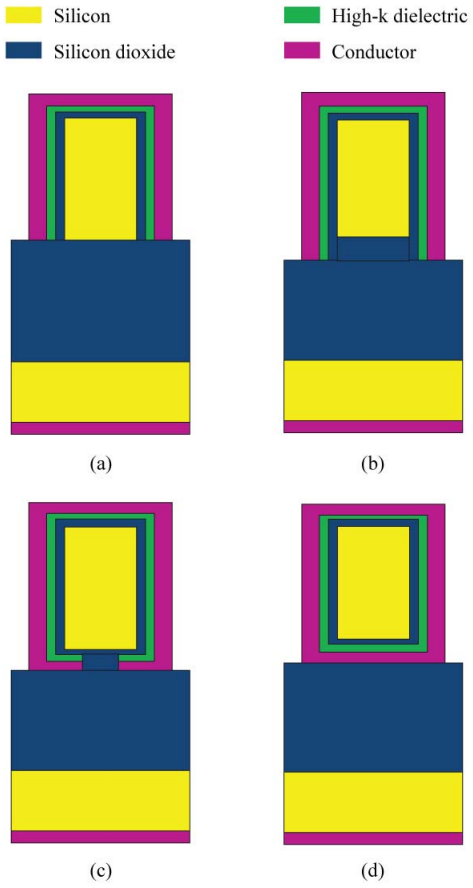


Fig. 1. Cross sectional view of (a) Tri gate, (b) Pi gate, (c) Omega gate, and (d) GAA FinFET structure

TABLE I. DIELECTRIC MATERIALS AND THEIR DIELECTRIC CONSTANTS

Dielectric Material	Dielectric Constant (k)
Silicon Oxide (SiO_2)	3.9
Silicon Nitride (Si_3N_4)	7.5
Hafnium Silicate ($HfSiO_4$)	12.0
Hafnium Oxide (HfO_2)	22.0

The dielectric materials employed are Silicon Oxide, Silicon Nitride, Hafnium Oxide, and Hafnium Silicate. The dielectric constant of these materials is listed on table I.

III. RESULTS AND DISCUSSION

The key electrical parameters Subthreshold Slope (SS), Drain Induced Barrier Lowering (DIBL), On current (I_{on}), Off current (I_{off}) and I_{on}/I_{off} ratio are evaluated using Silvaco Atlas. The high-k dielectrics can be deposited in two methods-first one is direct deposition which can deteriorate the electrical characteristics of device and second one is known as gate-stack (GS) which contains a very thin layer of SiO_2 grown on Si substrate, over which a thick layer of high-k dielectric deposited [15].In this section, the parameters and their interrelation with multigate FinFET devices having varying stack oxide will be discussed.

A. Impact on Subthresold Slope (SS) and DIBL

Subthreshold slope (SS) is a key electrical parameter for evaluating the current below the threshold voltage. SS is defined by the relationship [12]:

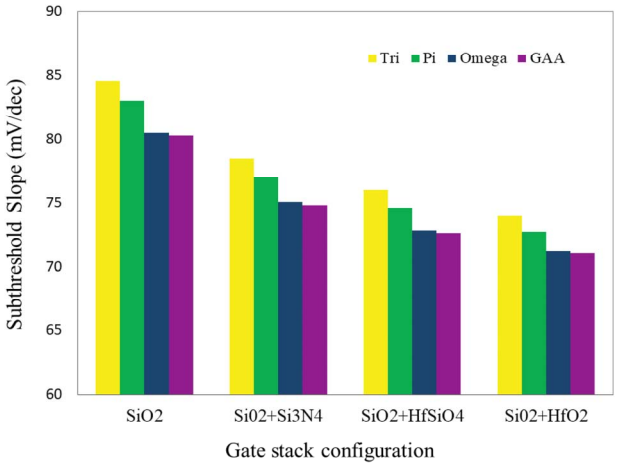


Fig. 2. Impact of gate stack configuration on SS

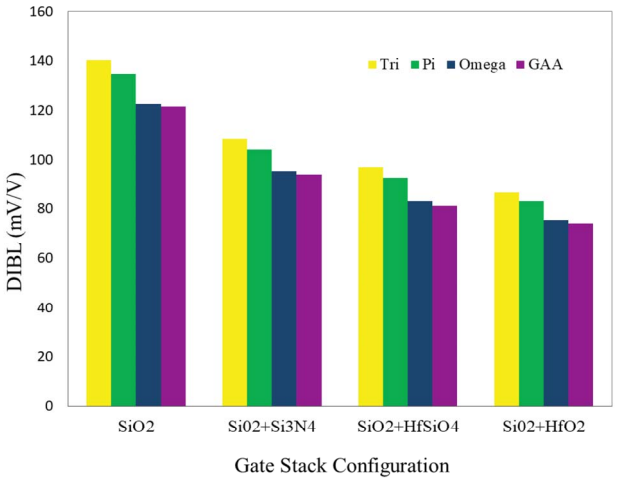


Fig. 3. Effect of gate stack configuration on DIBL

$$SS = dV_{GS} / d(\log_{10} I_{DS})$$

where V_G is the gate voltage and I_D is the drain current. For a FinFET device, an SS value of 80mV/decade means that an 80 mV change in the gate voltage will bring about a tenfold increase in the drain current.

The variation in SS for afore-mentioned FinFET structures is shown in Fig. 2 the value of SS is higher for FinFET devices using only SiO_2 for gate dielectric, with the SS being 84.54 mV/dec, 83 mV/dec, 80.44 mV/dec and 80.26 mV/dec for Tri, Pi, Omega and GAA FinFET respectively. However, for structures using high-k dielectric, the value of SS is significantly reduced, with Hafnium Oxide and Silicon Oxide stack exhibiting better SS than other configurations. The use of HfO_2 as gate dielectric yields the SS value of 74 mV/dec for Tri gate, 72.76 mv/dec for Pi gate, 71.2 mV/dec for Omega and 70.99 mV/dec for GAA FinFET. Furthermore, due to enhanced electrostatic control, in GAA FinFET, leakage current is abated and hence, SS is lower than other three multigate structures discussed.

Drain Induced Barrier Lowering (DIBL) is one of the short channel effects that arises as a result of channel length reduction. It is defined as the decline in potential barrier with the increase of drain bias. DIBL is expressed by the equation as in [13]:

$$DIBL\ (mV/V) = \Delta V_{TH} / \Delta V_{DS}$$

Fig.3 shows the DIBL of multigate FinFET devices with varying gate stacks. The DIBL follows a downward trend as we shift from SiO_2 to high-k dielectric gate stacks, with the DIBL ranging from 140.25 mV/V for SiO_2 to 86.84 mV/V for Hafnium Oxide stack considering Tri gate FinFET. GAA FinFET along with HfO_2 gate stack provides an even better result, suppressing DIBL to 74.24 mV/V.

From both fig. 2 and fig. 3 it is evident that higher relative permittivity of the high-k layer in the gate oxide stack results in enhanced DIBL and Subthreshold Slope.

B. Impact on current, off current and I_{on}/I_{off} ratio

The on current (I_{on}) and off current (I_{off}) are major device electrical parameters to estimate the device performance, especially switching speed of the circuit and standby time. In order to achieve robust switching performance, I_{on} is required to be higher and I_{off} should be as close to zero as possible, meaning a high I_{on}/I_{off} ratio.

But in short channel transistors, the prevalence of high leakage current hinders switching performance by reducing the I_{on}/I_{off} ratio. Use of multigate structure together with high-k dielectric proves to mitigate the problem and demonstrates superior switching behavior [14].

The on-state current is defined as the drain current at which gate voltage (V_{GS}) equals the drain voltage (V_{DS}). In this work I_{on} is measured at $V_{DS} = 0.1V$. Fg.4 shows the fluctuation in I_{on} when the gate oxide stack is altered. FinFETs with SiO_2 as gate dielectric does not indicate significant on current. But, substituting SiO_2 with high-k dielectric stack exhibits a gradual increase in on current. The best value of I_{on} is observed with GAA FinFET having HfO_2-SiO_2 stack. The mobility degradation is a serious issue when the high-k dielectrics are used instead of conventional gate oxide SiO_2 and which is due to Remote Coulomb Scattering (RCS), Remote Phonon Scattering (RPS) [15]. The net mobility of GS high-k dielectric depends on thickness, a dielectric constant of Interfacial-Layer

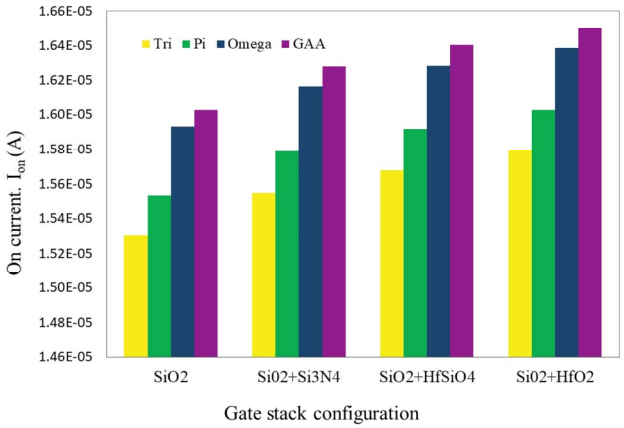


Fig. 4. Effect of gate stack configuration on I_{on}

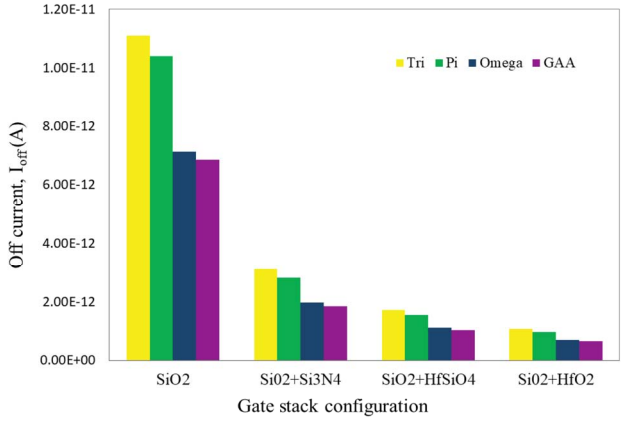


Fig. 5. Effect of gate stack configuration on I_{off}

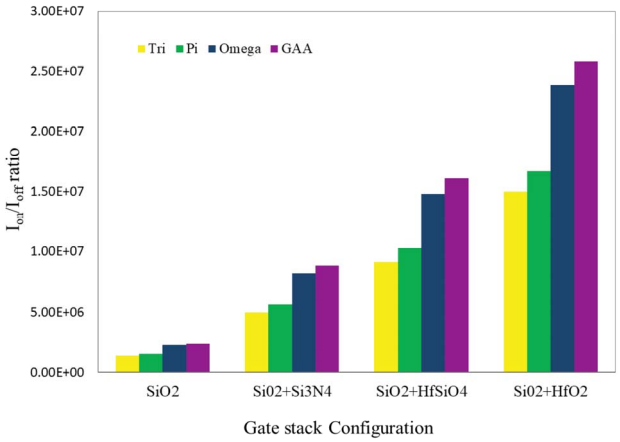


Fig. 6. Effect of gate stack configuration on I_{on}/I_{off} ratio

(IL) and of high-k. The mobility reduced due to coupling between the charges and channel (increased RCS and RPS) when the dielectric constant of IL increase [16]. In high-k FinFETs, increased relative permittivity increases the fringing field, which, in turn, reduces scattering. As a result, mobility is enhanced, making it possible to observe improved I_{on} even at 8nm channel length [17].

On the other hand, off current is the current that flows in the absence of gate voltage (V_{GS} = 0V). FinFETs with SiO_2 as dielectric, as shown in fig. 5, demonstrates I_{off} as high as in the range of 10^{-11}. The alteration in the dielectric shows a drastic reduction in I_{off}, with I_{off} in the range of 10^{-13} for HfO_2 stack.

In Fig. 6, the change in I_{on}/I_{off} ratio in accordance with the gate oxide stack variation is shown. The figure represents an overall upward trend as oxide stack with higher relative permittivity (k) is used. The I_{on}/I_{off} ratio is in the order of 10^6 when SiO_2 is used as dielectric in multigate FinFET devices. For Si_3N_4 and $HfSiO_4$ gate stack, I_{on}/I_{off} ratio increases exponentially. The I_{on}/I_{off} current ratio reaches the order of 10^7 when HfO2 oxide stack is used. GAA FinFETs, with ultimate control over the channel region, provides improved I_{on}/I_{off} ratio than other multigate FinFETs.

IV. Conclusion

A 3D simulation analysis of multigate FinFETs with high-k stack oxide is performed. From the simulation it is observed that, high-k stack oxide shows reduced short channel effects, lower leakage current, improved SS and higher I_{on}/I_{off} ratio than SiO_2. Also, gate stacks having dielectric with higher relative permittivity (k) exhibit better short channel effect suppression. It is also found that, GAA FinFET indicates enhanced performance than Tri gate, pi gate and Omega gate FinFETs.

References

[1] G. E. Moore, "Cramming more components onto integrated circuits", Electronics, vol. 8, 1965.

[2] Wann, C. H., et al, "A Comparative Study of Advanced MOSFET Concepts." IEEE *Trans. Electron Dev.*, Vol. 43, no. 10, pp. 1742-1753, 1996.

[3] T. Skotnicki, J.A. Hutchby, T.J. King, H.S.P. Wong and F. Boeuf, "The end of CMOS scaling", *IEEE Circuits & Devices magazine*, Vo. 21, no. 1, pp. 16-26, 2005.

[4] J. P. Colinge, "Multi-gate SOI MOSFETs", Microelectron. Eng., vol. 84, no. 9/10, pp. 2071-2076, 2007.

[5] R. Ritzenthaler, C. Dupré, X. Mescot, O. Faynot, T. Ernst, J.C. Barbé, C. Jahan, L. Brévard, F. Andrieu, S. Deleonibus, S Cristoloveanu, "Mobility behavior in narrow Ω-gate FET devices," Proceedings IEEE International SOI Conference, (2006) 77.

[6] N. Boukortt, B. H. S. Patan, and A. C. G. Crupi, "Electrical Characteristics of 8-nm SOI n-FinFETs," Springer Sci., vol. 8, pp. 497–503, 2016.

[7] X. Liu, S. Lou, Z. Xia, et al. "Characteristics of different structure sub-100nm MOSFETs with high-k gate dielectrics," Solid-State and Integrated-Circuit Technology, pp. 333-336, Jan 2001.

[8] H. S. Momose M. Ono and T. Yositomi "1.5 nm direct tunneling gate oxide si MOSFETs" IEEE Trans. on Electron Devices vol. 43 no. 8 pp. 1233-1242 1996.

[9] SILVACO International 2000, ATLAS: 2-D Device Simulation Software.

[10] R.N. Hall Phys. Rev., 87 (1952), p. 387.

[11] W. Shockley, W. T. Read, "Statistics of the Recombination of Holes & Electrons." Phy. Rev.vol. 87, p.p. 835-842,1952.

[12] H.-S. Wong, "Beyond the conventional transistor" in IBM J. of Research and Development, vol. 46, no. 2/3, 2002.

[13] Ferain, I., Colinge, C. A. & Colinge, J.-P., "Multigate transistors as the future of classical metal–oxide–semiconductor field-effect transistors," Nature 479, 310–316 (2011)

[14] V. Narendar, R. A. Mishra, "Analytical modeling and simulation of multigate FinFET devices and the impact of high-k dielectrics on short channel effects (SCEs)", Superlattices and Microstruct, vol. 85, pp. 357-369, 2015.

[15] J. Koh, Y. Lu, C. R. Wronski, Y. Kuang, R. W. Collins, T. T. Tsong and Y. E. Strausser, "Correlation of real time spectroellipsometry and atomic force microscopy measurements of surface roughness on amorphous semiconductor thin films" *Appl. Phys. Lett.* 69, 1297(1996)

[16] G.S. Lujana, W. Magnus, L-Å Ragnarsson, S. Kubicek, S. De Gendt, c,, M. Heyns, K. De Meyer, "Modelling mobility degradation due to remote Coulomb scattering from dielectric charges and its impact on MOS device performance",*MicroelectronicsReliability*Vol.45,pp. 794–797, 2005

[17] G. Ribes, J. Mitard, M. Denais, S. Bruyere, F. Monsieur, C. Parthasarathy, E. Vincent, and G. Ghibaudo, "Review on high-k dielectrics reliability issues," IEEE Trans. Dev. Mater. Reliab. Vol. 5, no. 1, pp. 5–19, 2005.

Effect of Solar Concentration on Performance of Perovskite Solar Cell: A detailed balance study

Md. Masud Rana
Department of EEE, Jashore University of Science and Technology
Jashore-7408, Bangladesh
mmrana.3012@gmail.com

Shyamal Kumar Mandal
Department of EEE, Jashore University of Science and Technology
Jashore-7408, Bangladesh
skumardotorg@gmail.com

Md. Tariquzzaman
Department of EEE, Jashore University of Science and Technology
Jashore-7408, Bangladesh
topu31.eee.just@gmail.com

***Abstract*—As a perovskite material, methylammonium lead iodide has become more popular for its flexibility, low-cost, lightweight, thin film and high power conversion capability in photovoltaic research community. In this work, we have explored the effect of solar concentration on PV performance for methylammonium lead iodide ($CH_3NH_3PbI_3$) perovskite solar cell computationally. To calculate the band structure, bandgap and density of states (DOS) of $CH_3NH_3PbI_3$, density functional theory (DFT) is used here. The detailed balance method is employed to figure out the effect of concentration on performance of perovskite solar cell. We have calculated performance parameters of $CH_3NH_3PbI_3$ solar cell for two solar spectra AM 1.5 G and AM 1.5 D. All calculations have been performed at different solar concentrations ranging from no concentration to full concentration for both types solar spectra. It is found that the solar cell efficiency at 1.5 G spectrum is higher than 1.5 D solar spectrum at any solar concentration. These results may impact future research on perovskite solar cell.**

Keywords—Perovskite solar cell (PSC), solar concentration, detailed balance limit, density functional theory (DFT), PCE.

I. INTRODUCTION

The worldwide average energy utilization rate is augmenting. It has been found that a daily solar radiation approximately 2×10^{17} W incident on earth at sea level is 12500 times over than the average global power consumption in 2006 [1]. It indicates the immense potential of solar energy to deliver the energy demand of the world. With the rapid usage of conventional energy, the world climate is being hampered tremendously. Hence we are to search for environment friendly sources of energy. Solar energy is abundant and it creates no harmful effect to the environment. Solar cell is the most efficient device to utilize the energy of the sun. Scientists are working on solar cells and trying to increase their efficiencies by using different types of materials. For perovskite materials, the efficiencies of solar cell have increased from 3.8% in 2009 [2] to 23.3% in late 2018 [3, 4]. Methylammonium lead tri-iodide ($CH_3NH_3PbI_3$) perovskite solar cells have unique features such as direct bandgap, lightweight, thin and flexibility, low cost of raw materials and low fabrication cost, high absorption coefficient, high efficiency [5-9]. These are very amazing in the field of solar cell research.

At present researchers are thinking of third generation photovoltaic cell due to efficiency limit. Perovskite solar cell is in the third generation category. The highest experimental efficiency of $CH_3NH_3PbI_3$ solar cell is 19.3% [10]. For a single-bandgap material, the detailed balance efficiency is 31% with no concentration (i.e. 1 sun) [11] for bandgap of 1.31 eV and 41% with full concentration (i.e.46,198 suns) [12] for bandgap of 1.1 eV. The AM 1.5 G and AM 1.5 D spectra have been considered in this work. The AM 1.5 D spectrum accounts for solar radiation as it is received directly from the sun after experiencing loss through the atmosphere. The AM 1.5 G spectrum adds extra solar radiance to the AM 1.5 D spectrum to account for additional light received from a 2π steradian field of view due to Rayleigh scattering.

Density functional theory (DFT) has been used in this research to find out the electronic properties of methylammonium lead iodide ($CH_3NH_3PbI_3$). We have found bangap of 1.65 eV for this organometal lead trihalide. The detailed balance efficiency according to SQ model [11] has been calculated to get maximum possible efficiency. The PCE (Power Conversion Efficiency) is 32.01% for experimental value, 29.87% for calculated value at AM 1.5 G and 31.17% for experimental value, 28.81% for calculated value at AM 1.5 D under no concentration. Under full concentration, the power conversion efficiency (PCE) is 39.86% for experimental value, 36.40% for calculated value at AM 1.5 G and 38.84% for experimental value, 35.13% for calculated value at AM 1.5 D.

II. SIMULATION MODEL

A. Detailed Balance Model

To calculate efficiency of solar cells many types of models are used by PV researchers. In this study, detailed balance limit according to Shockley Queisser model [11] has been adopted to estimate the best possible efficiency for the organometal lead iodide- $CH_3NH_3PbI_3$. The SQ model assumptions including ideal condition are as follow:

- 6000 K and 300 K are the sun and solar cell temperature respectively.
- No overlapping in inter band transition.
- Between two bands there must be radiative recombination.
- Carriers with infinite mobilities are extracted from VB or CB.

The current-voltage characteristic for single gap solar cell is given by

$$I(V) = q[Xf_s\varphi(E_g,\infty,T_s,0)+(f_c - Xf_s)\varphi(E_g,\infty,T_c,0) - f_c\varphi(E_g,\infty,T_c,qV)] \quad (1)$$

where E is energy, X is solar concentration, q is electron charge, T_s and T_c are the sun and solar cell temperature

respectively and, f_s and f_c are the sun and solar cell geometrical factors respectively. For ideal behaviour of cell, the photon flux expression [13] is represented as

$$\varphi(E_l, E_h, \mu, T) = \frac{2\pi}{h^3 c^2} \int_{E_l}^{E_h} \frac{E^2}{\exp(^{E-\mu}/_{kT}) - 1} dE \qquad (2)$$

where, k is Boltzmann's constant, μ is chemical potential, c is the speed of light and h is Planck's constant. So the PCE (power conversion efficiency) is found by

$$PCE = \frac{V.I(V)}{X.P_{in}} \qquad (3)$$

where P_{in} is given by

$$P_{in} = q \int_0^{\infty} E \left(\frac{2\pi q^3}{h^3 c^2} \right) \frac{E^2}{\exp(^{E}/_{kT}) - 1} dE \qquad (4)$$

We have calculated the PCE for $CH_3NH_3PbI_3$ solar cell using this equation.

B. DFT Calculation

DFT is a computational technique of obtaining an approximate solution to the Schrodinger equation of many body systems. Investigation of magnetic, structural and electronic properties of molecules, materials and defects are computed in DFT. Electron density is used to calculate one part of energy and other part of energy computation is done by a wave function in DFT method. We can find the total electron density if the squared wave function is divided by the number of electrons. Hohenburg-Kohn stated a theorem that all ground state properties of a system are determined by the density of the system [14, 15]. The total ground state energy is a functional of density of a many electron system in this case.

III. Results and Discussion

A. Band Gap and Density of states

Quantum Espresso software simulation is used for calculation of band gap and density of states (DOS) of $CH_3NH_3PbI_3$ based on DFT. We have used the so-called Generalized Gradient Approximation (GGA) to describe the exchange correlation functional. The schematic diagram of unit cell, band structure and DOS (density of states) are shown in Fig. 1, Fig. 2 and Fig. 3 respectively. The lattice constants for symmetric group of Pm are a = 6.437 Å, b = 6.439 Å and c = 6.492 Å for the monoclinic phase material $CH_3NH_3PbI_3$ at room temperature. The calculated bandgap of methylamonium lead iodide ($CH_3NH_3PbI_3$) is 1.65eV which has difference in value with the given experimental bandgap in other research [16, 17].

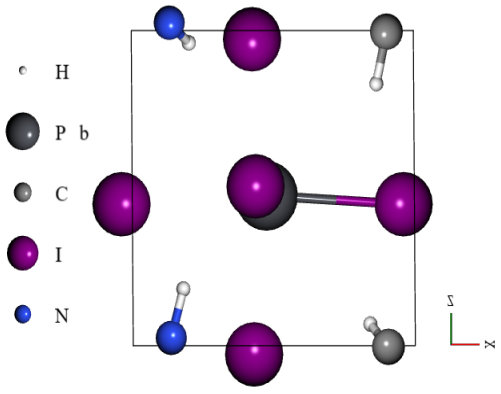


Fig. 1. Schematic diagram of methylammonium lead iodide ($CH_3NH_3PbI_3$) unit cell

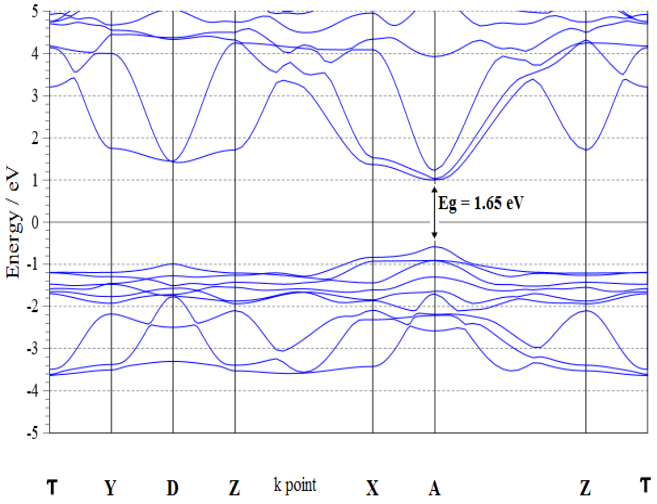


Fig. 2. Bandgap and band structure of methylammonium lead iodide ($CH_3NH_3PbI_3$).

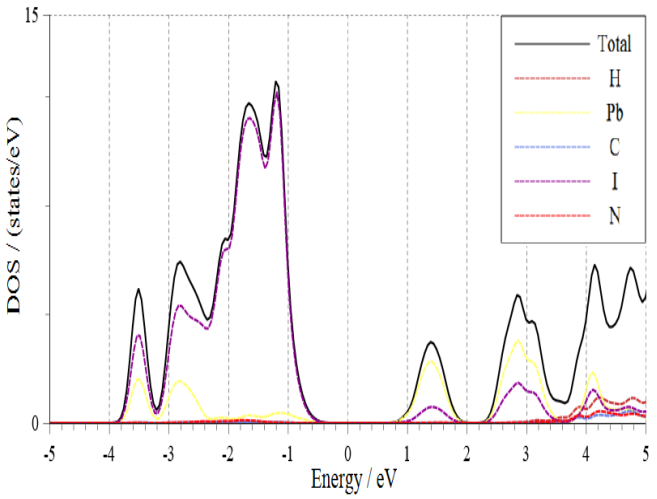


Fig. 3. Density of states (DOS) at different enrgy (eV) for methylammonium lead iodide ($CH_3NH_3PbI_3$).

B. Detailed Balance Study

The detailed balance analysis of the solar cell is performed here using the AM 1.5 D and AM 1.5 G solar spectra. The operational response of a solar cell depends on the spectrum of light that illuminates the cell. The sun is often modeled as a 6000 K blackbody. But the sun's temperature varies across the solar disc and light originating from the sun will be attenuated by the solar atmosphere that may make blackbody approximation questioned. Considering earth's atmospheric condition solar cell performances are taken at AM 1.5 solar spectrum. These effects are increased for longer optical path lengths traversed through the

atmosphere. So, the AM 1.5-Direct (AM 1.5 D) and AM 1.5-Global (AM 1.5 G) solar spectra are used for terrestrial applications [4]. These spectra are used by the photovoltaic community for terrestrial based applications.

In our calculation, we have considered the solar spectra AM 1.5 D and AM 1.5 G. The geometric parameters f_c and f_s have been defined to be equal to 1 and 2.1646×10^{-5} [18] respectively. The performance of a photovoltaic cell is shown at different solar concentration (i.e. 1 sun = No concentration to 46,198 sun = full concentration). The detailed balance performance of $CH_3NH_3PbI_3$ solar cell for experimental bandgap (E_g = 1.5 eV) and calculated bandgap (E_g = 1.65 eV) are plotted in Fig. 4 and Fig. 5 respectively as

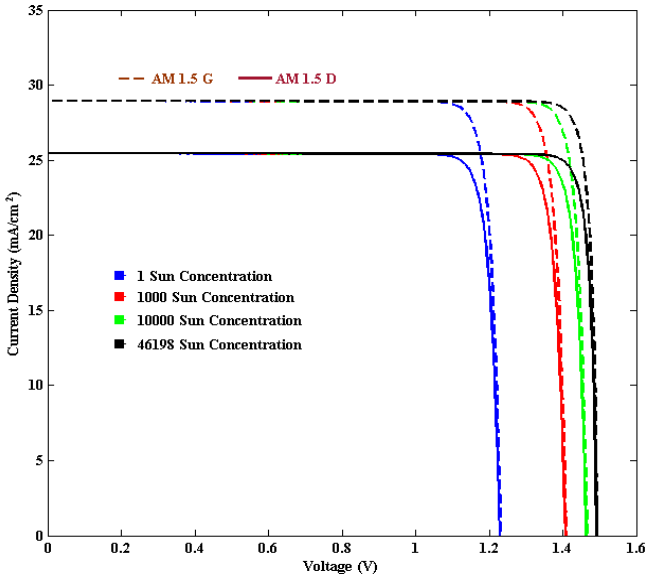


Fig. 4. I-V characteristics of single-gap perovskite (Methylammonium lead iodide, $CH_3NH_3PbI_3$) solar cell performed in detailed balance limit for experimental bandgap at different solar concentration (dotted lines and solid lines show results of Air Mass 1.5 G and Air Mass 1.5 D respectively).

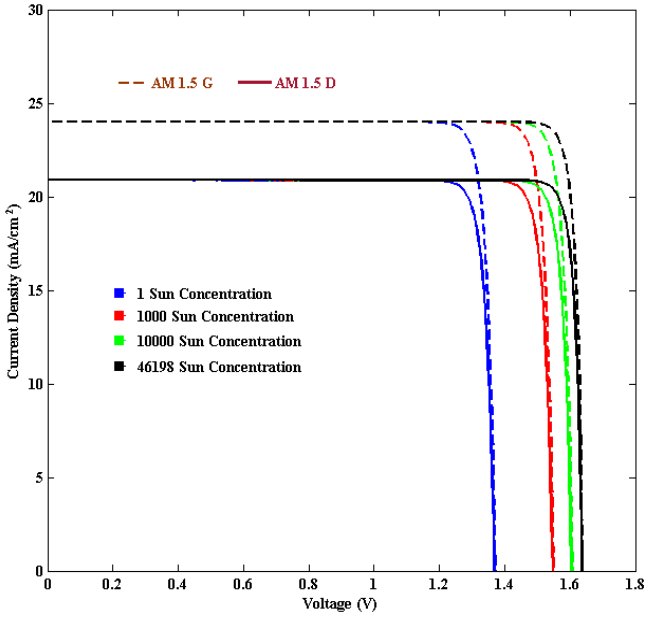


Fig. 5. I-V characteristics of single-gap perovskite (Methylammonium lead iodide, $CH_3NH_3PbI_3$) solar cell performed in detailed balance limit for calculated bandgap at different solar concentration (dotted lines and solid lines show results of Air Mass 1.5 G and Air Mass 1.5 D respectively).

TABLE I. DETAILED BALANCE EFFICIENCY OF METHYLAMMONIUM LEAD IODIDE SOLAR CELL FOR EXPERIMENTAL BANDGAP AT DIFFERENT SOLAR CONCENTRATION

Spectrum	Efficiency			
	1 sun	*800 sun*	*9000 sun*	*46198 sun*
AM 1.5 G	32.01	36.90	38.67	39.86
AM 1.5 D	31.17	35.95	37.68	38.84

TABLE II. DETAILED BALANCE EFFICIENCY OF METHYLAMMONIUM LEAD IODIDE SOLAR CELL FOR CALCULATED BANDGAP AT DIFFERENT SOLAR CONCENTRATION

Spectrum	Efficiency			
	1 sun	*800 sun*	*9000 sun*	*46198 sun*
AM 1.5 G	29.87	33.93	35.41	36.40
AM 1.5 D	28.81	32.74	34.17	35.13

a set of I-V characteristics. The experimental bandgap allows larger short-circuit current (I_{sc}) than calculated bandgap because the experimental bandgap value absorbs a larger portion of solar spectrum than calculated bandgap value. But experimental bandgap places a fundamental limit on the open-circuit voltage i.e. it has lower open circuit voltage (V_{oc}) where the calculated value has higher open circuit voltage (V_{oc}). We have used MATLAB coding for the simulation process to obtain the efficiencies, fill factors and current-voltage characteristics. The power conversion efficiency of this methylammonium lead iodide solar cell at different solar concentration for both AM 1.5 G and AM 1.5 D are tabulated in TABLE I and TABLE II respectively. The efficiencies for experimental and calculated bandgap for different solar concentrations are shown in Fig. 6 and Fig. 7 respectively. All values are plotted based on detailed balance efficiency limit from 1 sun concentration (no concentration) to 46,198 sun concentration (full concentration). The solar cell efficiency

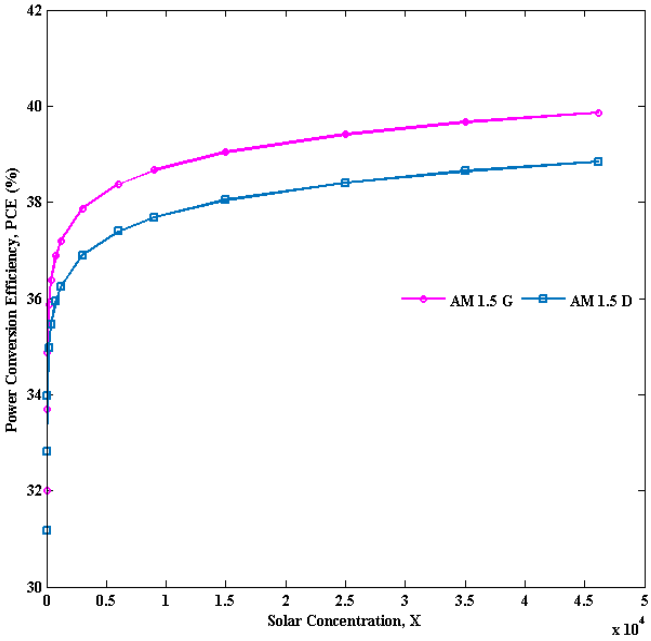


Fig. 6. Power conversion efficiency (PCE) of a single-gap perovskite (Methylammonium lead iodide, $CH_3NH_3PbI_3$) solar cell with different solar concentration at experimental bandgap for both AM 1.5 G and AM 1.5 D.

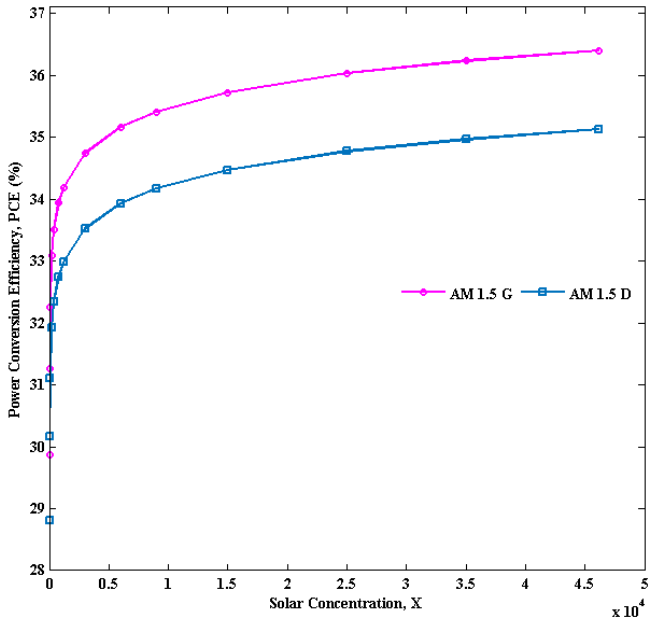


Fig. 7. Power conversion efficiency (PCE) of a single-gap perovskite (Methylammonium lead iodide, $CH_3NH_3PbI_3$) solar cell with different solar concentration at calculated bandgap for both AM 1.5 G and AM 1.5 D.

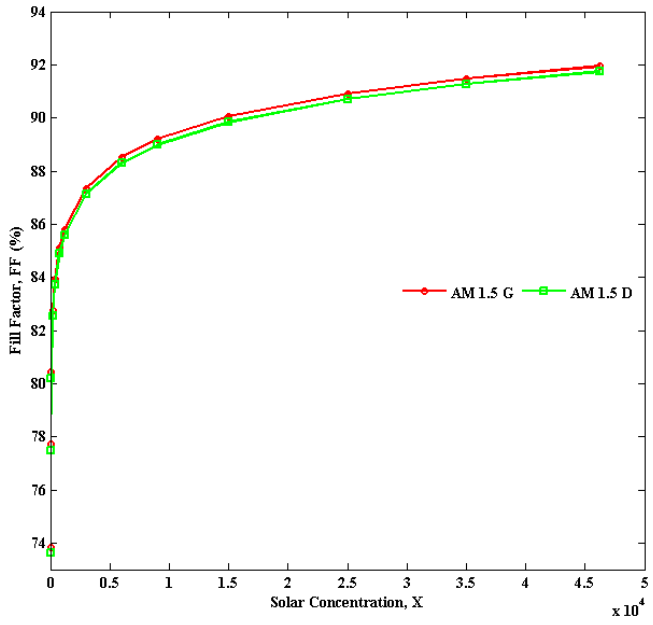


Fig. 8. Fill factor of a single-gap perovskite (Methylammonium lead iodide, $CH_3NH_3PbI_3$) solar cell with different solar concentration at experimental bandgap for both AM 1.5 D and AM 1.5 G.

increases with solar concentration where highest concentration shows highest value of solar cell efficiency. For the experimental bandgap (E_g = 1.5 eV), the maximum obtained efficiency is 39.86% for AM 1.5 G and 38.84% for AM 1.5 D. For the calculated bandgap (E_g = 1.65 eV), we have found the maximum efficiency 36.40% for AM 1.5 G and 35.13% for AM 1.5 D. The fill factors at different solar concentrations are shown in Fig. 8 and Fig. 9 for experimental and calculated bandgap respectively. The AM 1.5 G and AM 1.5 D give approximately same values of fill factor at different concentrations. We have evaluated the fill factors as a function of solar concentration for both experimental and calculated bandgap. Under no concentration, the fill factors are 73.83%, 75.45% for experimental and calculated bandgap respectively at AM 1.5 G. Under no concentration, fill factors are 73.61 %, 75.23 % for experimental and calculated bandgap respectively at AM 1.5 D. Under full concentration fill factors are 91.94% and 91.96% for experimental and calculated bandgap respectively for AM 1.5 G. Under full concentration fill factors are 91.73% and 91.74% for experimental and calculated bandgap respectively for AM 1.5 D. The improvement in cell performance is given by increment of fill factors. Methylammonium lead iodide ($CH_3NH_3PbI_3$) solar cell gives high fill factors for both solar spectra. The fill factors are increased with increasing of solar concentration.

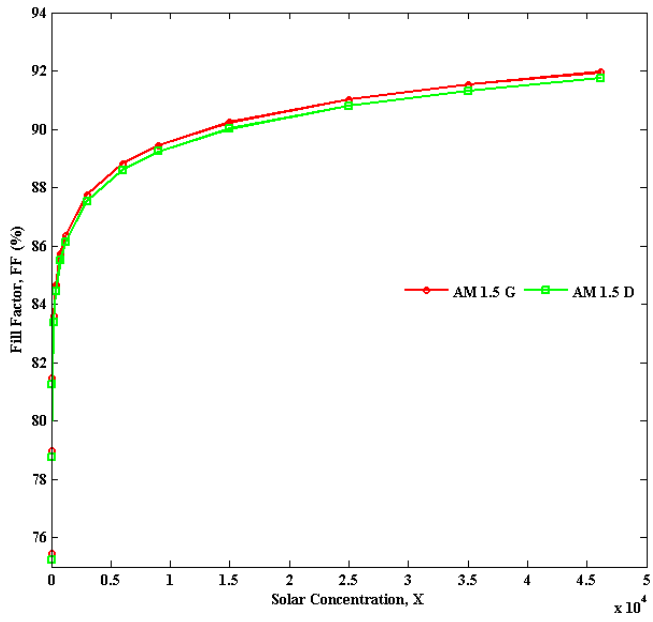


Fig. 9. Fill factor of a single-gap perovskite (Methylammonium lead iodide, $CH_3NH_3PbI_3$) solar cell with different solar concentration at calculated bandgap for both AM 1.5 G and AM 1.5 D.

IV. Conclusion

In this work, we have probed the effects of solar concentration on different cell performance parameters. The DFT method is used to calculate the bandgap, band structure and density of states of methylammonium lead iodide perovskite. Detailed balance method is employed to calculate efficiency of $CH_3NH_3PbI_3$ solar cell. We have performed our investigations for both AM 1.5 G and AM 1.5 D conditions. We have obtained the optimum possible efficiency for the bandgap. It is found that the AM 1.5 G solar spectrum produces better efficiency compared to AM 1.5 D. Our results will be fruitful for further research in $CH_3NH_3PbI_3$ photovoltaics.

References

[1] Twidell, J. and Weir, T. (n.d.). *Renewable energy resources.*

[2] A. Kojima, K. Teshima, Y. Shirai and T. Miyasaka, "Organometal Halide Perovskites as Visible-Light Sensitizers for Photovoltaic Cells", *Journal of the American Chemical Society,* vol. 131, no. 17, pp. 6050-6051, 2009.

[3] Djurišić, A., Liu, F., Tam, H., Wong, M., Ng, A., Surya, C., Chen, W. and He, Z. (2017). Perovskite solar cells - An overview of critical issues. *Progress in Quantum Electronics,* 53, pp.1-37.

[4] "National Renewable Energy Laboratory (NREL) Home Page | NREL", Nrel.gov, 2018.

[5] G. Hodes, "Perovskite-Based Solar Cells", *Science*, vol. 342, no. 6156, pp. 317-318, 2013.

[6] M. Green, A. Ho-Baillie and H. Snaith, "The emergence of perovskite solar cells", *Nature Photonics*, vol. 8, no. 7, pp. 506-514, 2014.

[7] P. Gao, M. Grätzel and M. Nazeeruddin, "Organohalide lead perovskites for photovoltaic applications", *Energy Environ. Sci.*, vol. 7, no. 8, pp. 2448-2463, 2014.

[8] H. Snaith et al., "Anomalous Hysteresis in Perovskite Solar Cells", *The Journal of Physical Chemistry Letters*, vol. 5, no. 9, pp. 1511-1515, 2014.

[9] N. Park, "Perovskite solar cells: an emerging photovoltaic technology", *Materials Today*, vol. 18, no. 2, pp. 65-72, 2015.

[10] Zhou, H., Chen, Q., Li, G., Luo, S., Song, T., Duan, H., Hong, Z., You, J., Liu, Y. and Yang, Y. (2014). Interface engineering of highly efficient perovskite solar cells. *Science,* 345(6196), pp.542-546.

[11] W. Shockley and H. Queisser, "Detailed Balance Limit of Efficiency of p-n Junction Solar Cells", Journal of Applied Physics, vol. 32, no. 3, pp. 510-519, 1961.

[12] G. Araújo and A. Martí, "Absolute limiting efficiencies for photovoltaic energy conversion", Solar Energy Materials and Solar Cells, vol. 33, no. 2, pp. 213-240, 1994.

[13] A. Luque and A. Martí, "Increasing the Efficiency of Ideal Solar Cells by Photon Induced Transitions at Intermediate Levels", *Physical Review Letters,* vol. 78, no. 26, pp. 5014-5017, 1997.

[14] A. Görling, "Density-functional theory beyond the Hohenberg-Kohn theorem", *Physical Review A,* vol. 59, no. 5, pp. 3359-3374, 1999.

[15] Sahni, V. (2004). The Hohenberg-Kohn Theorems and Kohn-Sham Density Functional Theory. Quantal Density Functional Theory, pp.99-123.

[16] Shi, D., Adinolfi, V., Comin, R., Yuan, M., Alarousu, E., Buin, A., Chen, Y., Hoogland, S., Rothenberger, A., Katsiev, K., Losovyj, Y., Zhang, X., Dowben, P., Mohammed, O., Sargent, E. and Bakr, O. (2015). Low trap-state density and long carrier diffusion in organolead trihalide perovskite single crystals. *Science,* 347(6221), pp.519-522.

[17] M. Saidaminov et al., "High-quality bulk hybrid perovskite single crystals within minutes by inverse temperature crystallization", *Nature Communications,* vol. 6, no. 1, 2015.

[18] M. Green, "Multilayer thin film silicon solar cells", *Natural Resources Forum*, vol. 19, no. 4, pp. 269-273, 1995.

Details semiconductor losses calculation of single phase dual active bridge converter for fast charger station application

1st Mohammad Jalal Uddin
Department of electrical and electronic engineering
International Islamic University Chittagong (IIUC)
Chittagong, Bangladesh
jalaliiuc@gmail.com

***Abstract*—Single phase dual active bridge (DAB) dc-dc converter is an ambitious candidate for the fast charger technology of the electrified vehicles. It has two dc-ac converter connected back to back through a bridge inductor isolated by a transformer. In particular, this paper presents the semiconductor loss calculation and the efficiency of the converter considering two different models and zero voltage switching (ZVS) switching. It also shows the optimal frequency range of $20\ kHz$ to $70\ kHz$ for the highest efficiency above 97% at different power level of operation between $124 - 35\ kW$. Finally, the study recommends a suitable cooling arrangement for the fast charger station to keep the junction temperature within the permissible limit. The simulation results confirm theoretical values.**

***Index Terms*—single phase DAB, dc-dc converter, semiconductor loss, fast charger station, efficiency of DAB, switching frequency of DAB**

I. Introduction

The modern world is moving towards electric vehicle solution from conventional one for sustainable transportation, reduced traditional consumption of fuel to lessen emission, and to reduce environmental pollution reported in [1] and [2].

One of the significant problems in an electric vehicle is its short driving range due to the lack of battery capacity, inadequate charging infrastructure and long charging time evaluated in [10].

Reducing charging time or fast charging of vehicle technology requires a high power charger. Increasing the power level increases the size of the charger. Hence, reduced size, compact design, and fast charging technology can be achieved by high power and high switching frequency application. High power silicon carbide (SiC) module makes it achievable to get the high-frequency operation that reduced the size of the components of the charger for a compact design. On the contrary, high switching frequency causes high switching and conduction loss and affects the efficiency of the system mentioned in [3] and [4].

Several control schemes, described in [11] - [14], suggest the possibility to control the dc-link voltage at the desired level. Besides, current SiC power module, detailed in [5], makes it possible of high voltage application up to 1.7 kV shown in [15] and [16]. It is also mentioned in [16] that maximum efficiency happens at an operating power level lower than the rated power level of the converter for high voltage high-frequency application. For example, it informed about 98.7% efficiency for 40 kW operation without the inclusion of auxiliary power loss of the converter designed for 100 kW of rated power. Furthermore, in [7] and [8], different DAB models have been proposed to estimate the current for loss calculation. Hence, it is required to find a suitable DAB model for loss calculation of high power, high voltage and high-frequency application needed for todays fast charging application with a frequency range of efficient operation.

In this paper, the semiconductor losses of a 100 kW dc-dc converter and it's efficiency are calculated in details considering the two different models of single-phase DAB converter. Based on the result, the optimal switching frequency range has been determined to reach the highest efficiency at different power levels. Finally, the maximum junction temperature is investigated for determining heat sink requirement.

II. Essential Parameters For Semiconductor Losses

Conduction and switching losses are the two losses contribute to the semiconductor losses. Accurate estimation of these losses is fundamental to find the heat source as a power loss that causes the rise of the junction temperature of the semiconductor devices. Conduction loss is approximated according to

$$P_c = V_f I_{avg} + R_{on} I_{rms}^2 \tag{1}$$

where P_c, V_f, R_{on}, I_{avg}, I_{rms} indicate conduction loss, forward voltage drop, on-state resistance, average current, and RMS current respectively. Similarly, switching loss is calculated according to

$$P_s = f_s \left(E_{on} + E_{off}\right) \tag{2}$$

$$E_{on} = E'_{on} \left(\frac{I_s}{I'_s}\right)^{K_{I,on}} \left(\frac{V_s}{V'_s}\right)^{K_{V,on}} \tag{2a}$$

$$E_{off} = E'_{off} \left(\frac{I_s}{I'_s}\right)^{K_{I,off}} \left(\frac{V_s}{V'_s}\right)^{K_{V,off}} \tag{2b}$$

where P_s, E, f_s, V, I, K_V, K_I stand for switching loss, switching energy, switching frequency, switching voltage, switching current, voltage coefficient, and current coefficient respectively. E', V', and I' indicate the corresponding datasheet value of the power module in [5].

A. Parameters for Switching loss

The necessary switching loss parameters are f_s, selected based on application and device requirement, E_{on} and E_{off} calculated or estimated from the datasheet. E'_{on}, E'_{off}, V'_s, and I'_s are collected from device datasheet [5]. K_V and K_I are estimated by using some selected values of Vs and Is from the voltage-current plot of the device from the datasheet [5].

B. Parameters for Conduction loss

The datasheet parameters for the conduction loss calculation are V_f and Ron whereas V_f could be a fixed value or zero according to the modern device manufacturer. I_{avg} and I_{rms} are the currents depend on the topology of the circuit and the operating point of the application specific design. Mathematical techniques could be used to find the required average and RMS current for calculating conduction loss of the switch.

III. Mathematical Models Of Single Phase Dual Active Bridge Converter

The single phase dual active bridge (DAB) converter is a dc-dc converter uses for many application especially for the recent technology of e-mobility shown in Fig. 1. The input voltage, $V1$ is controlled constant using VSC based rectifier [6], and the output voltage, $V2$ is controlled at a specific ranges of voltages value based on the particular application. There are two half-bridge legs on each side of the isolating transformer, and an inductor, L with a leakage resistance, bridges them both. Each leg consists of two switches, and each switch contains one MOSFET and one antiparallel DIODE in it. The inductor, L could be leakage inductance of the transformer or a separate one.

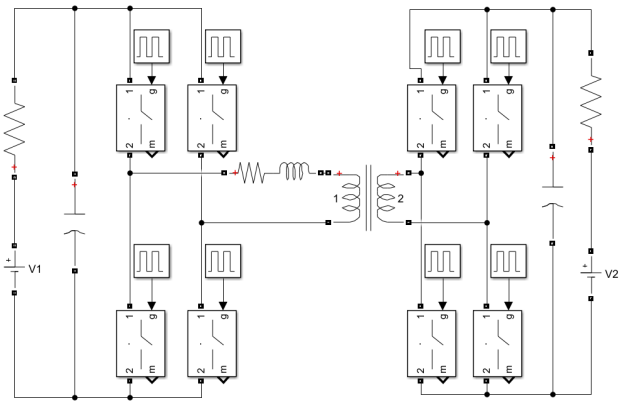


Fig. 1. Single phase dual active bridge dc-dc converter.

In Fig. 2, single phase DAB is modelled. The left terminal voltage of the bridge is V_{ac1}, which is equal to $V1$ or $-V1$ or $zero$ depending on the connection and combination of switches. Similarly, the right terminal voltage of the bridge is V_{ac2}, which is equal to $nV2$ or $-nV2$ or $zero$. n refers to the turn ratio of the transformer. Figure 3 shows the voltage and current of the inductor for unity turn ratio and equal terminal voltage of the DAB.

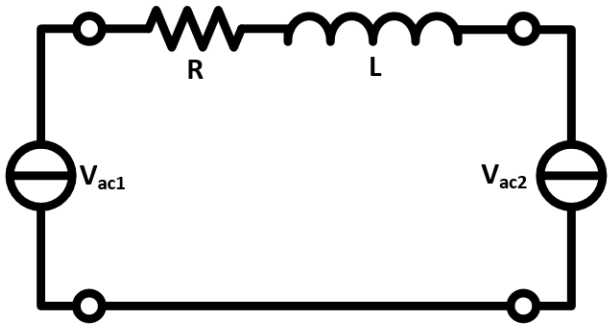


Fig. 2. Mathematical model of single phase DAB where $R = 0$ presents DAB loss-less (DAB-LL) model and $R \neq 0$ presents DAB conduction loss (DAB-CL) model.

Hence, the DAB model presented in Fig. 2 modelled two voltage sources at the inductor terminal along with inductor and resistor connected in series. Here amplitude of V_{ac1} = $V1$ and V_{ac2} = $nV2$. Value of L is selected in a way to be able to supply the maximum current to the load. R includes the on-state resistance of two switches in conduction, the leakage resistance of the inductor, and the resistance of the transformer. Based on the value of R whether it is $zero$ or non $zero$, the model in Figure 2 can be considered as DAB Loss-less (DAB LL) model and DAB conduction loss (DAB CL) model respectively [7]. These models are used to estimate the switching current, average and RMS current values for estimating semiconductor loss of DAB.

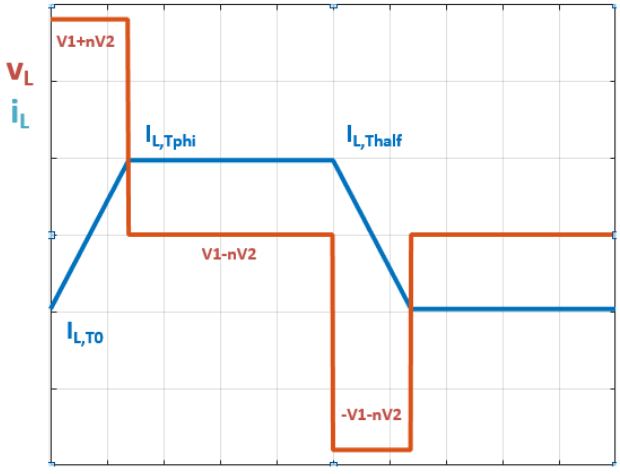


Fig. 3. Voltage and current of the DAB

A. DAB-LL Model

If $R = 0$ is considered then the DAB model in Fig. 2 becomes DAB-LL model. This model is useful to get the maximum possible stress value of the current for selection of the switches of the converter. The principal current equation for this model is described as

$$i_L(t) = \begin{cases} i_L(T_0) + \frac{M}{L}(t - T_0); & T_0 \le t \le T_\phi \\ i_L(T_\phi) + \frac{N}{L}(t - T_\phi) & T_\phi \le t \le T_{\frac{1}{2}} \end{cases} \quad (3)$$

where $M = V1 + nV2$, $N = V1 - nV2$, $T_0 = 0$, T_{phi} is the time of the phase shift which is about 49 degree for $100\ kW$ system in this study. This equation (3) is used for getting the switching currents and the conduction currents of the switches.

B. DAB CL Model

When $R \neq 0$, the DAB model in Fig. 2 becomes DAB-CL model and $R = 2Ron + R_L + R_{tr1} + n^2 R_{tr2}$. The base current equation, in this case, is derived as

$$i_L(t) = \begin{cases} e^{\frac{-t}{\tau}} i_L(T_0) + \frac{M}{R}(1 - e^{\frac{-t}{\tau}}); & T_0 < t < T_\phi \\ e^{\frac{-t+T_\phi}{\tau}} i_L(T_\phi) + \frac{N}{R}(1 - e^{\frac{-t+T_\phi}{\tau}}); & T_\phi < t < T_{\frac{1}{2}} \end{cases} \quad (4)$$

where $\tau = \frac{L}{R}$. This equation (4) gives the switching currents and the conduction currents for the semiconductor loss calculation for DAB-CL model.

The power in and power out of the DAB can be calculated using (5) and (6) below along with the related current from (3) and (4).

$$Pin = \frac{2V1}{T_s} \int_0^{\frac{T_s}{2}} i_L(t)dt \quad (5)$$

$$Pout = \frac{2V2}{T_s} \int_{\frac{T_s}{2}}^{T_s} i_L(t)dt. \quad (6)$$

IV. Derivation Of Required Current Equations

There are two switching techniques might be happened during the operation of the DAB converter known as soft switching and the hard switching. During soft switching, the current through the switch becomes zero before the turn-on of the switch because of the conduction of the anti-parallel diode. Neither voltage nor current become zero during hard switching. In the operation of the DAB converter, as long as there is a good control for the switching circuit, soft switching is evident, and hard switching is rare. Nevertheless, hard switching is considered to get the worst case possible power loss for the semiconductor device. It would be useful for the thermal analysis of many applications of DAB converter [9].

For the DAB-LL model, the switching currents to calculate the switching losses are listed below:

$$I_{L,T0} = \frac{-\pi N - 2\phi nV2}{4\pi f_s L} \quad (7)$$

$$I_{L,Tphi} = I_{L,T0} + \frac{M}{L} T_\phi \quad (8)$$

$$I_{L,Thalf} = I_{L,Tphi} + \frac{N}{L}\left(T_{\frac{1}{2}} - T_\phi\right) \quad (9)$$

For the DAB-CL model, the switching current equations are used for calculating switching losess as follows:

$$I_{L,T0} = \frac{-I_{L,T_{half}(I_{L,T0}=0)}}{1 + e^{(\frac{-T_{phi}+T_0}{\tau})} e^{(\frac{-T_{half}}{\tau})}} \quad (10)$$

$$I_{L,T_{phi}} = e^{(\frac{-T_\Phi}{\tau})} I_{L,T0} + \frac{M}{R}(1 - e^{(\frac{-T_\phi}{\tau})}) \quad (11)$$

$$I_{L,T_{half}} = e^{(\frac{-\frac{T}{2}-T_\phi}{\tau})} I_{L,T_{phi}} + \frac{N}{R}(1 - e^{(\frac{-\frac{T}{2}-T_\phi}{\tau})}) \quad (12)$$

where

$$I_{L,T_{half}(I_{L,T0}=0)} = e^{-\frac{T}{2\tau}} I_{L,T_{phi}(I_{L,T0}=0)} + \frac{N}{R}\left(1 - e^{-\frac{T}{2\tau}}\right) \quad (13)$$

and

$$I_{L,T_{phi}(I_{L,T0}=0)} = \frac{M}{R}(1 - exp(\frac{-T1}{\tau})) \quad (14)$$

The average and RMS current can be calculated using the piece-wise linear integration method. The integration of the current equation for a specific model would be done for the calculation of average value of the conduction current using the general equation below for a particular limit as:

$$I_{ab}(AVG) = \int_a^b i_L(t)\, dt \quad (15)$$

where a is the conduction start time of the switch and b is the conduction stop time of the switch.

The square integration of the current equation is done to find the RMS value of the current for loss calculation using equation below:

$$I_{ab}(RMS) = \int_a^b i_L^2(t)\, dt \quad (16)$$

Hence, the average and RMS value of the conduction currents for specific limits can easily be calculated according to:

$$I_{avg} = \frac{1}{T_s} I_{ab}(AVG) \quad (17)$$

$$I_{rms} = \sqrt{\frac{1}{T_s} I_{ab}(RMS)} \quad (18)$$

The switching currents, average currents, and RMS currents are calculated for the DAB-LL and DAB-CL models separately using corresponding current equations described earlier for the calculation of semiconductor losses of the DAB converter.

V. Results

The operating point considered for the converter to calculate the losses are as below:

$V1 = 700V$, $V2 = 700V$,

$I_o = 140A$, $P_o = 100kW$, and $f_s = 25kHz$.

The semiconductor device used is CREE (CAS300M17BM2) and its datasheet parameters for switching loss calculation are below:

TABLE I
PARAMETERS OF SWITCHING LOSS CALCULATION

Parmeters	Datasheet values	CREE (CAS300M17BM2)
E'_{on}	Turn-on Switching Energy (mJ)	13
E'_{off}	Turn-off Switching Energy (mJ)	10
V'_s	switching voltage (V)	900
I'_s	switching current (A)	300
$K_{V,on}$	Turn-on voltage coefficient	1.83
$K_{I,on}$	Turn-on current coefficient	0.647
$K_{V,off}$	Turn-off voltage coefficient	1.17
$K_{I,off}$	Turn-off current coefficient	1.26

The conduction loss parameters from data-sheet are forward voltage drop, V_f which is equal to zero for the chosen power module and the on-state resistance, R_{on} which is typically $8m\Omega$. Taking the variation of junction temperature into account, on-state resistance is considered as $12m\Omega$ for the maximum junction temperature of about 100 to 110^0C.

TABLE II
DAB EFFICIENCY FOR $R_{ds,on} = 8\ m\Omega$

Model		$P_{sc,DAB}$ (kW)	P_{RL} (kW)	P_T (kW)	P_{aux} (kW)	Efficiency (%)
DAB-LL	HS	3.070	0.083	0.158	0.020	96.60
	SS	1.707	0.083	0.158	0.020	97.99
DAB-CL	HS	3.0795	0.083	0.159	0.020	96.62
	SS	1.7146	0.083	0.159	0.020	98.00

The result of the analysis of the semiconductor losses and others losses are tabulated in Table II and Table III for the on-state resistance of $8m\Omega$ and $12m\Omega$. Soft switching and hard switching technique are considered along with the DAB-LL and DAB-CL model for the calculation of the semiconductor losses. Efficiency calculation considers an auxiliary loss of 20 W in all cases.

TABLE III
DAB EFFICIENCY FOR $R_{ds,on} = 12\ m\Omega$

Model		$P_{sc,DAB}$ (kW)	P_{RL} (kW)	P_T (kW)	P_{aux} (kW)	Efficiency (%)
DAB-LL	HS	3.5595	0.083	0.158	0.020	96.10
	SS	1.9518	0.083	0.158	0.020	97.74
DAB-CL	HS	3.5744	0.083	0.159	0.020	96.13
	SS	1.9644	0.083	0.159	0.020	97.75

It is evident from the tables that worst-case loss happened at the DAB-CL model and hard switching case. Losses could be further increased if the operating junction temperature rises beyond 110^0C. It is because of the positive temperature effect on on-state resistance of the switch. Hence, the efficiency of the DAB converter would be in between 96% to 98% for the models and switching schemes as mentioned earlier.

DAB-CL model is the more practical model then the DAB-LL model as each practical converter has losses during the operation. Presence of snubber capacitance and sound control circuit ensures the soft switching in the converter. Therefore DAB-CL and SS (soft switching) model is chosen for the determination of switching frequency range.

Fig. 4 shows the switching frequency versus efficiency plot. Changing frequency changes the operating power level as well because of the fixed parameters, especially inductance and phase shift, of the converter circuit. Fig. 4 also depicts that switching frequency range between 20 kHz to 70 kHz deliver the efficiency more than 97.56% with a power level between 124 kW to 98 kW. The reason for decreasing efficiency with increasing frequency is of fixed inductance, and phase shift.

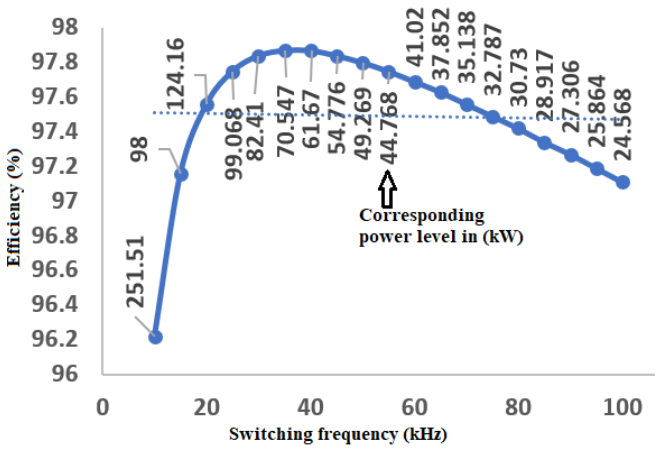


Fig. 4. Efficiency of the single phase DAB converter for different switching frequency and operating power level considering DAB-CL and SS model.

Fig. 5 shows the junction temperature for the same condition of the power level at the different switching frequencies. Junction temperature is 214^0C at 20 kHz which is 171^0C at 25 kHz, and that is about 94^0C at 70 kHz. As a result, to operate the converter at the rated power of 100 kW, there must have a heat sink to take care of the increased junction temperature.

VI. Conclusions

In this paper, semiconductor losses of the single-phase DAB converter of 100 kW are calculated for soft switching and hard switching scheme along with two models known as DAB-LL and DAB-CL models. It can be concluded from the analysis that higher operating junction temperature which raises the on-state resistance, and higher switching frequency causes higher semiconductor losses. The semiconductor losses have happened more for hard switching scheme than the soft switching scheme at the rated operating point of 100 kW, and 700 V. Achieving soft switching scheme requires a complicated control circuit that leads to additional cost but improves the efficiency of the converter. For a $100kW$ system

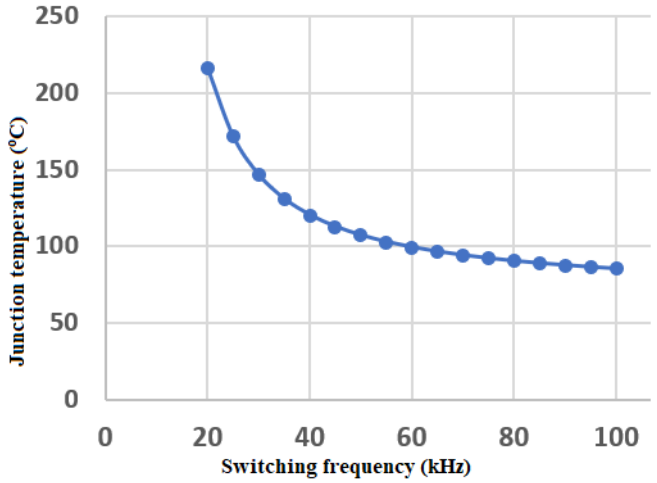


Fig. 5. Junction temperature at different switching frequency.

of DAB converter with $25kHz$ frequency, DAB-CL model soft switching scheme gives the efficiency of about $97-98\%$. On the other hand, the switching frequency can be ranged between 20 kHz to 70 kHz with a fixed phase shift angle of 49 degrees without sacrificing efficiency. However, the power level could be changed between 124 kW and 35 kW considering the on-state resistance of $12m\Omega$, junction to case thermal resistance of $0.061^oC/W$, and the maximum ambient temperature of 40^0C. Finally, at these conditions, the maximum junction temperature raises between 214 to 94^oC. Hence, a cooling arrangement must be needed for the rated operation of the converter to keep the junction temperature within the permissible limit.

ACKNOWLEDGMENT

Thanks to Saeid Haghbin of Elbind Elektronik AB, Sweden, Ola Carlson of Electric Power Engineering, Chalmers University of Technology, Sweden, Yasir Arafat of Department of Electrical and Electronic Engineering, International Islamic University Chittagong (IIUC), Bangladesh, and IIUC for their sincere support, facilitation, valuable comments, and cooperation.

REFERENCES

[1] Chapman, Lee. "Transport and climate change: a review." Journal of transport geography 15.5 (2007): 354-367.
[2] Santos, Georgina. "Road transport and CO2 emissions: What are the challenges?." Transport Policy 59 (2017): 71-74.
[3] Tate, E. D., Michael O. Harpster, and Peter J. Savagian. "The electrification of the automobile: from conventional hybrid, to plug-in hybrids, to extended-range electric vehicles." SAE international journal of passenger cars-electronic and electrical systems 1.2008-01-0458 (2008): 156-166.
[4] Haghbin, Saeid. "Design of a 300 kW Compact and Efficient Fast Charger Station Utilizing High-Power SiC Modules and Nanocrystalline Magnetic Materials." 2018 20th European Conference on Power Electronics and Applications (EPE'18 ECCE Europe). IEEE, 2018.
[5] Datasheet of CAS300M17BM2, [online] Available: www.cree.com
[6] Wu, Rusong, Shashi B. Dewan, and Gordon R. Slemon. "A PWM AC-to-DC converter with fixed switching frequency." IEEE Transactions on Industry Applications 26.5 (1990): 880-885.
[7] Krismer, Florian. "Modeling and optimization of bidirectional dual active bridge DC-DC converter topologies." Diss. ETH Zurich, 2010.
[8] Uddin, Mohammad Jalal (2018) Loss and thermal analysis of a 300 kW fast charging station. Gteborg : Chalmers University of Technology (Kandidatarbete - Institutionen fr elektroteknik, Chalmers tekniska hgskola, no: 256164).
[9] Mobarrez, Maziar, et al. "Performance and loss evaluation of a hard and soft switched 2.4 MW, 4 kV to 6 kV isolated DC-DC converter for a wind energy application." IECON 2012-38th Annual Conference on IEEE Industrial Electronics Society. IEEE, 2012.
[10] Artmeier, Andreas, Julian Haselmayr, Martin Leucker, and Martin Sachenbacher. "The shortest path problem revisited: Optimal routing for electric vehicles." In Annual conference on artificial intelligence, pp. 309-316. Springer, Berlin, Heidelberg, 2010.
[11] Parvez, M., S. Mekhilef, Nadia ML Tan, and Hirofumi Akagi. "Model predictive control of a bidirectional AC-DC converter for V2G and G2V applications in electric vehicle battery charger." In 2014 IEEE Transportation Electrification Conference and Expo (ITEC), pp. 1-6. IEEE, 2014.
[12] Lee, Dong-Choon, G-Myoung Lee, and Ki-Do Lee. "DC-bus voltage control of three-phase AC/DC PWM converters using feedback linearization." IEEE transactions on industry applications 36, no. 3 (2000): 826-833.
[13] Karlsson, Per, and Jrgen Svensson. "DC bus voltage control for a distributed power system." IEEE transactions on Power Electronics 18, no. 6 (2003): 1405-1412.
[14] Akter, Md Parvez, Saad Mekhilef, Nadia Mei Lin Tan, and Hirofumi Akagi. "Model predictive control of bidirectional AC-DC converter for energy storage system." Journal of Electrical Engineering Technology 10, no. 1 (2015): 165-175.
[15] Zhao, Tiefu, Jun Wang, Alex Q. Huang, and Anant Agarwal. "Comparisons of SiC MOSFET and Si IGBT based motor drive systems." In 2007 IEEE Industry Applications Annual Meeting, pp. 331-335. IEEE, 2007.
[16] Akagi, Hirofumi, Tatsuya Yamagishi, Nadia Mei Lin Tan, Shin-ichi Kinouchi, Yuji Miyazaki, and Masato Koyama. "Power-loss breakdown of a 750-V 100-kW 20-kHz bidirectional isolated DCDC converter using SiC-MOSFET/SBD dual modules." IEEE Transactions on Industry Applications 51, no. 1 (2015): 420-428.

Land-Use Detection Using Residual Convolutional Neural Network

Shuvendu Roy
dept. of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
bikash.shuvendu@gmail.com

Sneha Paul
dept. of Urban and Regional Planning
Khulna University of Engineering & Technology
Khulna, Bangladesh
snehapaulputul@gmail.com

***Abstract*—Land-Use detection is an important application in the context of remote sensing. This objective is to classify, a chunk of a satellite or high-earth orbit image of the earth, as the type of land-use. In this work, we have used recent advancement in computer vision with deep learning to improve the performance to the classification task. We have used a very deep convolutional neural network with the residual connection as a building block and the concept of transfer learning to train this model. The experiment shows that the performance of the classifier is very high with transfer learning on a pre-trained model rather than training from the sketch with such small dataset that is available for land use classification task. Our proposed approach generated a better result than the previously existing method in benchmark dataset.**

***Index Terms*—Convolutional Neural Network(CNN), Transfer Learning, Fine Tuning, Land-Use Classification**

I. Introduction

The land-use detection is the task of identifying what an area of land is used for. This is an important application of remote sensing. This is normally treated as a classification task. Recently few datasets are collected with a large number of labeled image of aerial photograph and many classification algorithms are used to improve the accuracy of such algorithm.

On the other hand, deep learning is an interesting and powerful development in recent years. It used the concept of neural network and backpropagation to train it to approximate any function. Convolutional neural network is the most efficient building block to deal with the image in deep learning. Recently deep learning is used for land-use classification of HSR imagery [1]. The technique such as autoencoder [2] and a variation called sparse autoencoder is used previously. Also, the deep belief network [3] was used in this context. The performance was later improved with the introduction of convolutional neural enwtork [4]–[6]. Although these models improved the performance of the application these regular convolutional neural network could not better than that as those models were not very deep. It is a common idea in deep learning that the deeper models should have better performance. But in practice, we can not just stack convolution layer on top of one another to improve the performance.

Over the year many research has proposed different variations of regular CNN to improve its performance. Recently a model called AlexNet [7] has proposed a successful deep architecture to train a big CNN on ImageNet dataset. This model is very deep and the trained with this very large dataset. One other problem with this big model is that it does not generalize well with the small dataset as these models have too many parameters to train. For this reason, a new technique called transfer learning became popular. The idea of transfer learning is to use the knowledge learned in one application in other similar applications. The transfer learning was used in land-use classification by Zhao et. al. [8]. They used the AlexNet as the base architecture and applied transfer learning on it. This shows the impressive result of this application.

Recently a new development called Residual Connection introduced by Resnet [9] has shown very impressive performance in image classification application. This uses the concept of skip connection to pass extra information to the layer. That gives the opportunity to train even deeper models without overfitting. In this work, we have explored the transfer learning technique on the residual convolutional neural network to improve the accuracy of classification of land-use.

The rest of the paper is organized as follow- Section II describes the basic of CNN, followed by our model architecture and training procedure. Section III analyses the experimental result of our proposed approach.

II. Methods

The convolutional neural network is the central building block to deal with image data with a neural network. Over the years many research has proposed various variations of the convolutional neural network to improve the performance. In the following section, we start by describing the general principles of the convolutional neural network(CNN). Then we describe the specific architecture used in our method to improve the performance of land use classification using CNN.

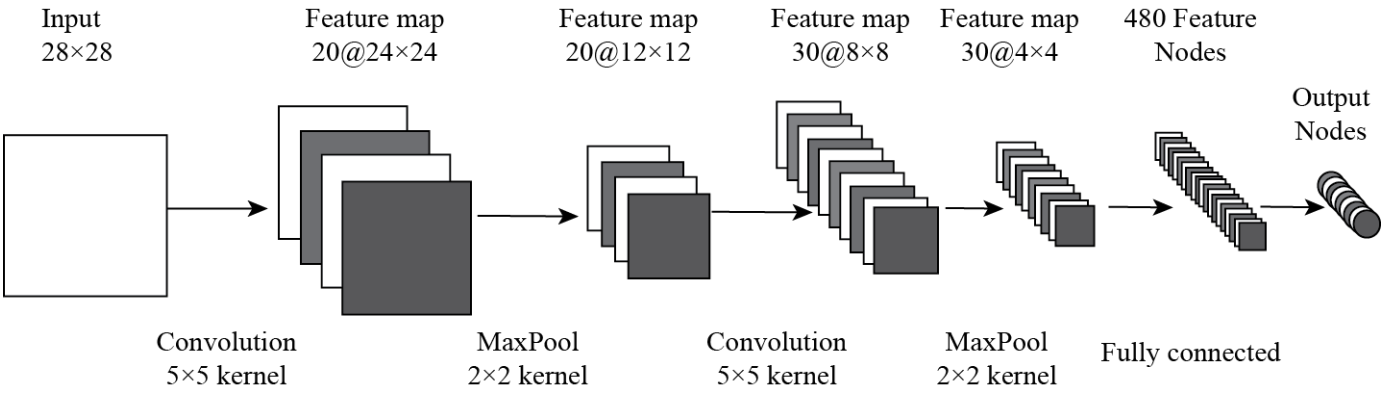


Fig. 1. An illustration of applying the convolutional neural network on a 28×28 size image.

A. Convolutional Neural Network(CNN)

The CNN [10] is a neural network based algorithm that takes an image as an input and process with learnable weights and biases to output the probability of various aspect about the image. The CNN based approach much less preprocessing on the actual image because the CNN can work with a plain image of any kind without any hand engineering. While in classical image processing the filters, that are used to processes the image is pre-defined and hand-engineered, the CNN treat the filters and learnable parameters and learn the values for those.

The kernel is basically a small size matrix that slides over the image column wise then row-wise. The step is controlled by a parameter called stride. The kernel has learnable weight. When the kernel is passed over an image the kernel weight values are multiplied with the corresponding values of the image pixels to calculate one value of the output feature map. The feature map is also a matrix which can be treated as an image but it is commonly called a feature map.

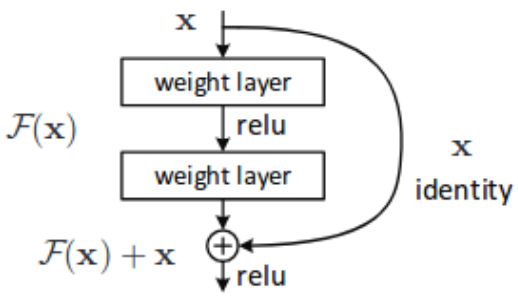


Fig. 2. A residual connection.

Another parameter of convolution is stride. The stride defines the amount of skip the kernel give for the next position. For example, with the stride size of 2, the kernel will mode 2 pixel right to generate the next pixel value in the feature map. As a result, the length of the feature map is half the size of the input length. One more parameter the determine the length of the output feature map is padding. When the kernel slide over the image, for the first position the kernel must fit the image at the corner to output the first position of the feature map. For a kernel size of 3×3, the first output of the feature map comes from the first 3 pixels. So, the feature size is 2 less than the original image for the first and last image as the kernel can not fit here. Padding defines whether to add extra dummy space in both side to keep the size of the output same as input or let the image size as it is generated by simply applying the convolution operation.

Like any other neural network, the CNN is consist of layers. The layers come on top of the previous one. The most common types of layers used in CNN are convolution layer, ReLU layer, pooling layer, and fully connected layer. The convolution layers have the kernel and the main convolution is done in this layer. Often the image size is changed in this layer. The ReLU layer is responsible for applying the nonlinearity. This is an element with activation function formulated as $max(0, x)$. The pooling layer is where the main downsampling occurs. One of the main concepts of CNN is to reduce the size of the image and finally output in some low dimensional space. The pooling layer reduces the size of the image along row and column dimension. The most common form of pooling is max pooling and average pooling. The max pooling takes the maximum values of the pixels in a given area. Finally, the fully connected layer converts the matrix into a linear vector and perform the regular linear combination as like the regular artificial neural network.

Fig. 1 shows one example of CNN on a 28×28 size image. The input size is $(28 \times 28 \times 1)$. Here 1 is the channel dimension. For a color image, it will be 3. The first convolutional layer contains the kernel of size 5×5. The output of this layer is $(24 \times 24 \times n1)$. Here n1 is the number of kernels applied in this layer. Each kernel generates one separate feature maps and all those are stacked vertically in

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				

Fig. 3. The model architectures of different variations of ResNets.

the depth dimension. The output has depth equals the size of kernels in the previous layer. The transition from 28 to 24 is due to the kernel size of 5 with no padding. This is the input to the first max polling layer. This is a max- polling with kernel 2×2. This reduces the output size to the half of the input size. The output of this layer is $(12 \times 12 \times n1)$. Two next convolution and max-polling of the same size reduces the feature map size to $(12 \times 12 \times n2)$ to $(8 \times 8 \times n2)$. This is then passed to a fully connected layer. The next fully connected layer output the probability of all classed in the dataset.

B. Land use detection using residual CNN

The basic concept of deep learning is to use a deeper model for getting better performance. But simply adding more layer does not help at all. For an image, the application is hardly useful to add a more convolutional layer into the network. But the deeper network is required for processing high dimensional input as it contains more information. Also, a deep network has more scope for learning more complex pattern in the data. The concept of ResNet [9] introduced a novel approach for that.

Instead of learning the mapping from x to y directly with a function $H(x)$ the residual block use $F(x) = H(x) - x$. Which implies $H(x) = F(x) + x$. The residual network suggests that the mapping of residual function $F(x)$ is easier to learn than the original $H(x)$. So, the way the residual connection is made is by adding the input to a layer with the output of that layer. The intuition of such connection is that the layer can see both the transition and the change with the original input. So, the layer has more information to learn from. In practice, such a residual connection is shown to avoid vanishing gradient problem in the large neural network. This is the way the regular deep CNN model with or without residual connection has a huge difference in performance.

Fig. 2 show a residual connection. The left shows a simple forward transform. The right shows the modification to have a residual block. This connection is called skip connection

The residual connection allows very deep convolutional neural network trainable. The original paper showed the performance of deep CNN with the residual connection. It proposed 5 difference architecture with residual connection, namely Resnet-18, Resnet-34, Resnet-50, Resnet-101, Resnet-152. The number after all the model is the number of layers in the model.

Fig. 3 shows the model architectures of all the ResNet variations. All the model contains the same kind of model structure with a variation in the number of repetition of the same layer.

In this work, we have used the residual neural network as a building block for our classifier. But instead of training the model from the scratch we have used the concept of transfer learining [11], [12]. Transfer learning is the concept of using the knowledge learned in one application in some other application. This is a useful application for dealing with large CNN. As a large network contains a lot of parameters, training the entire network from sketch requires a lot of data. For many application, such a huge amount of data is not always available. In such a case, transfer learning is a useful tool. The basic intuition of transfer learning in computer vision is that the patterns learned by the shallow(earlier) layer of CNN are basic shape or lines. These are the general idea of an image and not application specific. So the knowledge learned from facial image classification can be used in land use classification. But the features learned by the higher layers are application specific complex features.

To apply the concept of transfer learning in our application we have used pre-trained ResNet model which are trained

with large general purpose dataset called ImageNet. This is a very large dataset with 1000 classes. We have updated the final fully connected layer of this model to match our output class number which is 21 in this case. Re-training a fully trained model is called fine-tuning. In this case first, we have trained the final untrained layer. The gradually trained the pre-trained portion of the model.

III. Experiments

We have implemented the model with pytorch [13] int python. All experiments were done in NVIDIA 1060 GPU.

A. Experimental Setup



Fig. 4. Sample images from the UCM land-use dataset.

For the land use application, we have used UCM dataset [14]. The dataset contains images of 21 different classes each containing 100 images with a total of 2100 images. The classes are - agricultural, airplane, baseballdiamond, beach, building, chaparral, denseresidential, forest, freeway, golfcourse, harbor, intersection, overpass, parkinglot, mobilehomepark, runway, sparseresidential, storagetanks, mediumresidential, river, tenniscourt. The size of the images is 256×256. The image was collected from large images in the US Geological Survey National Map Urban Area image collection of various cities in the US. To train this model we have randomly split the data into 90% for training and 10% for testing. Few examples of the training data are shown in Fig. 4.

We have trained the model with the image size of 256×256 with the learning rate of 0.0005 in Adam optimizer. The Adam optimizer has two hyperparameters that control the momentum to decay the learning rate. They are called beta1 and beta2. The corresponding values are set to 0.9 and 0.999. Table I summarized the parameters of the training process.

To avoid over-fitting we have used the concept of augmentation. The idea of image augmentation is to transform the image slightly to generate a slightly new image. As a result, the model will never see the actual image twice. It helps to reduce over-fittings. The over-fitting is done by introducing a horizontal or vertical flip, rotation, translation, zooming etc. The augmentation setting for our experiment is summarized in the table. II

TABLE I
Values of Training Parameters

Parameters	Values
Image input size	256×256
Optimizer	Adam
Adam: beta1	0.5
Adam: beta2	0.999
Epochs	20
Batch size	16
Learning rate	0.0005
LR Decay rate	0.99

TABLE II
Values of Augmentation Applied to the Training Image

Type of Augmentation	Values
Rotation	-45° to 45°
Horizontal Flip	Yes
Scaling factor	×1.1

B. Result Analysis

As mentioned in the method, we have conducted an experiment on five different model architecture.

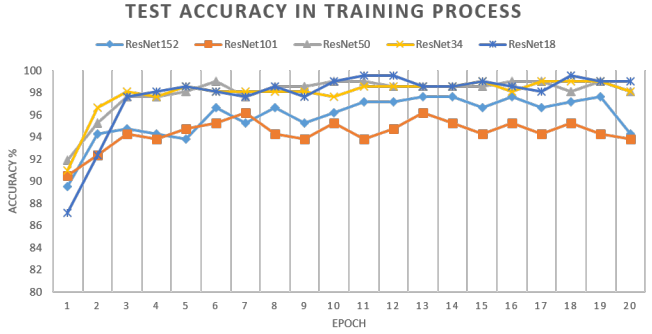


Fig. 5. Accuracy as a measurement in training process in all models.

Figure 5 shows the convergence of accuracy during training on all these model.

To further analyze the result we have reported the result in average precision, recall and accuracy of all model. Precision value for a class is the ratio of the total correctly classified image of that class and total output labeling as that class. Whereas the recall is the ratio of total correctly classifier image of that class and total images of that class. The reported precision and the recall is the average overall class. The accuracy is the overall accuracy of that model. Fig. 6 shows the result.

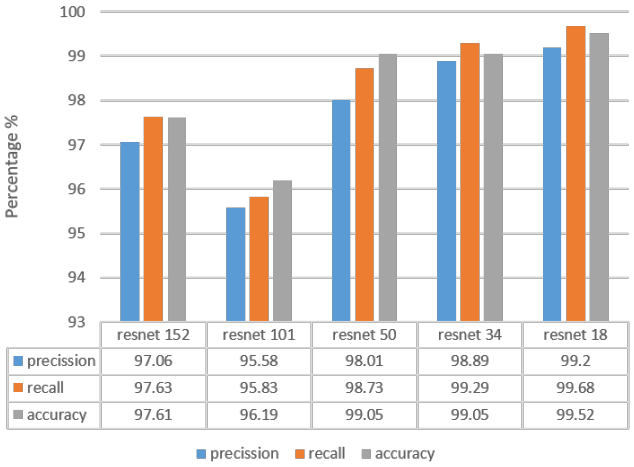


Fig. 6. Precision, Recall and the accuracy of all models.

From the result above we see that the best result is achieved from the ResNet18. Although the largest version is ResNet152. This is the consequence of small dataset. As we have mentioned above, a small dataset cannot generalize the model rather overfit the model. The whole point of using transfer learning is to solve this problem. Still, with the pre-trained model, we can not choose the largest available model for our application because of data shortage.

Table IV shows the accuracy of each class by each model. As we can see from the table, apart from a few complex class most of the classes have 100% accuracy on the test. For example, the class 'intersection' has low accuracy because 'intersection' and the 'freeway' are both images of roads. So, the model is confused by the similarity. In most other cases the features are easily separable.

Table III shows a comparison of accuracy on UCM dataset on all previous method. To match with the experiment of all other methods we have conducted the experiment with random 90% training and 10% test dataset for 10 times. The reported accuracy is the average of 10 runs.

TABLE III
ACCURACIES OF PREVIOUS METHOD ON UCM DATASET.

Methods	Accuracy
SPCK++ [14]	77.38
SSBFC [15]	91.67
DMTM [16]	92.92
LFK [17]	91.63
CNN with overfeat feature [18]	92.4
Multiview Deep learning [4]	93.48
GBRCN [6]	94.53
CNN-FT-FULL [8]	96.67
Our proposed method	97.55

From the table III above we can conclude that our method shows clear improvement over all existant method that is used in UCM dataset.

We have also compared the accuracy of different classed with different ResNet model architecture. The result is summarized in table IV.

IV. Conclusion

In this work, we have proposed a new approach for improving the classification accuracy of an important application of land-use classification. The experiments have shown the effectiveness of our approach compared to the previous approaches. Our work will help further improvement in the field of remote sensing with computer vision.

References

[1] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 22–40, 2016.

[2] X. Yao, J. Han, G. Cheng, X. Qian, and L. Guo, "Semantic annotation of high-resolution satellite images via weakly supervised learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3660–3671, 2016.

[3] Q. Zou, L. Ni, T. Zhang, and Q. Wang, "Deep learning based feature selection for remote sensing scene classification." *IEEE Geosci. Remote Sensing Lett.*, vol. 12, no. 11, pp. 2321–2325, 2015.

[4] F. P. Luus, B. P. Salmon, F. Van den Bergh, and B. T. J. Maharaj, "Multiview deep learning for land-use classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 12, pp. 2448–2452, 2015.

[5] A. Romero, C. Gatta, and G. Camps-Valls, "Unsupervised deep feature extraction for remote sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 3, pp. 1349–1362, 2016.

[6] F. Zhang, B. Du, and L. Zhang, "Scene classification via a gradient boosting random convolutional network framework," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 3, pp. 1793–1802, 2016.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[8] B. Zhao, B. Huang, and Y. Zhong, "Transfer learning with fully pretrained deep convolution networks for land-use classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 9, pp. 1436–1440, 2017.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[10] Y. LeCun, Y. Bengio *et al.*, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.

[11] L. Torrey and J. Shavlik, "Transfer learning," *Machine Learning*, pp. 1–22, 2009.

[12] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Advances in neural information processing systems*, 2014, pp. 3320–3328.

[13] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," 2017.

[14] Y. Yang and S. Newsam, "Spatial pyramid co-occurrence for image classification," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1465–1472.

[15] B. Zhao, Y. Zhong, and L. Zhang, "A spectral–structural bag-of-features scene classifier for very high spatial resolution remote sensing imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, pp. 73–85, 2016.

[16] B. Zhao, Y. Zhong, G.-S. Xia, and L. Zhang, "Dirichlet-derived multiple topic scene classification model for high spatial resolution remote sensing imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 4, pp. 2108–2123, 2016.

[17] B. Zhao, Y. Zhong, L. Zhang, and B. Huang, "The fisher kernel coding framework for high spatial resolution scene classification," *Remote Sensing*, vol. 8, no. 2, p. 157, 2016.

TABLE IV
EACH CLASS ACCURACY FOR ALL MODEL.

Class	ResNet152	ResNet101	ResNet50	ResNet34	ResNet18
agricultural	100.0	100.0	100.0	100.0	100.0
airplane	100.0	100.0	100.0	100.0	100.0
baseballdiamond	100.0	100.0	100.0	100.0	100.0
beach	76.92	84.61	100.0	100.0	100.0
buildings	100.0	92.3	100.0	100.0	100.0
chaparral	100.0	100.0	100.0	100.0	100.0
denseresidential	80.0	80.0	80.0	100.0	100.0
forest	100.0	100.0	100.0	100.0	100.0
freeway	100.0	100.0	100.0	100.0	100.0
golfcourse	100.0	83.33	100.0	100.0	100.0
harbor	100.0	100.0	100.0	100.0	100.0
intersection	93.33	86.66	93.33	93.33	93.33
mediumresidential	100.0	100.0	100.0	100.0	100.0
mobilehomepark	100.0	100.0	100.0	100.0	100.0
overpass	100.0	100.0	100.0	100.0	100.0
parkinglot	100.0	100.0	100.0	100.0	100.0
river	100.0	85.71	100.0	100.0	100.0
runway	100.0	100.0	100.0	100.0	100.0
sparseresidential	100.0	100.0	100.0	100.0	100.0
storagetanks	100.0	100.0	100.0	100.0	100.0
tenniscourt	100.0	100.0	100.0	100.0	91.67

[18] D. Marmanis, M. Datcu, T. Esch, and U. Stilla, "Deep learning earth observation classification using imagenet pretrained networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 105–109, 2016.

Pathological Voice Classification Using Deep Learning

Shuvendu Roy, Md Ijaj Sayim and M.A.H. Akhand
dept. of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
bikash.shuvendu@gmail.com, ijajsayim@gmail.com and akhand@cse.kuet.ac.bd

***Abstract*—Voice classification task deals with sequential data. This is well known that this type of data is well processed by a recurrent neural network. In this work, we showed that in case of longer sequence convolutional neural network can give better accuracy. Whereas the recurrent network suffers from vanishing gradient problem even with a complex model like Long Short-Term Memory(LSTM). To illustrate the method we used pathological voice detection task. It is a type of problem in human voice caused by the internal defect in the throat and very hard to detect. In this work, we experimented with low dimension feature to compare both models rather than focusing on improving the overall accuracy.**

***Index Terms*—CNN, RNN, Voice Classification**

I. Introduction

In the throat, there are some folds of tissues named vocal folds, also known as vocal cords which are the key sources for generating sound for human through vocalization. Human voice may be tempered by different phenomena like their profession- singers, manufacturing supervisors, and some organic pathologies. These pathologies may affect the human sound system that modifies the structure of vocal folds which results in unnatural vibration sequence which may cause pathological voices [1].

Several methods could be applied to evaluate whether the voice is from a normal vocal system or pathological vocal system. Usually, the medical community uses the direct method which requires observation of vocal folds which is not so comfortable for the patients. Along with these, there are some methods like GRBAS and RBH mentioned here [2], [3], that require a direct audition of doctors. To get a better approximation using these techniques highly trained and experienced doctors are required.

A normal human being can differentiate pathology voices to some extent but we can't rely on it int etiologic procedure. Human voice carries several features which can be used to differentiate or classify the voice pathology. Here come different modern deep NNs like ANN(Artificial Neural Network) and RNN(Recurrent Neural Network) which can classify the voices based on those features. There are several ways for extracting features from voice, for instance, modulation spectra mentioned here [4] and MFCC(Mel-Frequency Cepstral Coefficients) which is described here [5].

The short term power spectrum of a particular sound is represented by MFCC in the processing of sound. It depends on LCT(cosine transform) and log power spectrum on a frequency which is scaled as Mel-Frequency. MFCCs are derived from the cepstral representation of the input voice as sound. This may help to collect important acoustic events, for example, phonemes. These MFC coefficients are actually the input features in common speech recognition [6] systems, such as models that can automatically identify numbers spoken into a device like a telephone. The basic distinction between cepstrum frequency and the Mel-frequency cepstrum is that the input sound frequencies are equally separated on a predefined scale known as Mel-scale. To extract features from MFCC we need to go through some mathematical details. Nevertheless, we can easily obtain these attributes using a python package name LibROSA. We also used this technique, described later, to extract features. We used these features as the input of ANN and RNN in our approach.

The recurrent neural network is a well-known technique for working with sequential data. But it has some drawback. With length getting longer it suffers from vanishing gradient problem and it becomes harder to train longer sequence. To solve this issue many variations of RNN is proposed. Among them long short-term memory(LSTM) and gated recurrent unit(GRU) is common. In this work, we have shown that it is comparable or sometimes better to use CNN in such application than the RNN as CNN is faster and no such vanishing gradient problem.

II. Related work

In pathology voice detection literature several methodologies studied and implemented. Zulfiqar, Ali et al. worked on speaker identification and used Mel-frequency Cepstral Coefficients for attribute collection and VQ technique for differentiating speakers [5]. Firstly, they used the MFCC technique to obtain feature vectors. Then, they used the VQ technique for detecting speakers.VQ refers to classical to

a quantization method from the field of signal processing which helps to develop a model of PDF into vectors. It turns out partitioning a massive set of array into different groups having apparently the same amount of arrays nearest to them. Every group is referred to as centroid in various clustering algorithms.

Patricia Henrquez et al. studied the effectiveness of six non-linear chaotic measures on the basis of non-linear dynamics theory to differentiate between 2 levels of voice sound: healthy and pathological [7]. They found that human voice can be expressed as non-linear time series, so, they used a non-linear system such as chaos theory and extract features from the sound system. This method has several steps, including extraction of values from mutual information function, calculating correlation among feature dimension and entropies. Here, entropy tells the amount of disorder of a system. They studied Shanon and Rnyi etropy estimator. In short, they divided their work into three portion- pre-processing, here they basically process the signal, feature extraction, and classification. For classification, they used standard NN(neural network).

Zoran Milutinovic divided the voice pathology into an organic and functional group and they focused on etiologic factors for clinics [8]. Along with these, Maria Markaki and Yannis Stylianou considered modulated spectrum for detecting pathology and classifying it [4]. This method produces a high dimension space and for reduction hey proposed SVD and a method for selecting features on the basis of correlated information between a subjective signal of the input voice and calculated attributes. After feature extraction and dimension reduction, they applied a classifier named support vector machine, and used RBF as the kernel of the classifier.

Along with these, Al-Nasheri, Ahmed, et al. worked on pathology classification using auto-correlation and entropy for feature extraction [9]. They basically worked on the improvement of attribute extraction so that they could have increased accuracy of classification.

Their working pipeline could be summarized as follows, voice sound was used as signal and they pass it through bandpass filters, then passed it through small frames, after framing they used two methods to obtain attributes i.e; auto-correlation and entropy, finally, using these attributes they classify voice signal using a support vector machine(SVM).

Alireza Afshordi Dibazar also proposed a pipeline for detecting pathological speech [10]. They also used MFC coefficients as well as pitch dynamics and modeled it by a Gaussian mixture. After that, they used a Hidden Markov Classifier(HMM). Moreover, in [11] authors also proposed Gaussian mixture models and short term cepstral features for the same work mentioned. In [12] voice pathology differentiation system is implemented on the basis of local discriminant wavelet packet, for feature selection purpose they employed GA(genetic algorithm) and used SVM (support vector machine) and RBF(Radial Basis Function) kernel for the classification task.

III. Method

In this work, we experimented on two different approaches to voice classification. The data we worked with is variable length. Traditionally recurrent neural network is a preferred method for modeling variable length data. But we also experimented on Convolutional Neural Network(CNN) which is well known for fixed length data, like an image. Here we describe the theory of the used method starting with Long Short-Term Memory(LSTM) which is a variation of the recurrent neural network we worked with. Then we describe the CNN and the specific network architecture which we experimented on.

A. Recurrent Neural Network

The recurrent neural network is a sequence processing model. The input to this model is a metric of dimension sequence length(t)$\times$ length of each sequence(d). The model is basically a single layer which is folded over time step. The same block of computation is repeated for all time step. At each time step(i) one sequence vector of length d is passed through the model which outputs a hidden state h_i and an output o_i. Fig. 1 shows the unfolding process of RNN processing. Which shows that single RNN block is unfolded over time e handle the sequence of variable length. The recurrent neural network has memory to remember information from the previous time step that gives the recurrent neural network advantage over other conventional neural networks. These are the network with loops inside of it, allowing information to persist. RNNs combines input with a state vector to produce a new state vector. This new state vector contains the information about the entire past sequence. The formula for vanilla RNN is as follow

$$h_i = tanh(W_{hh}h_{i-1} + W_{xh}x_i) \quad (1)$$

$$y_i = W_{hy}h_i \quad (2)$$

To use RNN as a building block of a classification model the sequence is converted into a vector of size d by the RNN. On top of this, a dense layer is added. The input to the dense layer is size d and the output of the dense layer is size n, where n is the number of output classes. To represent the output as a probability as softmax layer is added on top of the dense layer which represents the output into a vector whose some of all elements is 1.

With the sequence getting longer the regular RNN shows a drawback of vanishing gradient, which means the gradient cannot flow properly from very last sequence to the start sequence and the information is lost. For this reason, the vanilla RNN is not very good at modeling long sequences.

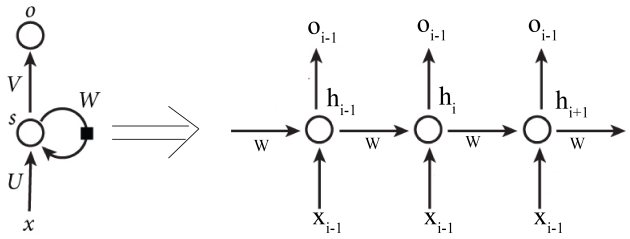


Fig. 1. Processing of sequential data in time step by unfolding a recurrent layer in the recurrent neural network. Which gives a single output for each time step that summarizes the information until that point of time. For our classification purposed we only need the last output as it summarizes the information of full sequence

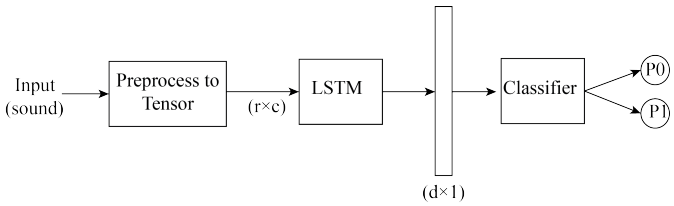


Fig. 2. RNN classifier. The input is a voice data, that is preprocessed to geenrate matrix of size (r×c) and passed through LSTM to generate vector of size $(d \times 1)$and it is the input to a classifier to generate the probability of each output classes

To solve this issue few variations of RNN were proposed. Among them LSTM [13] and GRU [14] are the two most popular model. Long short-term memory network(LSTM) is a variation of RNN, that is capable of learning long-term dependency. LSTM is ground up designed to solve long-term dependency problem. Remembering is its default behavior and it has a forget gate to forget information that is not important anymore. The full model is trained end to end. So, the model can learn that to remember and what to forget by itself. The key idea of the LSTM model is the cell state. With only some minor linear interaction the cell state flows straight through the sequence almost unchanged. One other important part of this model is the 'gate' concept. The information is added or removed from cell state carefully regulated by this gates. Formulas for updating the LSTM cell are

$$f_i = \sigma(W_f.[h_{i-1}, x_i] + b_f) \tag{3}$$

where x_t is the input and h_i is the hidden state from the previous time step $i-1$ that is passed to it. And the computed hidden state at this state will be kept in his variable to access it in the next time step.

The cell state is represented as C_i, which is responsible to calculating current output and hidden state. The cell state is updated as follow

$$I_i = \sigma(W_j.[h_{t-1}, x_t] + b_j) \tag{4}$$
$$C'_i = tanh(W_C.[h_{i-1}, x_t] + b_C) \tag{5}$$
$$C_i = f_t * C_{i-1} + I_i * C'_i \tag{6}$$

where * is a point-wise multiplication operation, W_j and W_C are weights of LSTM with b_j and b_C are the biases. σ and $tanh$ are two activation function of two layers. Finally, C_i is calculated as the weighted sum of the current cell state and previous cell state.

With this we can calculate the output and hidden state at time step t

$$o_i = \sigma(W_o.[h_{i-1}, x_i] + b_o) \tag{7}$$
$$h_i = o_i * tanh(C_i) \tag{8}$$

Here, W_o and b_o is the weights of the output layer. The sigmoid activation is used on top of the linear output. The output is calculated as the weighted multiplication of cell state and the output. A $tanh$ activation is used on top of that which pushes the valued between -1 and 1. We will treat this output of each time step as a feature vector for our application.

The output of the last layer of the LSTM block is of size d, which is a defined output vector length of the LSTM. This is passed to a classifier layer that outputs the vector of length n. Where n is the number of output class in the classifier application. In this case, the output is 2, which are the probability of normal voice and pathological voice. Fig. 2 shows how to RNN calculates added with a dense layer classify the pathological and voice data.

B. Convolutional Neural Network

Convolution neural network is based on the concept of convolution operation from image processing. But the main difference from classical image processing is that the kernels are not predefined to serve any specific purpose, rather it is learned by back propagation based on a loss function.

The CNN model is consists of an input layer, output layers, and some hidden layers. In this case, the hidden layers are all convolution layers and some helper layers of the convolution which are pooling layer and fully connected layers. In this experiment, we used two layers of convolution in the model. The first is passed to the first layer of the model and the generated output is passed to the next layers of the model and so on until the final fully connected layer.

The convolution block is the central building block of a CNN model. The layer is consists of a learnable kernel of filters whose parameters are adjusted during the training. The kernel has a small receptive field which is extended over the entire image volume by sliding through the row and column dimension. During the forward pass, the kernel is passed through the image to generate an activation map of two dimensions. As a result, the filter learns to activate at the particular scenario of having a specific feature at a particular portion of the image.

The size of the output generated by convolution operation is controlled by 3 parameters- depth of output layer, stride, and padding.

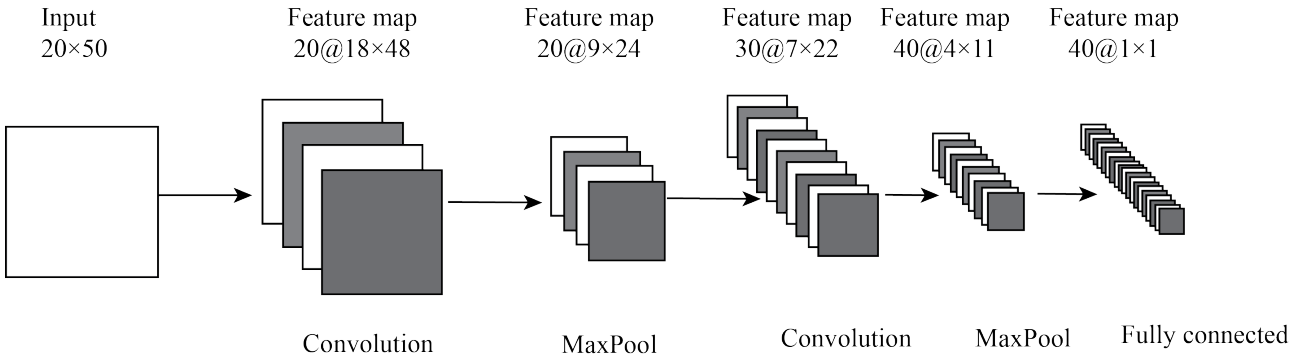


Fig. 3. Passing input data along the convolution layer. The oprations are given in the bottom and the size of the output feature map is shown of the top of the figure

Depth of the output controls the connected neurons in a layer to the same region of the input volume. These neurons combinedly learn to activate at a particular feature. If an input image is passed to this layer, a different combination of edges, shape, and texture of the input image will activate different neurons along the depth dimension.

Stride controls how the kernel slides around the image along the depth dimension. When the stride is 1, the kernel is moved one pixel to the right at a time and after going through all column it goes one pixel down in row dimension. When larger stride size is used the kernel jumps long over the image resulting in an output where the receptive field overlaps and has less correlation. That is why it is very uncommon to use stride size of more than 2.

Padding is used to add extra row and column around the image to let kernel output the first data at the first position of the image by providing the extra data for the kernel to fit outside the image. Sometimes it is important for output size of the image being equal to the input size. In such case, padding is applied.

If convolution operation is applied with a kernel of size K with padding of size P and Stride of S on an input image of size N. The dimension of the generated image will be $(N - K + 2P)/(S + 1)$

The padding size for output dimension to be equal to the input dimension is called zero-padding. Padding size for zero-padding with S=1 is $P = (K - 1)/2$

Pooling layer is another important concept of the CNN Pipeline. It is a nonlinear function used for downsampling the feature map. There are two common types of pooling-max pooling, min pooling, and average pooling. Among them, max pooling is the most common. It divides the data into a small fixed sized non-overlapping subregions. For each subregion, it outputs the maximum of all values in that area. The intuition is the rough apparent location of features of an image is more important than the exact location of features. That is why we can downsample the image with max poling without losing much information but giving a computational advantage for the next layers. It reduces the computational parameter for the next layer and also controls overfitting. Due to the aggressive compression nature pooling is conducted with small kernel size. The most common kernel is 2×2.

Fig. 3 shows the layers of the network and the output feature dimension of that layer. The input is compressed into small feature map by convolution and pooling layer which is the input to a fully connected layer. One final classifier layer is added with a softmax activation to generate the probability distribution of each output class.

IV. Experiment

A. Data

In our experiments, we have used 'Saarbruecken Voice Database' dataset. It's a collection voice recording of more than 2000 persons. Every session of recording includes the recording of the vowels [i, a, u] which was produced at a normal, high and low pitch, vowels [i, a,u] with rising falling pitch.

B. Feature Extraction Using Mel-frequency Cepstrum Coefficients

MFCC is a very common technique for obtaining feature vectors from a sound. MCCs The cepstral presentation is the inverse of the Fourier transform of log magnitude of Fourier transform. SpeechCepstral provides properties of the signal. Each tone of the sound is measured with frequency t in Hz, thus, an inward pitch is obtained on a scale is called Mel Scale. It is the frequency less than 1000Hz and more than 1KHz on the logarithmic scale. are commonly derived as follows:

Signal Strengthening

In this step, we pass a signal through a filter so that it highlights higher frequencies. It increases the energy of signal and leaves the low-frequency components in their original state. This is computed as

$$\phi = e^{(-2\pi F \Delta t)} \tag{9}$$

TABLE I
RESULT ANALYSIS ON ALL EXPERIMENT MODELS.

Model	True Positive	True Negative	False Positive	False Negative	Precision	Recall	Accuracy
GRU	105	123	46	38	73.42	69.53	73.07
LSTM	106	123	46	37	74.12	69.73	73.39
CNN (2-layers)	108	125	44	35	75.52	71.05	74.67
CNN (3-layers)	109	127	42	34	76.22	72.18	75.64

Here t represents the sampling period of our input sound.

Frame Fragmentation

This process involves segmentation of input samples which were converted from analog to digital data into small frames of length 20 to 40ms. Signals are divided into N fragment and adjacent frames are separated by M. Usually, the values of N and M are 100 and 256 respectively.

Windowing

In this step, the Hamming window is used as the next block that integrates the nearest frequency range. This window is used because it resists the signal not to be zero at the edges of window [15].

Lets say the window is Hw(n),0nN-1, where N=sample amount in each frame and equation is defined as follows:

$$y(n) = x(n) * Hw(n) \tag{10}$$

Fourier Transform

In this step we apply Fast Fourier Transform on the three component of equation (2) as x(w), H(w), and y(w) respectively.

$$y(w) = FFT(h(t) * X(t)) - H(w) * X(w) \tag{11}$$

Mel Filter Bank Processing

Since the extent of the output from the previous step is very wide, we apply Mel Filter to scale it. Triangular filters used to compute the weighted sum that is filter generated.

$$F(mel) = [2595 * log10(1 + f)700] \tag{12}$$

Discrete Cosine Transform

This step converts previous step's Mel Spectrum into spatial domain by DCT(discrete cosine transformation). The result is called MFFCs and a particular set of it is called features or attributes, used as the input of our models.

C. Train

We used mfcc to generate a matrix of size (20×50) from the voice data. For the experiment with RNN, we used one LSTM block followed by a classification layer. We chose the hidden size of 128 for LSTM hidden state and used the output of the last layer only as the input to the next layer. The output of the dense layer is 2.

For 2 layer CNN we used two convolution layer followed by two max pooling layer. For 3 layer CNN we have used one more convolution layer. The kernel size of the convolution layer was (2×2) with a ReLU activation function. We used two fully connected layers on top of convolution layer and a Dropout [16] layer in between them to prevent overfitting. The output of the first layer is 128 and the second layer is 2.

D. Result

We trained the model highly compressed data generated from the voice data. We trained four models and a few experimental settings of this. There is two RNN model with GRU and LSTM block, two are CNN with 2 and 3 CNN layers. The best-found accuracy of both models is given table II.

TABLE II
ACCURACIES ON DIFFERENT EXPERIMENTAL MODELS.

Model	Accuracy
GRU	73.07%
LSTM	73.39%
CNN (2-layers)	74.67%
CNN (3-layers)	75.64%

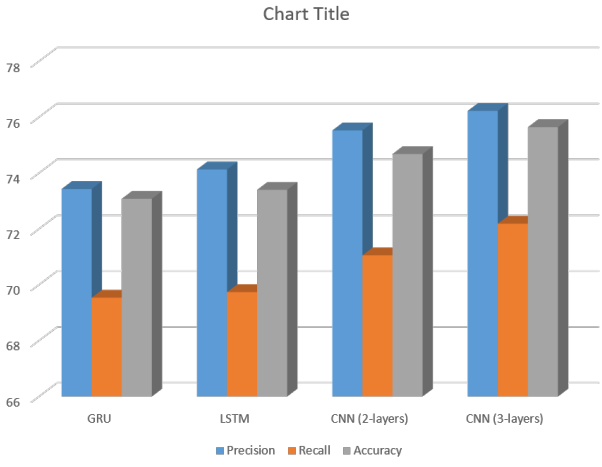


Fig. 4. Precision, Recall and Accuracy of different models used in experiment.

We can see from the data that the CNN gives better accuracy than the LSTM although the LSTM was the first candidate for sequential data. To further analyze the result we have reported the performance in precision and recall. The precision is formulated as $tp/(tp + fp)$ and the recall is calculated as $tp/(tp + fn)$. Table I shows the calculated precision, recall, and accuracy of all models in the experiment.

Fig 4 shows a visual summary.

We conducted multiple experiments of CNN model to find the best parameter for training the model for this task. We also trained on different feature vector size. We got a good result with reasonable training time at vector length 50. Fig. 5 shows the accuracy convergence of different training settings.

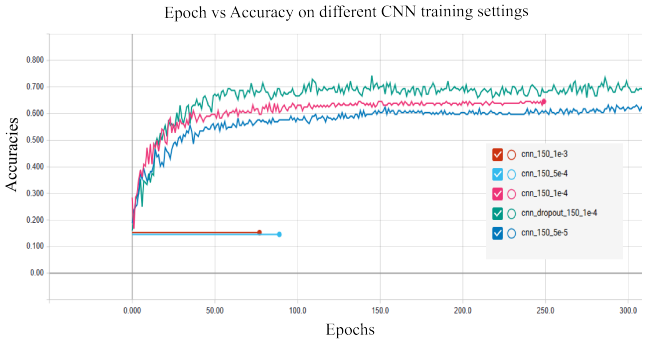


Fig. 5. Accuracy of different CNN model training approach. The first word of the experiment nace is the model name, in one case with extra layer name. The second word is the length of each sequence. The last word is the learning rate.

V. Conclusion

In this work, we build a classifier on voice dataset of pathology detection to illustrate the accuracy of different approach of possible models. Although the recurrent neural network is appropriate for sequence modeling we have shown that CNN can be a possible candidate for such a task and sometimes does better than LSTM if the sequence length gets longer. We have compared the accuracy of those models and also compared different hyperparameter settings to find the best parameters for the future experiment of relevant tasks.

References

[1] R. J. Baken and R. F. Orlikoff, *Clinical measurement of speech and voice*. Cengage Learning, 2000.

[2] M. S. De Bodt, F. L. Wuyts, P. H. Van de Heyning, and C. Croux, "Test-retest study of the grbas scale: influence of experience and professional background on perceptual rating of voice quality," *Journal of Voice*, vol. 11, no. 1, pp. 74–80, 1997.

[3] T. Nawka, L. C. Anders, and J. Wendler, "Die auditive beurteilung heiserer stimmen nach dem rbh-system," *Sprache Stimme Gehör*, vol. 18, no. 3, pp. 130–133, 1994.

[4] M. Markaki and Y. Stylianou, "Using modulation spectra for voice pathology detection and classification," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*. IEEE, 2009, pp. 2514–2517.

[5] A. Zulfiqar, A. Muhammad, and M. E. AM, "A speaker identification system using mfcc features with vq technique," in *2009 Third International Symposium on Intelligent Information Technology Application*. IEEE, 2009, pp. 115–118.

[6] T. Ganchev, N. Fakotakis, and G. Kokkinakis, "Comparative evaluation of various mfcc implementations on the speaker verification task," in *Proceedings of the SPECOM*, vol. 1, 2005, pp. 191–194.

[7] P. Henríquez, J. B. Alonso, M. A. Ferrer, C. M. Travieso, J. I. Godino-Llorente, and F. Díaz-de María, "Characterization of healthy and pathological voice through measures based on nonlinear dynamics," *IEEE Transactions on audio, speech, and language processing*, vol. 17, no. 6, pp. 1186–1195, 2009.

[8] Z. Milutinović, "Classification of voice pathology," *Folia phoniatrica et logopaedica*, vol. 48, no. 6, pp. 301–308, 1996.

[9] A. Al-Nasheri, G. Muhammad, M. Alsulaiman, Z. Ali, K. H. Malki, T. A. Mesallam, and M. F. Ibrahim, "Voice pathology detection and classification using auto-correlation and entropy features in different frequency regions," *IEEE Access*, vol. 6, pp. 6961–6974, 2018.

[10] A. A. Dibazar and S. Narayanan, "A system for automatic detection of pathological speech," in *Conference Signals, Systems, and Computers, Asilomar, CA*, 2002.

[11] J. I. Godino-Llorente, P. Gomez-Vilda, and M. Blanco-Velasco, "Dimensionality reduction of a pathological voice quality assessment system based on gaussian mixture models and short-term cepstral parameters," *IEEE transactions on biomedical engineering*, vol. 53, no. 10, pp. 1943–1953, 2006.

[12] P. T. Hosseini, F. Almasganj, T. Emami, R. Behroozmand, S. Gharibzade, and F. Torabinezhad, "Local discriminant wavelet packet basis for voice pathology classification," in *Bioinformatics and Biomedical Engineering, 2008. ICBBE 2008. The 2nd International Conference on*. IEEE, 2008, pp. 2052–2055.

[13] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[14] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[15] D. Instruments, "Waveform analysis using the fourier transform," 1995.

[16] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

An Efficient Method of Lane Detection and Tracking for Highway Safety

Bubly Barua
Department of Computer Science and Engineering
Chittagong University of Engineering and Technology
Chattogram- 4349, Bangladesh
bublycse@gmail.com

Shuva Biswas
Department of Computer Science and Engineering
University of Science and Technology Chittagong
Chattogram- 4202, Bangladesh
shuvabiswas12@gmail.com

Kaushik Deb
Department of Computer Science and Engineering
Chittagong University of Engineering and Technology
Chattogram- 4349, Bangladesh
debkaushik99@gmail.com

***Abstract*— Lane detection and tracking in highway road is one of the most crucial tasks for advanced driver assistance systems (ADAS). Advanced driver assistance systems are being developed to assist driver in driving process reducing road accidents. In this paper we present efficient lane detection and tracking method in highway roads. To do this, the first step is camera calibration used to remove the effect of lens distortion. Next canny edge detection algorithm finds the edges of the images. Then an adaptive region of interest (ROI) is selected. The ROI is actually based on the triangular shape appeared at the bottom of the image. ROI removes the unwanted region. After that possible lane markings are identified using Hough transformation analysing lane candidate pixels. Once the lane candidate pixels along with the lane are found, these pixels are continuously scanned to obtain the best fit using linear regression analysis. The system has been investigated under various illuminations in highway roads for detecting lane marking. The experiment shows that the proposed method can detect the lane quickly and accurately.**

Keywords — ROI segmentation, lane detection, Hough transformation, advanced driver assistance system

I. Introduction

Advanced driver assistance systems (ADAS) are being developed to assist driver in their driving process reducing road accidents. Car safety and road safety are thereby now-a-days a most important issue. Human errors are prone occurring most of the road accidents in highway roads. Drowsiness of driver is one of the main factors for road accidents. ADAS provides the automated system to automate, enhance and adapt with the environment to reduce road fatalities and avoid the collisions. The system either alerts the driver or take necessary actions of driving tasks [1] providing necessary information like lane departure warning (LDW), position of vehicles, road boundary. One of the scientific field involving in these tasks include computer vision that perform tasks automatically compared with human visual system task. Computer vision senses the environment and can be applied in many applications supported by intelligent vehicle systems (ITS). Autonomous vehicle is a newer arena of computer vision ranging its level from fully autonomous vehicle to vehicles where a driver is supported by a computer vision based systems. Road boundaries, texture and lane markings are the most perceptual cues for human driving. Autonomous systems are also likely to be expected to rely on the perceptual cues that human do. Computer vision is pioneered with artificial intelligence and digital image processing that extract 3D image for images achieving full scene understating. This paper presents an efficient lane tracking system which is capable of providing a real time performance in finding lane markings in highway road based on region of interest along with the triangular shape in various illuminations e.g. day, night and sunny. Finally this detected line is best fitted with curve fitting where the left and right lane boundaries smoothly progress.

The structure of the paper is organized as follows section I provides introductions, section II describes literature review, section III contains the proposed method, section IV presents results and discussions and in section V conclusion is discussed briefly.

II. Literature Review

Various approaches have been discussed for lane detection system. Understanding of lane marking includes detecting the structure of road, the position and number of lanes, splitting, merging and ending lanes and roads in highway road. Much progress has been made possible in recent years. Various sensing modalities like stereo monocular vision, LIDAR, GPS, has been used for better improvement in understanding lanes [2]. A top-view transformation model has been proposed for a vehicle parking assistant [3]. Here perspective projection image is transformed into corresponding bird's eye view followed by a fitting parameters search algorithm. A region of interest (ROI) segmentation technique has been applied for intelligent highway safety [4]. In this algorithm a top-hat transformation is being created as a preprocessing step for enhancing contrast of the image. Then Hough transformation has been used for lane departure decision. A vision based lane departure detection technique has been developed [5] for urban roads. This technique uses Sobel operator. Then straight lanes are detected based on Hough transformation estimating the lane

width. A horizontal line detection technique has been presented for detecting horizontal line under which the lanes are present [6]. This method uses a vertical mean distribution. Then canny edge detector is applied marking edges of lanes and noise candidate. Finally K-means clustering and RANSAC algorithm has been applied for detecting lanes. Edges based lane detection method has been analyzed step by step [7]. Few known and common edge detection algorithms like Sobel, Canny, Roberts and Pewit were analyzed and evaluated using image segmentation, morphology, image statistic and Hough transformation. A layered approach has been provided for lane detection at night [8]. The bottom part as ROI is set here eliminating the sky and other irrelevant objects. In further step the color channel is averaged to obtain gray scale. The dashed lines are elongated using temporal blurring. Then bright objects are extracted by threshold. Low resolution Hough transform is applied on the resulting binary image consisting of left and right halves. Finally a matched filter is iterated to extract lane markers. Three modules-lane model generation, parameter estimation and matching has been developed for lane detection based on geometrical and Gabor filter [9]. Lane gradient maximized image has been created from a color image for lane detection in various illuminations [10].

III. METHODOLOGY

This section illustrates the lane detection and tracking algorithm in highway road in various illuminations. The steps of our proposed method are drawn in "Fig. 1". The system consists of few subsystems: camera calibration, noise removal, Gaussian blurring, canny edge detection, ROI selection, Hough transformation and finally curve fitting.

A. Capturing Image

The first step involves capturing image. The input data is a color image sequence captured by a CCD camera. The camera is mounted just at the front view mirror along the central line of the road inside the vehicle. It captures image from the front view of vehicle including road curvature, trees, buildings, shadows, vehicles and so many objects which do not come under consideration. The real time computer system attached with image capturing card read the images in real time and store them in memory. After that the proposed system read the image sequences stored in memory and then start the further processing steps. Few captured images from highway road during day time, sunny and night depicted in "Fig. 2".

B. Camera Calibration

Camera calibration may be required to calibrate the captured image due to its distortion. Geometric camera calibration also referred to as camera resectioning, measures the parameters of camera lens and image sensor of the captured image. These parameters are used to correct because of lens distortion and measure the object's size in world unit as well as for determining the location of the camera in real application. To estimate camera parameters like intrinsic, extrinsic and distortion coefficients, 3D world coordinate system points and their corresponding 2D image coordinate system are measured. To do this, a checkerboard is a best example as a multiple image of calibration pattern. Calibration

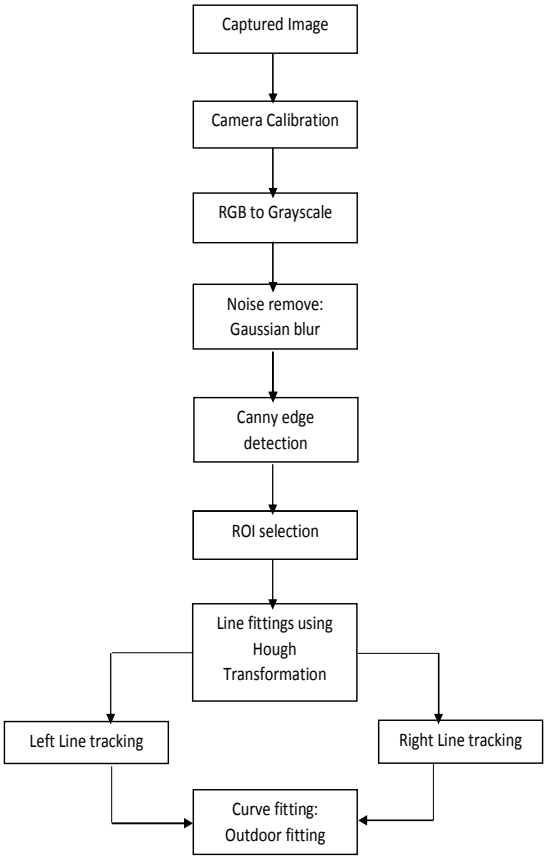
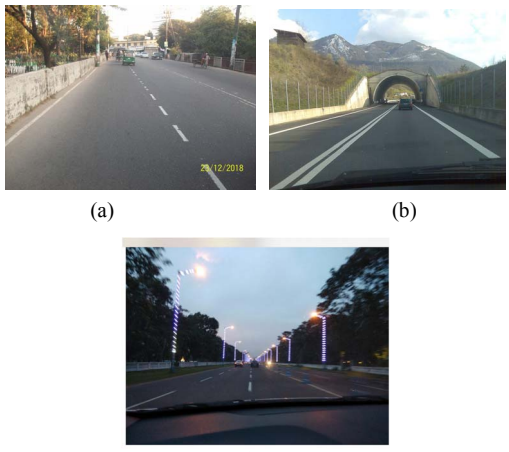


Fig. 1 The proposed framework for lane detection.

algorithm is used to calculate the camera matrix using the parameters of extrinsic and intrinsic shown in Eq. (1).



(a) (b)

(c)

Fig. 2. Road scene in different illuminations a) The captured image from highway road at day light b) at sunny c) at night view from highway road.

$$W\,[x\ y\ 1] = [X\ Y\ Z\ 1]\,P \quad (1)$$

Where W is scale factor, [x y 1] represent image points, [X Y Z 1] is world points and P is camera matrix shown in Eq. (2).

$$P = \begin{bmatrix} R \\ t \end{bmatrix} k \quad (2)$$

The optical centre of the camera points the origin of the camera's coordinate system that comprises the rotation R and translation t. However the intrinsic matrix K is calculated based on its focal length, principal length also known as optical centre and skew coefficient.

C. RGB to Grayscale conversion

Segmentation of lane boundary from the road surface using the color information becomes difficult for edge detection. The gradient of the road surface and lane markings change invariantly due to its shadows, buildings, trees and invariant illumination. Images are made of pixels and each pixel in the image is a combination of three intensities. Therefore images are converted into grayscale meaning that three channel color image consisting of red, green and blue channel-24 bit is now converted to one channel gray 8 bit. The point of using a grayscale image processing is a single channel is faster than processing a three channel color image and less computational intensive. The basic form of the conversion is defined in Eq. (3).

$$Y = 0.299R + 0.587G + 0.114B \quad (3)$$

Y, the gray scale representation of the color image is the weighted sum for the red, blue and green component value. "Fig 3" shows effective result of conversion.

D. Noise Removal

Noise is a real world problem and almost always undesirable in image processing. The goal of image filtering is to enhance the image contrast and to recover the image being degraded. Noise removal is a prerequisite step for edge detection and thus smoothen the image. When detecting edges it is important to accurately catch as many edges in the image as possible. Image noise must be filtered out as it creates false edges and ultimately it affects edge detection. A typical filter-Gaussian filter being used to filtering out noise of image. Grayscale image is represented by a single number that describes the brightness of the pixel. Basically the value of a pixel is modified by average value of the pixel intensities around it. Averaging out the pixels in the image to reduce noise is done by the kernel of normally distributed numbers and finally each pixel value equal to the weighted value average of its neighboring pixel is set. The probability density function of a Gaussian random variable is given by Eq. (4).

$$P(z) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(z-\bar{z})^2}{2\sigma^2}} \quad (4)$$

E. Edge Detection

Edges are the points where the gray scale values change sharply. Edge detection is a very important step in determining the location of lane boundaries as the edge is one of the major information contributors to the lane. The main factor of edge detection is the detection of the exact edges along with the good orientation of the lane boundary in the image. Because the memory required to store the edges of the image is less compared to the whole image view. However it contains all information of the lane boundary and also orientation of the lane. In our system we applied Canny edge detection, a standard and popular edge detection algorithm. The Canny edge detector works in a multi stage process. The method is based on the characteristics of intensity values of considered pixel. A signal to noise ratio (SNR) on the edge detector may be applied so that edges could be found even the data quality is poor. Gradient is the change in brightness or discontinuity over a series of pixels. So, Canny function computes the gradient in all directions of blurred image and then trace the strongest gradients as a series of white pixels as shown in "Fig. 3". Two arguments low threshold and high threshold allow us to isolate the adjacent pixel that follows the strongest gradient. The basic form of edge detection using canny is defined in Eq. (5) and (6). Edge Gradient,

$$G = \sqrt{Gx^2 + Gy^2} \quad (5)$$

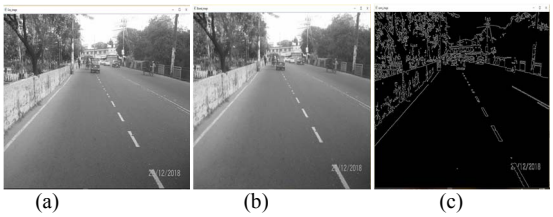
$$\theta = tan^{-1}\, Gy/Gx \quad (6)$$

The threshold value finds the optimum values for the lower and upper threshold in potential data analysis is defined in Eq. (7).

$$T_1 = \min_{\substack{X=0,1,\ldots\ldots,m \\ Y=0,1,\ldots\ldots,n}}\{f(x,y)\} \text{ and } T_2 = \max_{\substack{X=0,1,\ldots\ldots,m \\ Y=0,1,\ldots\ldots,n}}\{f(x,y)\} \quad (7)$$

F. Region of Interest

The region of interest is basically a polygonal selection from 2D image. ROI defines the border of the lanes that comprises individual points of interest. Preprocessing examples of selecting region of interest using matplot are shown in "Fig.4".



(a) (b) (c)

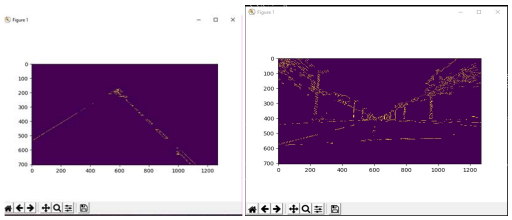
Fig. 3. Processing example of detecting edges a) Gray image b) Blurred image (c) Canny image.

In this step lane lines are detected using ROI that excludes the other irrelevant objects. There might be several objects in the image that are not of interest like trees, buildings, human

etc. A ROI is isolated based on the region of interest that traces a triangle limiting the extent view. Ultimately this triangle isolates the lane lines which are shown in "Fig. 5" and "Fig. 6". The enclosed of the triangle shape is completely filled white as the mask image. Here the canny image is masked with the mask image to trace the ROI triangle shape by bitwise –and operator. Thereby masking the entire region, the operation occurs element wise, so all the white pixels in the region will be operated against the corresponding region of canny image. So, by bitwise and operation which means successfully masked our canny image to ultimately show the ROI the region traced by the polygon contour.

G. Line Finding Using Standard Hough Transform

This step involves using Hough Transform technique to detect straight lines in our region of interest and thus identify the lane lines. This technique uses a parametric space called Hough space. Hough transform is a mechanism that gives more weightage to pixels that are already in line by a voting procedure. This voting procedure is carried out in a parameter space from which lines are obtained as local maxima. The lines that are outside the space are rejected. For example any horizontal line that is not part of the lane boundary can be neglected as well. Hough transform halves the image splitted in yielding left and right side of the image and limits the



(a) (b)

Fig. 4. Preprocessing examples of selecting region of interest (a) Matplot image for selecting mask at day light, (b) Matplot image used for selecting mask at night view.

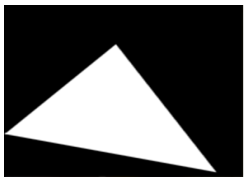


Fig. 5. Processing example of extaracted mask calculated from mat plot for ROI selection.



Fig. 6. Processing example of extracting region of interest (ROI) selection based on mask of fig.11 at day, sunny and night.

search space to 45^0 for each side. Every line has two quantities its slope m and intersect c. With these two numbers, a line can be described completely. Every line in the xy space is equivalent to a single point in mc space shown in "Fig 7". However Hough transform is to convert points in the xy space to line in mc space. For canny image, every point that is nonblack are converted into lines in mc space. Obviously some lines intersect shown in "Fig.8". And these intersections are the parameters (m, c) of the line. Hough transform is a technique that can be used to record vote for each point lies i.e. to look for lines that get many votes. These points of intersections are inside a bin of Hough space for every point of intersection there are to cast of votes inside the bin that it belongs to bin with the maximum number of votes as shown in "Fig. 9". And this is how the line that can be drawn .Polar coordinate system in replace of Cartesian coordinate system using rho and theta is expressed in "Fig.10". Whatever the values of theta, the line is drawn since it was voted as the line of best fit scribing the data based on the threshold. Threshold is the minimum number of intersections to detect a line previously mentioned in a series of points. Hough space represents the theta and rho values of lines that are common between a series of points. For every point of intersection of vote inside the bin it belongs to a line with some value of rho and theta. Finally lines are detected from a series of points in the gradient image using Hough transform and these lines are shown in the original color image by blending the lines with the original image as shown in "Fig 11".

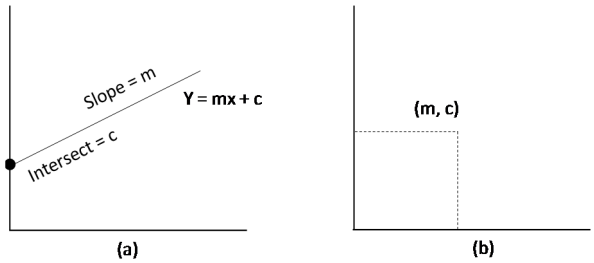


Fig. 7. a) xy space b) mc space

H. Finding Lane Line

Multiple lines with different slope and length are formed using Hough transform. The ultimate solution is to have only two lines, one for right and other for left lane .And it is important that the lines depending on either side are all roughly more or less going to the same direction. Left line has a negative slope and right one has a positive slope. Averaging out the slope and y intercept of multiple lines into a single line traces out both of the right and left lane. Left fit contains the averaged lines on the left and intuitively right fits containing coordinate of the line at right. So both the lines are going to have the same vertical coordinates are going to start at the bottom and go upward. Hough lines vary in length, so the most important and appropriate lines are those which are longest in length. Any detected line traced by less than minimum line length will be rejected. Since shorter lines may be due to the noise or other random calculations.

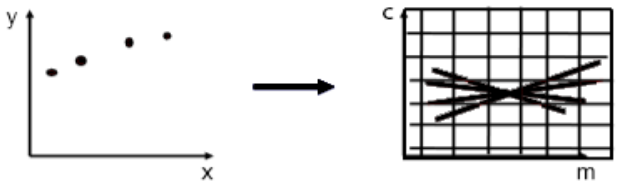


Fig. 8. Finding lines in an image using Hough Transform -each edge point in image space vote for a set of possible parameters in Hough space

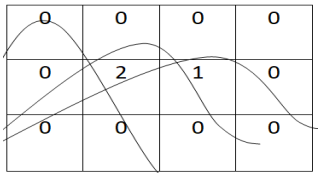
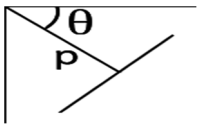


Fig 9. Processing example of Hough Transform voting system obtained as local maxima.



$$p = \cos\theta + \sin\theta$$

Fig. 10. Polar coordinate system

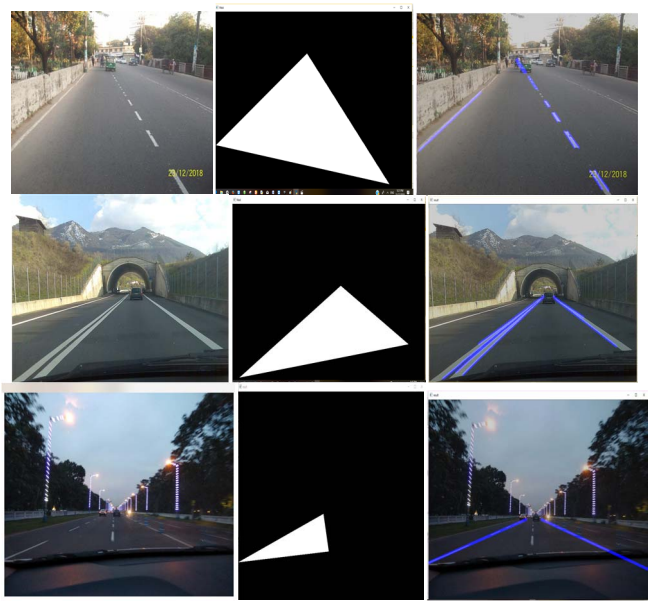
I. Curve Fitting

For curve fitting least square method is used to fit the data points. This phase uses the two vectors of data points from the lane scan as input. For simplifying the data points (x_1, y_1), (x_2, y_2), ... (xn, yn) in lane boundary are observed. The function f(x) = ax + b is the formula commonly used to fit the data. The coefficients a and b in the formula f(x) are the parameters that are adjusted. However, there may be inevitable measurement errors in the observation. The best try is to choose the parameter of the function so to minimize the fitting error-the distance between the data values yi and those given by the fitted curve at the x-values where the data was originally collected ri = yi - f(x) for i = 1,2,.....n. For optimization the least square fitting the parameters of f are chosen to minimize mean square error. “Fig. 12” shows the processing examples of curve fitting for lane tracking straight and dashed lane detection in highway.

IV. Results And Discussions

The experimental results of some lane detection images are explained in this section. All the experiments were executed on the intel(R) Core(TM) i5-4200U CPU@1.60 GHz 2.30 GHz processor with 4 GB RAM. All our proposed systems were implemented using python 3.7 and OpenCV library. The system was built up on recorded RGB images from a digital camera in 2560x1920 resolutions. For processing the resolution were resized in 704x1269 formats.

The experiment is performed on moving vehicles shooting in highway roads under various illuminations with different distances ranging from 10-60 meters. The camera height from the ground should be a certain value that is not more than two meters. The experimental results of lane mark detection and the time complexity of the proposed system is summarized in Table I. As observed from the table the detection rate for the tested image is 98.6 %.The captured images from different distances and the detected resultant images are shown in “Fig. 13”. The blue lines in the images show the result of final processing of lane detection. As observed from the figures that the lanes can be easily detected in different distances and different illuminations with so many noises. The results reveal that the proposed system is not affected by the environment of the streets; have the ROI selected suitable for any set of frames recorded by the camera. However, some problems are not solved yet such as sharp curves in the foreground of the image and the accurate detection of the lanes under the heavy rain. The overall performance of the proposed system is performed qualitatively and quantitatively. This evaluation ranges in highway during day, sunny and night few frames had been shown with detected lane during various illuminations.



(a) (b) (c)

Fig. 11. Processing example for extracting lane line (a) input image at day, sunny and night (b) their corresponding mask based on polygon shape (c) detected lane output image .

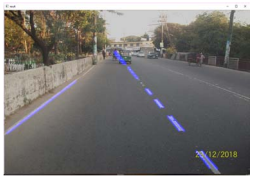


Fig. 12. Processing example of curve fitting for lane tracking straight and dashed lane detection in highway



(a)

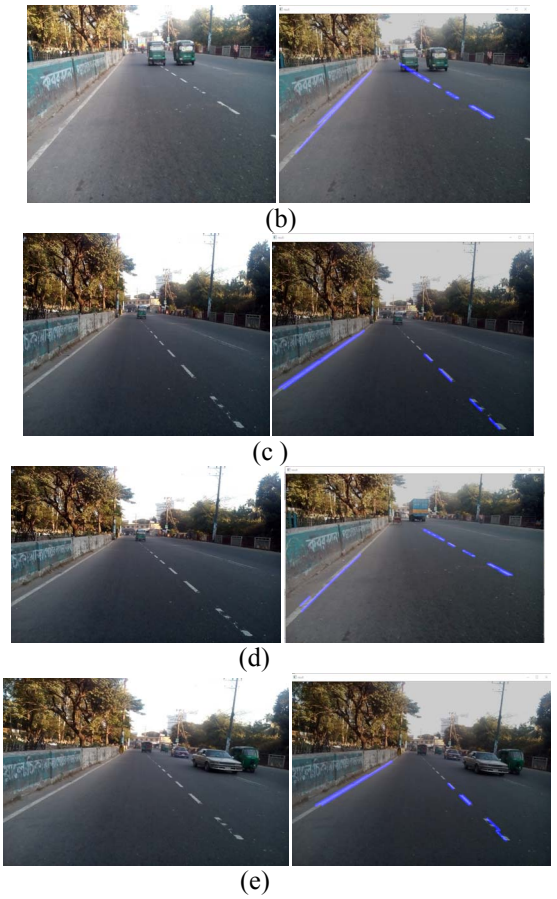


Fig. 13. The lane image with vehicles from distance (a) -(e) 0m,20m,30m,40m,60m respectively and Final resultant image

Table I.
EXPERIMENTAL RESULTS OF THE PROPOSED ALGORITHM

Input Images	Distance in meter	Time complexity	Detection Rate
a	10	0.23	100
b	20	0.28	100
c	30	0.35	98
d	40	0.43	98
e	60	0.48	97

V. CONCLUSION

In this review, we have presented a frame work which is robust for lane detection and tracking in highway road. This presents a detailed analysis of the use of lane position in various environments with different illuminations. This led to our presentation of the ROI based lane-tracking system operated in a wide range of environments. For safe lane systems, this approach developed a safety system for avoiding lane departure for a large and complex set of traffic scenarios. Besides the application of driver assistance, the Google driverless car can also be included along with lane keeping support system for future work [11]. Google's fleet of robotic Toyota cruises which drives in busy highway roads, city traffic and mountains roads, steering by itself with no driver. According to our experiment result of the proposed method in this paper, we can conclude that this system can be used under most of highway road in daylight, night time, sunny and so on. Also it can be used under various – straight lines as well as with dash lines, lane markings and vehicles for lane boundary recognition. Based on ROI and Hough transform voting system, the system can recognize lane boundary by means of computer vision and forward collision functions. Gray scale statistic, ROI, Hough transform-voting based approaches are applied to recognize the lane boundaries. The system provides high availability, reliability and accuracy. In addition to this, computer vision system supports full installation conveniently. To simplify the installation step, the systems are almost adaptive to most of the vehicles.

REFERENCES

[1] Ammu M Kumar and Philomina Simon ," Review Of Lane Detection And Tracking Algorithms In Advanced Driver Assistance System", International Journal of Computer Science & Information Technology (IJCSIT) Vol 7, No 4, August 2015.

[2] Aharon Bar Hillel · Ronen Lerner · Dan Levi · Guy Raz," Recent progress in road and lane detection: a survey", Machine Vision and Applications (2014) 25:727–745 DOI 10.1007/s00138-011-0404-2, Received: 28 October 2010 / Revised: 12 December 2011 / Accepted: 19 December 2011 / Published online: 7 February 2012 © Springer-Verlag 2012.

[3] Chian-Chuan Lin, and Ming-Shi Wang,"A vision Based Top-View Transformation Model for a Vehicle Parking Assistant",Sensors 2012,12,4431-4446,doi:10.3390/s120404431,ISSN 1424-8220,www.mdpi.om/journal/sensors,30 March 2012.

[4] Prof. Sachin Sharma and Dr. D. J. Shah, "A Much Advanced and Efficient Lane Detection Algorithm for Intelligent Highway Safety", Sundarapandian et al. (Eds) : ITCS, SIP, CS & IT 09, pp. 51–59, 2013. © CS & IT-CSCP 2013, DOI : 10.5121/csit.2013.3106.

[5] Y.-C. Leng and C.-L. Chen, "Vision-based lane departure detection system in urban traffic scenes," in Control Automation Robotics Vision (ICARCV), 2010 11th International Conference on , 2010, pp. 1875-1880.

[6] T.T. Tran, H.M.Cho, S.B.Cho, "A robust method for detecting lane boundary in challenging scenes." Information Technology Journal. Image Process., vol. 10, no.12, pp. 2300-2307, 2011.

[7] Moveh Samuel, Maziah Mohamad, Shaharil Mad Saad, Mohamed Hussein," Development of Edge-Based Lane Detection Algorithm Using Image Processing", International Journal On Informatics Visualization ,Vol 2 (2018) No 1 e-ISSN : 2549-9904 ISSN : 2549-9610.

[8] A.Borkar, M. Hayes, M. Smith, and S. Pankanti, "A layered approach to robust lane detection at night," in Proc. IEEE Workshop Comput. Intell. Vehicles Veh. Syst., 2009, pp. 51-57.

[9] S.Zhou, Y.Jiang, J. Xi, J. Gong, G. Xiong, and H. Chen, "A novel lane detection based on geometrical model and gabor filter," in Intelligent Vehicles Symposium (IV), 2010 IEEE , June 2010, pp. 59-64. www.powerdivision.gov.bd/user/brec1/30/1, accessed on August 29, (2015).

[10] H.Yoo, U.Yang, and K. Sohn, "Gradient-enhancing conversion for illumination-robust lane detection," IEEE Transactions on Intelligent Transportation Systems , vol. 14, pp. 1083 -1094, September 2013.

[11] Shweta N. Dethe, Varsha S. Shevatkar, Prof. R. P. Bijwe," Google driverless car",IJSRSET | Volume 2 | Issue 2 | Print ISSN : 2395-1990 | Online ISSN : 2394-4099,2016.

Bangla Sign Language Conversation Interpreter Using Image Processing

Prosenjit Roy
Daffodil International Universitym(CSE)
antor.roy.roy@gmail.com

S.M.Miraj Uddin
American International University Bangladesh(EEE)
uddinsmmiraj@gmail.com

Md. Arifur Rahman
Daffodil International Universitym(CSE)
coder.roby@gmail.com

Md. Musfiqur Rahman
Daffodil International Universitym(CSE)
musfiqforhad@gmail.com

Mohammad Shahin Alam
American International University-Bangladesh(EEE)
the.shahinalam@gmail.com

Md.Saidur Rashid Mahin
American International University-Bangladesh(EEE)
mahin770 @gmail.com

Abstract—This project about sign language conversation interpreter system. It is made for speak disable people. Here a man who don't know the sign language can make conversation with a speak disable people. In this project we used efficient methods to convert sign language into text. We used our own dataset and collected dataset of Bangla Sign Languages using hand gestures. Inputs will be taken by webcam and recognized by neural network system. System will show text output by using gestures. We have created a CNN which is similar model of the MNIST classifying model. It can use both Tensorflow and Keras. This model was trained using Keras by video streaming. In a video stream sign language, interpreter can turn it fast into text. We used histogram back projection method to recognize people's hand as object. For better result we calculated the grey level of hands. *Index Terms—Video streaming, CNN model, Keras, Tensorflow, OpenCV, Histogram backprojection.*

I. INTRODUCTION

In this modern era, we have advance technology to utilize human ability more and more. But some people are physically disable by birth or for unexpected accidents. In this world about 15% people are speak disable (according to WHO reports). It's about 8% of the Bangladesh's population. They have to lead their life as burden in society. This burden word is applied because they can't make conversation with Normal people and that's why they can't gain skills of any work. Basically, people who are hear disable and speak disable, they need to learn sign language for communication. At this matter, for communication, sign language is the preferred form. But most of the common people do not understand sign languages. For making communication between the people who use sign language and the people who do not understand sign language, sign language interpreter's service can be arranged. In this way of communication, it is important to know how to make use of the services of the system effectively. Every country has their own sign languages. Our country has also well-formed sign language. Many researchers are working on sign language recognition systems. In our country most researchers worked on Bangla sign language, a system developed by Pavel et al. [1] analyzes video clips of different gestures of sign languages taken as input and gives audio output.

II. GOALS AND OBJECTIVE

We have a huge population who are speaking and hearing impaired. In our development country they a very little chance to make them self-dependent as they can't communicate with others. If they can communicate, they can do all works. They have to face under humiliation situation in every step of life. Now-a day's people mostly depend on technology. We want to speak disable and hearing disable people can lead a modern life through this advanced technology. Our main goal, they can communicate all common people and take parts in all of work of our life. We want to make this people our assets not burden in our society.

III. LITERATURE REVIEW

In 17th century western countries sign language developing was started for a visual language where the conventional gesture, hand signs was, finger spelling also includes the positions to represent a meaningful line. In Europe, Laurent Clerc first taught French sign language. Then Thomas Hopkins Gallaudet brought Clerc to America. His plan was to start American school for deaf. After that from all over the world children came to attend this school and brought their home sign language. Before 19th century whole sign language system was confined only by fixed words using finger spelling system [13]. They probably figured and created signs to understand and speak with deaf people. Khedkkar Safaya and Prof. (Dr.). J.W.Bakal of Mumbai researched on Real time based hand gestures in 2013. They used a dynamic vision censor which can recognize bare hand gesture. They figured out the delivery point track number for every frame and compared with given threshold values. They took this point to detect frame and find the hand position. They calculated the wrist point in the extraction phase used their own algorithm to define the hand sign. In our country our Bengali sign language is formalized in 2003 by CDC (Center for Disability in Development).

Bengali language is the world's 5th language in terms of population. There are 47 alphabets in Bengali language where is 11 vowels called "SOROBORNO" and 36 constants called "BENJONBORNO". But it's true that in conversation people don't use alphabets. That's why sign words should interpreted. Bikash Chandra Karmokar[14] and his team of Technology did a research employing Neural Network Ensemble. They took input from webcam. They also detected skin color of hand. After that captured image had been converted into its threshold value. In 30x33 scale pixels, images were converted by applying normalization process. Then feature extraction method has been applied and they used NCL algorithm to train these images.

IV. METHODOLOGY

Our main target is to capture image by our system and match it with our data set and show the outputs as Bangla language. For this work we take a lot of images of different people's hand sign in different specific gesture. Then we created 15 different gesture samples using OpenCV Library. We convert those captured images into 50x50 pixels and made our dataset. This system works in two steps. At first this system will take dataset images then it applies the skin detection algorithm and detect the skin and remove the background. We used Convolutional Neural Network (CNN) in this system and we created a CNN is looks a lot of similar to classifying model using both Tensorflow and Keras. If anyone want to add more gesture images in this project, they need to make their own layer and some parameters. We trained this model using Keras on videos stream. Which trained by 4 Bangla font, 5 numbers and 6 words. For run this project have to meet with some requirements. The instruments and requirements are given:

1. Install Python 3.x operating system

2. Tensorflow 1.5 using command

3. OpenCV 3.4

4. KerasAPI

5. CPU

For collecting data, Bangla sign language recognition is not still a widely researched topic so we did not find any dataset on any resources. So that, we made our own dataset. We captured images from hundred people bare hands for fifteen different signs and some examples are shown in Fig.1.

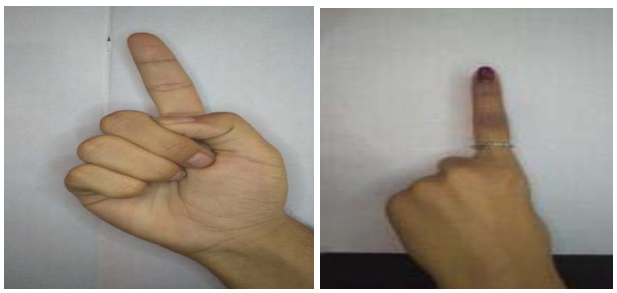


Fig 1: Image Conversion of ১



Fig 2: Data collection (Image Conversion of অ & আ)

But for our collecting image we have to apply skin detection algorithm which detect the skin color pixel from it and identify the blob and remove all binary objects. Then it converts the image to inverted image. So, the hand gestures turned into a black and white image. That is shown in Fig.2.

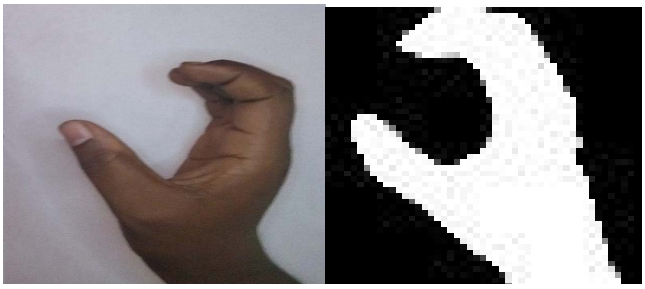


Fig 3: Image Conversion of গ

To create gesture, we used python create_gesture.py command. By using this command, we can create gesture instantly. These gestures can take instance image from us for defining specific signs. As in our country, skin color vary man to man we inputted our dataset in each gesture for better result. All these images are in grayscale which is stored in the gestures. The pictures were flipped by using flip_images.py. The script flips every image along the vertical axis.

For training our dataset we created a CNN model which is similar to MNIST Classifying model. For starting our model, we used some basic code of google tensorflow. We divided our model in three layers. First layer is for recognizing gestures. Second layer helps first layer to show the right output. Third layer is for decreasing processing time. we use two library of CNN model. One is Tensorflow and another is Keras. Keras can run over Tensorflow. So that, anyone can use either Tensorflow or Keras. Keras is a modified model of CNN. In tensorflow there is default command for training dataset. We trained our dataset in every gesture by running cnn_tf.py file. By using these we trained our dataset and we got good result to convert the sings to text which can make a conversation between deaf and numb in Bangla in future.

In this set_hand_hist.py file we write the code to set make histogram in our camera position for 50x50 pixel from where the camera only hand gesture, for making real time gesture to compare with our dataset. This histogram setting is very important in light cases. We need to set hand histogram in every time for light changes. Because it is light sensitive. Which is shown in Fig.4.

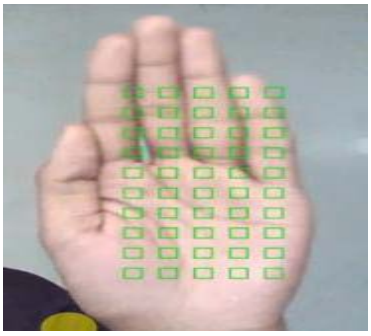


Fig 4: Hand set up

V. RESULTS AND DESCUSSION

In our proposed system to measure the performance we use fifteen types of signs and in future we want to increase the numbers of sign. We use 100 images from 15 datasets. Our first dataset contains 100 images and the average accuracy is 81%.

A. *Experimental Result*

Total class of images =15;
In one class number of images = 100;
Our final dataset contain 1500 images.

For testing our model, we make some gestures in front the camera to test our model. As we read a lot of papers and try some similarities projects, we understood that our model is working as well. Some of the results are given below

Table I: Resulting dataset

ক	
থাম	
১	
২	
৩	
৪	

Table II: Accuracy for train set classes

Class	Train set average accuracy for each class
আমি	82%
তুমি	79%
ভাল	80%
ভালোবাসি	83%
থাম	82%

Table III: Accuracy for train set classes

Class	Train set average accuracy for each class
১	81%
২	78%
৩	80%
৪	80%
৫	85%

Table IV: Accuracy for train set classes

Class	Train set average accuracy for each class
শুভকামনা	75%
ক	78%
খ	82%
গ	80%
ঘ	81%

B. *Get model report*

To get classification reports about the model we have test_images and test_levels file which have generated load_image.py. If we do not have them run load_images.py file again. Then we have to run below file.

Python get_model_report.py

We will get the confusion matrix, f scores, precision and recall for the prediction by the model.

Precision: precision is the piece of related instances among the retrieved instances. High precision means that an algorithm returned substantially more relevant results than irrelevant ones.

Table V: Accuracy measure based on performance matrix

Class	Recall %	False positive rate %	Specificity %	Precision %	F-measure %	Accuracy %
আমি	82	11	81	78	87	90
তুমি	81	9	78	83	90	91
ভাল	80	12	91	81	81	92
ভালোবাসি	78	11	79	84	78	87
বাম	79	10	93	82	78	86
১	82	7	82	86	88	83
২	82	9	82	82	79	86
৩	80	10	81	79	90	82
৪	83	15	89	81	87	83
৫	86	01	79	82	88	86
শুভকামনা	78	11	81	89	82	90
ক	79	12	82	88	88	84
খ	80	12	83	78	78	81
গ	82	0	81	81	78	91
ঘ	82	11	78	82	89	81
Average	81	9	83	82	84	86

Statically analysis of accuracy, F-measure, Precision, Specificity, false positive is shown in Fig.4.

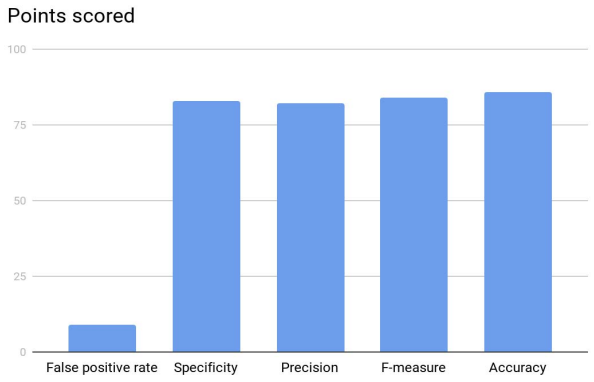


Fig 4: Statically analysis of accuracy, F-measure, Precision, Specificity, false positive

Overall, result is satisfying. Our project can show result over 80%. Our model and gestures are well formed and project can process result in high accuracy and with fast processing time. We need to set hand histogram for every

light change condition. Our project also lightly backgrounds sensitive. Background noise sometimes effect on results.

VI. CONCLUSION AND FUTURE PLAN

In Bangladesh about 4.8 milloin people who are disabled in hearing or speaking. This huge population can be our assest if they can make conversation with us. In this paper, we showed about 15 combined Bengali sign language where has Bengali sign language alphabets, some Bengali sign language Numeric numbers and individual signs too. Our main goal is interconnect thisspeaking and hearing imapired people with common people. Without this interconnection, We can not turn them into skillfull and self-dependent population. In the future, Our goal is to make this system more efficient. We collected our own data-set. That's why It took huge time to collect data and train data. In future our target is to collect and combine all Bengali sign language alphabets , all Bengali sign language numeric numbers and other individual Bengali sign languages. This is modern era. Most of the people in our country use smart phones. We will work to make a phone-based application of this system for making easy communication with all.

REFERENCES

[1] D. Shahriar Hossain Pavel, Tanvir Mustafiz, Asif Iqbal Sarkar, M. Rokonuzzaman, "Geometrical Model Based Hand Gesture Recognition for Interpreting Bengali Sign Language Using Computer Vision", ICCIT, 2003.

[2] Md Azher Uddin; Shayhan Ameen Chowdhury, ''Hand sign language recognition for Bangla alphabet using Support Vector Machine'',2016 International Conference on Innovations in Science, Engineering and Technology (ICISET),2016.

[3] Mahmood Jasim, Md. Hasanuzzaman, "Sign Language Interpretation using Linear Discriminant Analysis and Local Binary Patterns", 2014 International Conference on Informatics Electronics & Vision (lCIEV), pp. 1-5, 2014.

[4] RahatYasir, RiasatAzim Khan, "Two-Handed Hand Gesture Recognition for Bangla Sign Language using LDA and ANN", 2014 8th International Conference on Software Knowledge Information Management and Applications (SKIMA), pp. 1-5, 2014.

[5] Tanzila Ferdous Ayshee, Sadia Afrin Raka, Quazi Ridwan Hasib, Md. Hossain, Rashedur M Rahman, "Fuzzy Rule-Based Hand Gesture Recognition for Bengali Characters", 2014 IEEE International Advance Computing Conference (IACC), pp. 484-489, 2014.

[6] Salma Begum, Md. Hasanuzzaman, "Computer Vision-based Bangladeshi Sign Language Recognition System", 12th International Conference on Computers and Information Technology (ICCIT) 2009, pp. 414-419, 2009.

[7] Matheesha Fernando, Janaka Wijayanayaka, "Low cost approach for Real Time Sign Language Recognition", 2013 8th IEEE International Conference on Industrial and Information Systems (lCIIS), pp. 637-642, 2013.

[8] Shreyashi Narayan Sawant, M. S. Kumbhar, "Real Time Sign Language Recognition using PCA", 2014 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), pp. 1412-1415, 2014.

[9] J.R.Movellan,"Tutorial on Gabor Filters", 2008, [online] Available: http://mplab.ucsd.edu/tutorials/gabor.pdf.

[10] M. Haghighat, S. Zonouz, M. Abdel-Mottaleb, "Identification using encrypted biometrics" in Computer analysis of images and patterns (CAIP), Springer, pp. 440-448, 2013.

[11] Milad Moghaddam, Manoochehr Nahvi, Reza Hassanzadeh Pak, "Static Persian Sign Language Recognition Using Kernel-Based Feature Extraction", 7th Iranian conference on Machine Vision and Image Processing (MVIP), pp. 1-5, 2011.

[12] A. Kocsor, L. Toth, "Kernel-based feature extraction with a speech technology application", Signal Processing IEEE Transactions on, vol. 52, pp. 2250-2263, 2004.

[13] Baker-Shenk, Charlotte Lee, and Dennis Cokely. American Sign Language: A teacher's resource text on grammar and culture. Gallaudet University Press, 1991.
© Daffodil International University 19

[14] Bikash Chandra Karmokar, Kazi Md. RokibulAlam and Md. KibriaSiddiquee: "Bangladeshi Sign Language Recognition employing Neural Network Ensemble ", International Journal of Computer Applications (0975 – 8887) Volume 58– No.16, November 2012

Depression Analysis of Bangla Social Media Data using Gated Recurrent Neural Network

Abdul Hasib Uddin[1], Durjoy Bapery[2], Abu Shamim Mohammad Arif[3]
Computer Science and Engineering Discipline, *Khulna University*
Khulna, Bangladesh
Email: [1]abdulhasibuddin@gmail.com, [2]durjoybapery@gmail.com, [3]shamimarif@yahoo.com

***Abstract*—Nowadays, micro-blogging sites like Twitter, Facebook, YouTube, etc., have become much popular for social interactions. People are expressing their depression over social media, which can be analyzed to identify the causes behind their depression. Most of the researches on emotion and depression analysis are based on questionnaires and academic interviews in non-Bengali languages, especially English. These traditional methods are not always suitable for detecting human depression. In this paper, we introduced a Gated Recurrent Neural Network based depression analysis approach on Bangla social media data. We collected Bangla data from Twitter, Facebook and other sources. We selected four hyper-parameters, namely, number of Gated Recurrent Unit (GRU) layers, layer size, batch size and number of epochs, and presented step by step tuning for these Hyper-parameters. The results show the effects of these tuning steps and how the steps can be beneficial in configuring GRU models for gaining high accuracy on a significantly smaller data set. This work will help psychologists and concerned authorities of society detect depression among Bangla speaking social media users. It will also help researchers to implement Natural Language Processing tasks with Deep Learning methods.**

***Index Terms*—Depression, Bangla, Social Media, Gated Recurrent Neural Network, Hyper-parameter Tuning**

I. Introduction

The objective of Artificial Intelligence is to imitate and analyze human behaviors. In this course, detecting human sentiment and emotion is an important part for which Machine and Deep Learning approaches are being widely used. Sentiment and emotion classification can be analyzed further from two different perspectives, specifically, detecting sentiment and emotion from image data, and sentiment and emotion detection from textual data. In both cases, psychological and technical knowledge is essential to analyze the data. In a general sense, sentiment analysis covers the overall area of positive, negative, and neutral sentiment classification tasks. Emotions, such as happiness, sadness, depression, disgust, etc., are rather deep sentiments which are much more difficult to analyze. Some of these emotions are deeper than others which requires high-level psychological knowledge and more sophisticated technical approaches to study.

Depression is one of the deepest human emotions. With a mechanical way of life, more and more people are falling into depression [1]. Depression is a mental disorder which destroys not only the corresponding person but also affects the morality of that person causing him/her to commit unsocial activities, even suicide and murder. According to the World Health Organization (WHO), over 300 million people were suffering from depression in 2017, which indicates an increase of more than 18% between 2005 and 2015 [1].

Most of the research work on depression analysis are survey and one-to-one communication-based. Technical knowledge of depression analysis can help psychologists to detect and analyze depression and causes behind it. Micro-blogging sites like Facebook, Twitter, LinkedIn are now more popular than ever before. On these sites, people frequently share their daily activities and emotional reactions. Wang et al., used Sina, a Chinese micro-blogging website, to collect data and applied psychological knowledge to extract features from it [2]. They demonstrated the use of Naive Bayes, Decision Tree and Rule-based classifiers for their depression detection task.

Bangla is the seventh most spoken language in the world [3]. However, depression analysis in Bangla is still to be done. Riyadh et al., used Naive Bayes classifier with Laplacian Add-1 Smoothing for emotion classification in Bangla [4]. They collected Twitter data from Sentiment140 and used only the core texts for their research work. Their emotion classification covered happiness, sadness, surprise, and disgust. In another research work, Chowdhury et al. performed sentiment analysis in Bangla [5]. They used a hybrid mechanism consisting of both Lexicon based and Machine Learning approaches. They acted upon a semi-supervised bootstrapping approach with Support Vector Machine (SVM) and Maximum Entropy (ME). They collected their training corpus from Twitter.

Neural Network based Deep Learning approaches are becoming more and more popular for semantic analysis, such as emotion classification. Recently, in 2017, Suhara et al., introduced a self-reported history based depression detection procedure using Recurrent Neural Network (RNN) [6]. Hassan et al., used Long Short Term Memory Recurrent Neural Network (LSTM-RNN), a Deep Learning model, for performing sentiment analysis in Bangla [7]. Their work involved both normal and Romanized Bangla.

Can et al., performed multilingual sentiment analysis on limited data using RNN based model [8]. They assembled data from Amazon reviews, Yelp restaurant reviews, and Competition restaurant reviews. Their work covered sentiment analysis in English, Spanish, Turkish, Dutch, and Russian languages. Ayata et al. used both Deep Learning and Machine Learning approach for sentiment analysis [9]. They gathered data from Twitter and trained Support Vector Machine (SVM), Random Forest (RF), Naive Bayes, and LSTM-RNN. Their work concluded with the decision that in most of the cases, Deep Neural Network based approaches perform better than Machine Learning approaches. In their work, only SVM accomplished slightly better performance than LSTM.

In this paper, we introduced the Gated Recurrent Unit (GRU) Neural Network based depression analysis on small Bangla dataset, collected from Twitter, Facebook, and native Bangla speakers using Google form. Labeling human emotion related data requires knowledge of human behaviors. Hence, we hired a Sociology student, who is experienced in dealing with human emotions, to label our dataset. We applied step by step Hyper-parameter tuning while implementing GRU model. Then, we compared each implementation of Hyper-parameter tuning and showed how to tune different Hyper-parameters in an easy way to accomplish better performance on a very small dataset. We have manipulated four different Hyper-parameters, specifically, model size, no. of layers, batch size, and no. of epochs. Apart from applying the GRU model for depression detection from Bangla social media data, our other aim was to focus specifically on GRU model performance over a significantly small dataset. In Bangla, it is reasonably difficult to manage a large dataset due to lack of enough research works. Hence, for Hyper-parameter tuning, we only focused on the aforementioned four Hyper-parameters, without following any specific Hyper-parameter optimization approach.

The rest of this paper is arranged as follows. In Section II, we have discussed some relevant existing research works, our proposed method is described in Section III, and in section IV, the results of our method and relevant discussions are presented. Finally, we conclude our work in Section V along with some future plans.

II. Related Works

In this section, some available research works related to Machine and Deep Learning based sentiment and emotion analyses are described.

A. Machine Learning for Emotion Analysis in Bangla

A Machine Learning based human emotion analysis approach is described by Riyadh et al., in [4]. In this research work, authors selected happiness, sadness, surprise, and disgust for their classification task. They collected tweets from Sentiment140, labeled them manually, eliminated tweets with no emotion, and created a balanced dataset containing 3,750 tweets. 3,500 tweets were selected as their training dataset and 250 tweets as the testing dataset. For feature extraction, Unigram model and Unigram model with POS tagging were used. Authors used the frequency of Bag of Words model as a feature to train their classifier. They used Multinomial Naïve Bayes classifier for the classification task. To avoid zero probability problem of Naïve Bayes classifier, they used Laplacian Add-1 Smoothing by assuming the training dataset to be very large and by adding one to each count. Against 4-way classification, they gained an average of 81% accuracy for Unigram model. For the Unigram model with POS tagging, they gained an average of 79.5% accuracy. For the 5-way classification task, they gained an average of 66% for Uni-gram model. For Uni-gram model with POS tagging, they gained an average of 64.8% accuracy.

B. RNN for Depression Forecasting

A novel method for depression forecasting was established by Suhara et al., using RNN [6]. The authors developed LSTM-RNN based deep learning algorithm. They used their model to produce embedding layers regarding every categorical variable, which also incorporates a day-of-the-week variable to determine the day-of-the-week consequences in their model. They collected depressed data from 2,382 self-declared depressed persons, covering 22 months time span, via a smartphone application. Their model was successfully able to forecast 84.6%, 82.1%, and 80.0% severe depression cases in 1, 3, and 7 days beforehand, respectively. Additionally, the model was successful in detecting overall depression cases with 88.6%, 86.0%, and 84.2% accuracy for forecasting 1, 3, and 7 days beforehand, respectively.

C. Machine Learning for Depression Analysis

Wang et al. conducted an experiment on Sina micro-blog, a Chinese micro-blog, which is one of the most influential social media services in China [2]. They emerged both Psychological and Machine Learning knowledge for their experiment. From Psychological perspective, they used ten features, such as 1st person singular, 1st person plural, positive emoticons, negative emoticons, mentioning, being forwarded, being commented, original blogs, blogs posted between 0:00 - 6:00 o'clock, and one other. From the technical perspective, Machine Learning techniques, such as Decision Tree, Naive Bayes, and Rule-based classifiers were used. Their described method contained mainly three steps, namely, sentence and word segmentation, polarity calculation of sub-sentences, and polarity calculation of sentences. Their model was able to accomplish 80% precision.

D. CNN and RNN for Natural Language Processing

Yin et al. presented a systematic comparison between CNN and RNN in [10]. Their research work covered a wide range of representative NLP tasks, providing basic

guidance for Deep Neural Network (DNN) selection. They systematically compared CNN, LSTM, and GRU models. It covered a broad range of NLP tasks, such as sentiment classification, relation classification, textual entailment, answer selection, question relation match, path query answering, part-of-speech tagging, etc. For their experiment, on each step, they trained their model from scratch, used basic setup without complex tricks, like batch normalization, searched for optimal Hyper-parameters, analyzed the basic architecture and utilization of the models. This experiment concluded that, generally, RNNs perform better and robust for NLP tasks, except for some cases, like keyphrase recognition task.

III. Methodology

Our proposed approach for depression detection from Bangla dataset is divided into two steps - *Creating Bangla Dataset* and *Training GRU Model*. We put special efforts for preparing Bangla dataset and shown steps for Hyperparameter tuning.

A. Creating Bangla Dataset

We used Twitter as our primary data source. We collected 5,000 Bangla data from Twitter and 210 depressed Bangla statements from native Bangla speakers using google form. We prepared our dataset into three steps - *Data Pre-processing*, *Data Labeling*, and *Data Post-processing*.

1) Data Pre-processing: For data pre-processing task, we created a white list containing Bangla alphanumeric characters, punctuations, and space. We scanned each tweet character by character, filtered out all non-white listed characters, and removed multiple consecutive white spaces. An example of twitter data cleaning is given in Fig 2. We only had to pre-process the tweets, as data collected via google form were clean.

Table I: Part of the labeled and stratified dataset

Data no.	Data	Label
34	ওরাল রিহাইড্রেশন সলিউশন ও আর এস প্রস্তুত করার ১২ ঘন্টার মধ্যে বা রেফ্রিজারেটরে রাখলে তা ২৪ ঘন্টার মধ্যে পান করে ফেলা উচিত ।	Non-depressive
33	রাজনাথ সিং সংসদে বললেন এই নাকি ফাইনাল নয় । এই লক্ষ লোক কে আবার নাকি সুযোগ দেওয়া হবে নাগরিত্ব প্রমাণ করার । এটা কি আইওয়াশ নয় ? এরা আর কি প্রমাণ দেবেন ? পাসপোর্ট , আধার ভোটার কার্ড সব ই তো বাতিল করে দিয়েছে । আবার অনেক পুরুষের নাম আছে অথচ তাদের পরিবারের নাম নেই । এরা কি করবে ?	Depressive
118	আহা ! গভীর গহী নের নিঃশব্দ সুর সা থে আঁধা রের ঘ্রাণ !	Non-depressive
117	প্রায় ৪০ লক্ষ বাঙালি দেশ ছাড়া হতে চলেছে সামলাতে পারবে তো পৃথিবী ?	Depressive

2) Data Labeling: Deep human emotions, like depression, are extremely difficult to analyze. In order to make our depression detection dataset more reliable, we hired a Bangla speaking Sociology student, experienced in dealing with human emotions, to manually label our Bangla dataset. A total of 5,000 Bangla tweets were labeled. After labeling, we eliminated incomplete and ambiguous tweets. Our initial tweet set consisted of 984 depressive tweets and 2,930 non-depressive tweets.

3) Data Post-processing: After pre-processing and labeling our dataset, we eliminated redundant data from both tweet set and depressed statement set collected via a google form. After removing redundancy, our dataset was imbalanced with 1,289 non-depressive and only 588 depressive data. An imbalanced dataset may lead to wrong accuracy and over-fitting problems, where accuracy for data within the training dataset would be high, while accuracy for data outside of the training dataset would be low. Hence, we randomly selected only 588 non-depressive data from the labeled dataset to balance with 588 depressive data. In this step, our balanced dataset consisted of only 1,176 data of which 588 were depressive and 588 were non-depressive.

Stratifying Dataset. The balanced dataset was very small to train with a neural network. Therefore, we stratified our dataset to reduce the effect of its small size while training the GRU model. In this step, data were rearranged into a one-to-one approach, that is, one depressive data followed by one non-depressive data, and continuously repeated this process all over the balanced dataset. After stratifying our dataset, we considered this dataset as final and used it to train our GRU model.

Part of an example of labeled and stratified data set is given in Table I. The table shows some Bangla data along with their original labels. The 'Data no.' signifies the positions of the data in the final dataset, and how the final dataset was stratified with one non-depressive data followed by one depressive data.

B. Training GRU Model

For training our Gated Recurrent Neural Network Model, we again divided the training phase into two substeps - *Dataset Splitting*, and *Hyper-parameter Tuning and Training*.

1) Dataset Splitting: Each time before training our model, we split our dataset into 80% training, 10% validation and 10% testing set. Validation dataset was only used to avoid over-fitting. We compared different trained models according to their corresponding testing accuracy.

2) Hyper-parameter Tuning and Training: For implementing our model, we utilized tensorflow implementation described in [11]. To get an accurate result using our small dataset, we fixed learning rate of our model to a very low value of 0.0001. Low learning rate helps the model to avoid over-fitting problem so that our model detects depression accurately from any given data. We tuned 4 GRU model

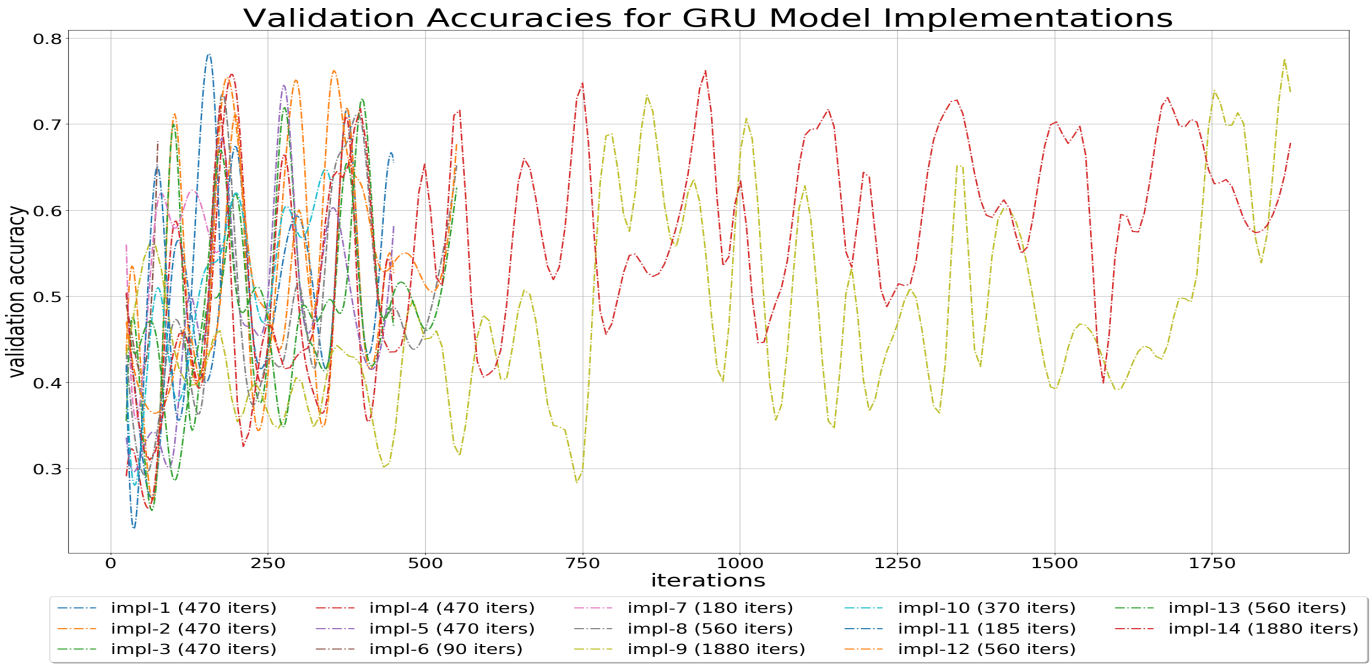


Figure 1: Comparing GRU implementation performances in terms of validation accuracies.

Original data:
নদীর নীচে দেশের প্রথম রেল সুড়ঙ্গ তৈরি হয়েছে কলকাতায়। এবার মেট্রো রাইলের সৌজন্যে দেশের সবথেকে বড় ভূগর্ভস্থ রেল ইয়ার্ড পেতে চলেছে কলকাতায়। যা তৈরি হচ্ছে কলকাতা বিমানবন্দরের ঠিক পাশেই।https://ebela.in/state/kolkata-metro-railway-is-constructing-the-country-s-biggest-underground-rail-yard-near-airport-dgtl-1.839389?ref=state-new-stry
...

After pre-processing:
নদীর নীচে দেশের প্রথম রেল সুড়ঙ্গ তৈরি হয়েছে কলকাতায়।এবার মেট্রো রাইলের সৌজন্যে দেশের সবথেকে বড় ভূগর্ভস্থ রেল ইয়ার্ড পেতে চলেছে কলকাতায়।যা তৈরি হচ্ছে কলকাতা বিমানবন্দরের ঠিক পাশেই। : ?

Figure 2: Example of data pre-processing.

Hyper-parameters into three steps - *Tuning for GRU Size*, *Tuning Batch Size with No. of Epochs*, and *Tuning No. of GRU layers with No. of Epochs*. Fig 1 represents the effects of Hyper-parameter tuning in each implementation (impl) in terms of validation accuracies.

a) Tuning for GRU Size: We initially fixed no. of GRU layers to 5, batch size to 10 and, no. of epochs to 5 (total 470 iterations) and tuned GRU size over these parameters. We trained our GRU model with size 64, 128, 256, 512, and 1024. Impacts for tuning GRU size are shown in Fig 1 (impl 1 to 5). Our model gained the highest accuracy for GRU size 512 (impl 4).

b) Tuning Batch Size with No. of Epochs: After tuning GRU size, we fixed its size to 512, no. of layers to 5, and tuned batch size along with no. of epochs. We trained our model with batch size and no. of epochs set to 50 and 5 (90 iters), 50 and 10 (180 iters), 5 and 3 (560 iters), 1 and 2 (1880 iters), 25 and 10 (370 iters), 25 and 5 (185 iters), respectively. As time complexity increases while batch size decreases, we had to decrease no. of epochs corresponding to small batch size. With batch size 5 over 3 epochs, our model accomplished the highest accuracy for tuning batch size with no. of epochs. Effects for tuning batch size with corresponding no. of epochs are shown in Fig 1 (impl 6 to 11). It indicates how accuracy changes with different batch size and no. of epochs.

c) Tuning No. of GRU layers with No. of Epochs: In the final step of Hyper-parameter tuning, we fixed GRU size to 512, batch size to 5, and tuned no. of GRU layers of our model. To reduce memory complexity, we decreased no. of epochs while increasing no. of GRU layers. We trained our model with GRU layers and no. of epochs set to 3 and 3 (560 iters), 10 and 3 (560 iters), and 5 and 10 (1880 iters), respectively. For tuning GRU layers with no. of epochs, our model attained the highest result for 3 layers over 3 epochs. The influences of tuning no. of GRU layers on learning curve are shown in Fig 1 (impl 12 to 14).

IV. Result and Discussion

All the steps of Hyper-parameter tuning along with their corresponding test accuracies are given in Table II. On the first step of Hyper-parameter tuning, we started with a low GRU size of only 64 and gradually increased the size to 1024. Our GRU model achieved 59.1%, 70.0%, 67.3%, 74.5%, and 69.1% accuracy for GRU size 64, 128, 256, 512, and 1024, respectively. According to these results, with

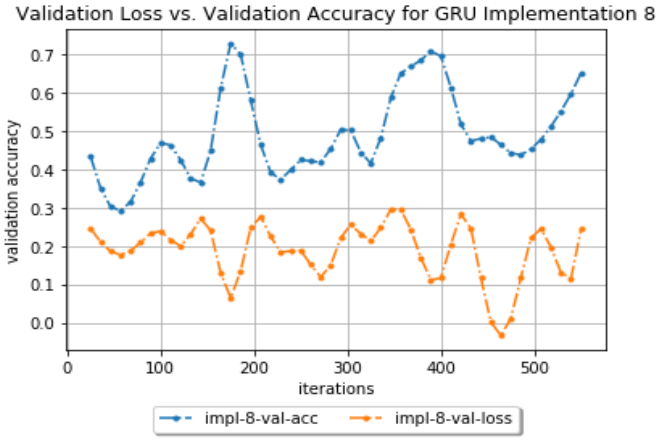


Figure 3: Validation loss vs. validation accuracy for GRU best implementation.

Table II: GRU Hyper-parameter tuning results

GRU implementation no.	GRU size	No. of GRU layers	Batch size	No. of epochs	Test accuracy
1	64	5	10	5	59.1%
2	128	5	10	5	70.0%
3	256	5	10	5	67.3%
4	512	5	10	5	74.5%
5	1024	5	10	5	69.1%
6	512	5	50	5	52.0%
7	512	5	50	10	61.0%
8	512	5	5	3	75.7%
9	512	5	1	2	70.3%
10	512	5	25	10	57.0%
11	512	5	25	5	61.0%
12	512	3	5	3	74.8%
13	512	10	5	3	69.6%
14	512	5	5	10	56.5%

size 512, the accuracy was the highest for this step [impl 4]. These results show that, in most of the cases, accuracy and GRU size acts proportionally.

As on the first step of Hyper-parameter tuning, the highest accuracy was generated for GRU size 512, we fixed this size on the second step, and tuned batch size along with no. of epochs. Our GRU model gained 52.0% accuracy for batch size 50 over 5 epochs, 61.0% accuracy for batch size 50 over 10 epochs, 75.7% accuracy for batch size 5 over 3 epochs, 70.3% accuracy for batch size 1 (that is, online learning) over 2 epochs, 57.0% accuracy for batch size 25 over 10 epochs, and 61.0% accuracy for batch size 25 over 5 epochs. Here, the results signify that accuracy falls for a very large batch size. Also, it is shown that the required no. of epochs is highly dependent on the batch size. However, for fixed batch size, accuracy increases with increasing no. of epochs. Hence, in case of applying GRU, a set of balanced batch size with reasonably large no. of epochs is required for high accuracy. The highest accuracy for this step was 75.7% for batch size 5 over 3 epochs [impl 8]. Accuracy for training with the online learning approach is 70.3% for over 2 epochs.

On the last step of Hyper-parameter tuning, we fixed GRU size to 512, batch size to 5, and tuned no. of GRU layers along with no. of epochs. Our GRU model accomplished 74.8% accuracy with 5 layers over 3 epochs, 69.6% accuracy with 10 layers over 3 epochs, and 56.5% accuracy for 5 layers over 10 epochs. These results led us to the decision that large no. of GRU layers does not necessarily help to perform better on a small dataset. The highest accuracy for this step was 74.8% for 3 layers over 3 epochs [impl 12].

Comparison among all GRU model implementation test accuracies is presented in Fig 4. The validation loss and validation accuracy of the best implementation [impl 8] for Hyper-parameter tuning are shown in Figure 3. The best accuracy we obtained was 75.7% for 5 layered GRU with size 512, batch size 5, and the learning rate 0.0001 over 3 epochs.

This work is the first attempt to utilize a Deep Learning approach for depression analysis in Bangla. Therefore, we are unable to directly compare our results with any relevant works. However, the method described by Suhara et al., is quite similar to ours [6]. They also used a Deep Learning approach for depression analysis. Nonetheless, they were able to create one of the largest depression datasets with 345,158 records over 22 months time span. Hence, they were able to apply LSTM-RNN with their huge dataset. On the contrary, our small dataset contains only 1,176 data, which is crucially insufficient to train LSTM model. This led us to utilize GRU model. Also, their work was to forecast depression in English. On the other hand, instead of forecasting, our model detects depression, which is in Bangla. Embracing all these constraints, our depression detection model still is able to gain maximum 75.7% accuracy, comparing to their LSTM model, which was able to gain maximum 88.6% accuracy. If we were able to create such kind of large dataset, we might be able to accomplish higher performances.

V. Conclusion and Future Works

In this research, we established a GRU model based depression detection technique by analyzing Bangla text data collected from social media (Twitter, Facebook, and Google form). We applied a sequence of Hyper-parameter tuning and showed the corresponding effects of these operations on a small dataset. The results indicate that

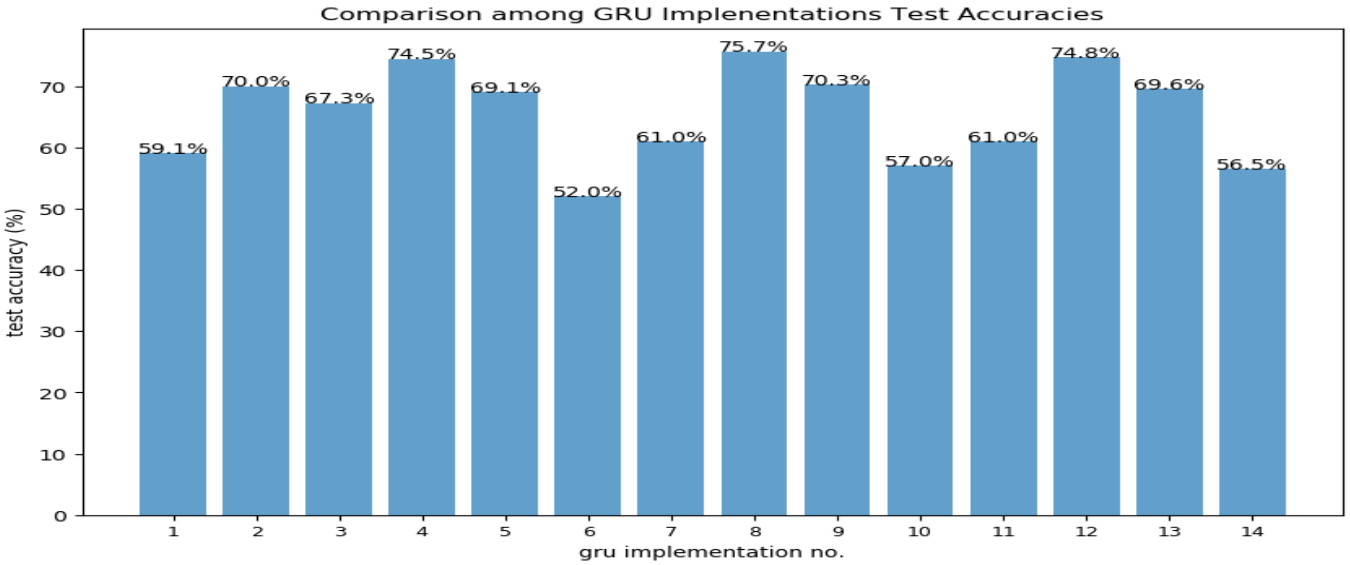


Figure 4: Comparing GRU implementation test accuracies.

in the case of a small dataset, accuracy depends on the variations of batch size and the number of epochs. We have also shown that the required number of epochs is highly dependent on the batch size. However, for fixed batch size, accuracy increases with increasing number of epochs. The results conclude that in case of applying the GRU model on a significantly small dataset, a set of balanced batch size with a reasonably large number of epochs is required to obtain adequate high performance.

This research will guide further investigations to analyze Bangla social media data for detecting depression from both small and large dataset using Deep Learning models. Additionally, multiple experienced persons can be used for labeling the dataset to make it more reliable. Furthermore, we have tuned only four Hyper-parameters without following any specific Hyper-parameter optimization approach. However, such kind of approaches, like Evolutionary optimization, Random search, Gradient-based optimization, and others can be applied for more in-depth analysis.

References

[1] "Depression," World Health Organization, 04-Jul-2017. [Online]. Available: https://www.who.int/mental_health/management/depression/en/. [Accessed: 21-Dec-2018].

[2] X. Wang, C. Zhang, Y. Ji, L. Sun, L. Wu, and Z. Bao, "A depression detection model based on sentiment analysis in micro-blog social network," in Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 201–213, Springer, 2013.

[3] "Languages of the World," Ethnologue. [Online]. Available: https://www.ethnologue.com/. [Accessed: 02-Feb-2019].

[4] A. Z. Riyadh, N. Alvi, and K. H. Talukder, "Exploring human emotion via twitter," in 2017 20th International Conference of Computer and Information Technology (ICCIT), pp. 1–5, IEEE, 2017.

[5] S. Chowdhury and W. Chowdhury, "Performing sentiment analysis in bangla microblog posts," in 2014 International Conference on Informatics, Electronics & Vision (ICIEV), pp. 1–6, IEEE, 2014.

[6] Y. Suhara, Y. Xu, and A. Pentland, "Deepmood: Forecasting depressed mood based on self-reported histories via recurrent neural networks," in Proceedings of the 26th International Conference on World Wide Web, pp. 715–724, International World Wide Web Conferences Steering Committee, 2017.

[7] A. Hassan, M. R. Amin, A. K. Al Azad, and N. Mo- hammed, "Sentiment analysis on bangla and roman-ized bangla text using deep recurrent models," in 2016 International Workshop on Computational Intelligence (IWCI), pp. 51–56, IEEE, 2016.

[8] E. F. Can, A. Ezen-Can, and F. Can, "Multilingual sentiment analysis: An rnn-based framework for limited data," arXiv preprint arXiv:1806.04511, 2018.

[9] D. Ayata, M. Saraclar, and A. Ozgur, "Busem at semeval-2017 task 4a sentiment analysis with word em- bedding and long short term memory rnn approaches," in Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), pp. 777–783, 2017.

[10] W. Yin, K. Kann, M. Yu, and H. Schütze, "Comparative study of cnn and rnn for natural language processing," arXiv preprint arXiv:1702.01923, 2017.

[11] Mchablani, "mchablani/deep-learning," GitHub. [Online]. Available: https://github.com/mchablani/deep-learning/blob/master/sentiment-rnn/Sentiment_RNN.ipynb. [Accessed: 30-Jan-2019].

Designing an automatic workplace recommendation system for garment workers using fog computing

1st Arunavo Dey
Lecturer, Department of Computer Science and Engineering
Bangladesh University of Business and Technology (BUBT)
Dhaka, Bangladesh
arunavo071@gmail.com

2nd Md. Nahid Newaz
Lecturer, Department of Computer Science and Engineering
Bangladesh University of Business and Technology (BUBT)
Dhaka, Bangladesh
md.nahidnewaz@gmail.com

***Abstract*—Garments industry is one of the biggest industries of Bangladesh and a large portion of our population comprises of garments workers who face several problems starting from health and financial problems, discrimination, safety issues and lack of skill development opportunities – which make them more vulnerable and obstruct them from finding a better tomorrow. A country can't march forward with a large portion of it leaving behind and thus we target to improve their conditions by designing an intervention which propels to achieve the target by maintaining a predictive model which predicts on the data collected from the users. It suggests the users who are garment workers, about the available and better opportunities for workspace in the easiest way.**

Keywords—Human Computer Interaction, Mobile, fog computing

I. INTRODUCTION

RMGs are the finished textile product from clothing factories and the Bangladeshi RMG sector is one of the fastest growing sectors in the Bangladeshi economy, with a growth rate of 55% from 2002 to 2012 [1]. Exports of textiles, clothing, and ready-made garments (RMG) accounted for 77% of Bangladesh's total merchandise exports in 2002. Bangladesh's overall export earnings rose by 5.8% to $36.66 billion in FY17 where Bangladesh's export earnings from the apparel sector registered 8.76% growth, reaching $30.61 billion, thus the apparel sector contributing 83.49%, to the country's total exports of $36.66 billion [2]. This not only indicates how much important the sector is for our economy but also how much people are dependent on this and how large the industry is. By 2013, there were approximately 5000 factories. And 3.6 million people out of a total population of over 150 million are being employed by this industry [3]. This huge population suffer from various problems including health issues, economic crisis, life risk, job uncertainty, lack of socio economic development and many others. In this paper we tried to identify and categorize their problems and proposed an interface based on their basic needs, ease of use and usability of this proposed interface. We also proposed a prediction model that predicts suitable workplaces to the users based on the data from all the users.

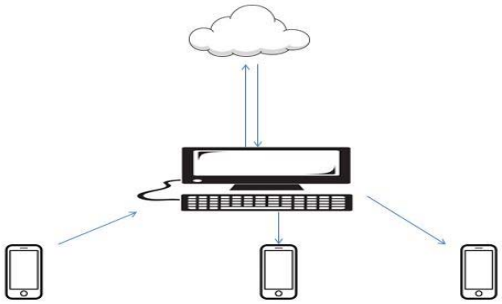


Fig: 1 Automated Model with cloud Architecture

II. RELATED WORK

Finding better and suitable workspace is one of the demanding aspects for almost every employee. So a significant amount of research has been carried out in these fields. Moreover fog computing has been applied to solve numerous problems of HCI too. In order to present our study in more easier and simplistic form we separate our related study into two groups – one is related existing interventions and another is related research works.

A. Related Existing Applications

There are a lot of existing interventions that provide recommendations to the users in diverse fields of interests. But there is none for garment workers for recommending better workplace suitable for them. Criticker is a movie recommendation system which allows users to rank and rate movies and then compare this rating with those of other users and employ "*Taste compatibility Index*" to see how closely it matches with others and then recommend movies that his/her similar testes users prefer [4]. Jinni is another movie recommendation system which uses consumption behavior to create "*unique taste profile*" for each individual or household and also provide user's mood based and effective group based recommendation [5]. There are also a lot of recommendation systems based on user's profile, mood, feedback, ratings and reviews in diverse fields of interests like music, books, games [6] and news [7].There are also a lot of applications for recommending best hotels, restaurants, hospitals etc.

B. Related Research Works

Many researches has been pursued for incorporating IOT with HCI. In 2003, Shou-Hsia Cheng, Ming-Chin Yang and Tung-Liang Chiang uses patient's

ratings of interpersonal and technical performance of the hospital in their study of patient's satisfaction and hospital recommendation system [8].

In 2009, Zei-Mei Wang and Fan Yang utilized GPS information, E-map, comments, tags and ratings for their study of personalized location based restaurant recommendation and navigation services [9].

In 2012, Keunho Choi, Donghee Yoo, Gunwoo Kim and Yongmoo Suh used a hybrid approach of collaborative *filtering* and *sequential pattern analysis* for recommending online shopping malls to the users [10]. In 2015,Kai Zhang, Keqiang Wang, Xiaoling Wang, Cheqing Jin and Aoying Zhou applied collaborative filtering, content based methods, travel intents and diversity techniques in their study of hotel recommendation system based on user preference analysis [11]. In 2015, Binita Gupta proposed a recommendation engine that provides recommendation by collecting information from various real time IOT devices [12]. In 2017, Agostino Forestiero built a distributed recommendation system which uses multi agent algorithm and decentralized and self-organizing strategy [13].

III. PROTOTYPING AND FEATURES SELECTION

Before implementing the interface for the intervention we have done brainstorming, concept mapping, storyboarding and paper prototyping for features selection purposes and we have done two different surveys. One is for validity of our selected features and another is for testing the usability of our designed interface.

A. *Brainstorming and Concept Mapping*

Before starting our work for designing the intervention for garment workers, we first conducts brainstorming for identifying the problems of the garment workers. Then we perform another brainstorming session for selecting important features for the intervention that can provide solutions for those identified problems. After that we groups the number of ideas and thoughts that we have generated in our first and second brainstorming into various concepts.

B. *Storyboarding and Paper Prototyping*

After finishing our concept mapping part, we creates a storyboarding session for our intervention design. This storyboarding helps us to visualize the problems and our proposed solutions for overcoming those problems in a story telling manner. This also helps us to visualize how our solutions will work in real time scenarios. After done with storyboarding we create multiple paper prototypes of our intervention for designing and testing user interfaces. This paper prototyping helps us for easy modification and also eliminates our chance of missing any key features from the design. After finalizing our paper prototype, we designed our application interface in android platform for surveying the usability of our interfaces.

C. *Survey for Features Selection*

We conducted a survey for validating our identified features from the real garments workers. We collected survey results from 30 garments workers. Our survey includes male and female, young and old garment workers. We created 15 key questions to validate our identified problems from the users. We used Likert scale questionnaire for this survey. Likert scale questionnaire is a psychometric scale questionnaire that is used to measure a respondent's attitude or opinion towards a given subject [12]. As most of the garment workers were illiterate, we have to speak the questions to them in Bangla language and collect the response from them. We got highly positive responses for almost all of our questions from the respondents. The results of our survey questionnaires are presented in the following figure.

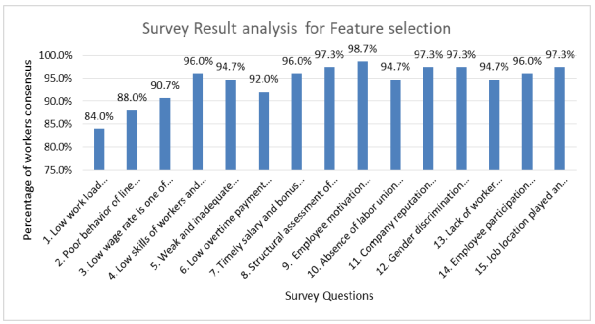


Figure 2: Analysis of survey results in features selection survey

D. *Survey for Usability Testing*

After designing our interfaces we also conducts a usability testing survey on garments workers to test the usability of our application. We used Think Aloud Protocol to conduct this survey [13]. In our survey we test the rollback, interactivity, heuristics, readability, maintainability, intuitiveness, usability, visualization and coherence of our interface. We conducted usability testing survey on 10 garments workers. The following figure shows the results of our usability testing survey.

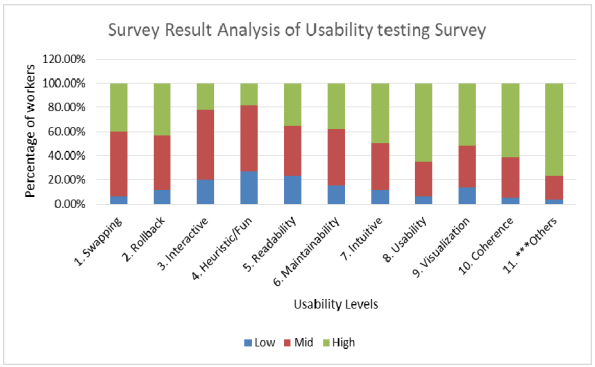


Figure 3: Survey result analysis of usability testing survey

IV. INTERFACE DESIGN

Analyzing user's convenience and usability, we designed the interface in such a way that it becomes at the same time user friendly to the targeted users and efficient in achieving its goal. From our survey we got that most of the users are familiar with form type registration system and thus we created an interface for log in with an extra button named sign up which generates a form when pressed. After filling the form, users get them registered. From our surveys we also saw that 60% users are used to see the main interface after log in and profile in another linked tab. Surveys also showed us that users are also used to find their profile by clicking their picture or any person icon. That's why we placed a person icon in the left corner by clicking

which a user can visit his or her profile. Main interface holds simply a button which finds the most suitable jobs by the predictive model when pressed. In user profile, there is two other tabs one for review where users can find different workplaces by their names and can also review them. In the other tab, there are messages from the system which can also be generated by the users in case of emergency. This software is targeted for workers, most of which are dropouts from primary and that's why we equipped each screen and symbol with Bengali words to make the best use of this app.

V. Data collection

As data plays a significant role in this model thus we propose collection of data from the interface in such a way that it becomes easy for being provided by the users as well as being highly efficient at the same time. While registration, we collect users location which is stored immediately to the nearest fog node of our proposed model. Again if any user decides to review any workplace in the review section, a questionnaire with five questions pops up immediately for each rating criterion of the workplace each is having five different ratings as answers. As we have to rely on these questions and answers for rating these workplaces, we conducted survey for setting which questions are to be placed. All that has to be done by the users is to rate that workplace against these questions which will be saved in the fog nodes. The fog nodes are the ones who hold the predictive model in our proposed system which is fed by these ratings and generates its own rating. Again these ratings are sent to the users and a list of workplaces according to this list and other metrics related to the worker is sent to the users. When any user presses the button named find job in main interface, it simply shows this list. All the data will also be updated periodically to the cloud for making the model better with feedbacks.

VI. Proposed model

We proposed a very simplistic technique to rate the workplaces. Each of the five answers with each of the questions related to the places possesses different values. We propose a rather straightforward approach where each question has ratings as answers to the questions or in other words each place can be rated against each question in scale of 1 to 5. After getting these ratings a very general model inspired from Naïve Bayes model calculates values for these places. When any user rates a workplace, its rating gets updated.

Fog part:

Upon submitting ratings in the review interface, we store them temporarily in user mobiles and in network availability; they are uploaded to fog nodes. Fog nodes initially assign values to all the workplaces. Upon receiving the ratings, it categorizes them by their criterions and respective questions where each question has different weights. For calculating the rating we proposed a novel algorithm as follows,

Algorithm

change_index = **1**
For each $r_i \in$ rating_criterions,R
  For each $q_i \in$ question, Q
    For each $p_i \in$ rating_type, P
      If given_rate equals p_i
        If submitted_rate, p_z exists then
          $p_z = p_z - (1 * change\_index)$
          If (number_of_times_chnaged > 2)
          change_index = 1/ number_of_times_changed
          $p_i = p_i + (1 * change\_index)$
        else $p_i = p_i + 1$
      Probability $[p_i] = p_i / \sum_{i=0}^{n} p_i$
    For all $p_i \in$ Probablity,P
      rating $[q_i]$ = max $p_i \in P$
  For each $q_i \in$ question, Q
    weighted$[q_i] = w_i *$ rating $[q_i] / \sum_{i=0}^{n} w_i *$ rating $[q_i]$
For each $r_i \in$ rating_criterions,R
  Probability $[r_i]$ = max weighted$[q_i]$ **For All** $q_i \in$ question, Q

All the weights are between 0 and 1. Whichever place gets the higher score moves forward in the list. All the ties are being broken arbitrarily. All the calculations are to be performed in fog nodes. These nodes, after performing these calculations, send the lists to the users which are being shown to them whenever they try to find a job. This list also gets some modifications with the distance metric. Ties in jobs scores are broken by their distance to the individual' s location searching suitable job where nearest jobs gets priority over jobs with equal score.

There is another interface named Messages. This has been proposed with the view of sending emergency messages to the nearest ones in times. If any garments worker writes in the box in the lower part of that tab, that message is being sent immediately to the fog node which can be sent to other users nearby the sender. Fog finds the suitable receivers and forwards them that message.

Alongside performing calculation, fog nodes can send periodic updates to the cloud where cloud can cross match the ratings and cloud can send updates to the list prepared by fog nodes.

Cloud part:

Cloud is periodically updated with the data from the fog node. When number of emergency messages from a certain place exceeds a certain threshold level, it decreases the rating and same occurrence of this event increases the rate of decrease in the rating. This rating is transferred to the fog node and eventually sent to the edge devices (user mobiles) without any modification.

VII. Experiment & performance evaluation

In our experimental setup, we set up a slum of garments workers to run our application in small scale. We setup a local host which acted as a fog node and installed our application in approximately in 50 workers mobiles and set

up a Wi-Fi under which the local host operated. From that, we tracked down the trend of number of users.

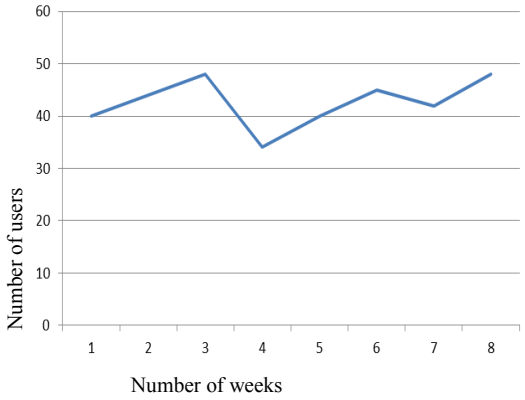


Figure 4: Survey result of number of users

The very first breakdown in the number of users accounts for their initial struggle with the interface. After making them more familiar with the interface, the second breakdown occurs for their concern about the security of their data as it contains highly sensitive information. Though making them understand that it's only for test purpose cleared this, the concern is real and it is a drawback of our system that it doesn't provide any strong measure of security which is needed to be addressed too. We also conducted a survey on 1.Usability, 2. Ease of use, 3. Features and 4.Prospects and the responses above good are projected in a manner below

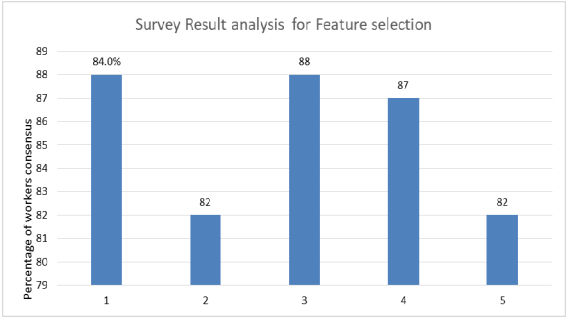


Figure 5: Survey result analysis of usability survey

VIII. FUTURE WORKS

This proposed model can be implemented for various rating prediction including other occupations and educational sectors as well. We proposed this model in job prediction context, but this proposed model can be of good work disaster based systems which we currently working on. Moreover the prediction model is too naïve which surely possess a great scope to improve. If a better model, trained effectively with a good set of training data can be formed it can predict more accurate result which is also our next motive. Moreover, this proposed rating shows suitable workplace based on everyone's ranking whereas suitability can vary from person to person. So suggesting suitable workplace to the user with the help of specific rating provided by that user is also our next motive.

IX. CONCLUSION

Though a lot of recommendation system has been proposed before, our approach in this topic is unique of its kind. The issue of security and implementation of proper neural network is somewhat neglected as this is only a conceptual model but needed to be addressed for a practical approach. Moreover proper analysis of this data can surely help to achieve the targeted goal.

REFERENCES

[1] https://en.wikipedia.org/wiki/Textile_industry_in_Bangladesh

[2] https://www.dhakatribune.com/business/2018/08/14/rmg-exports-record-21-72-growth-in-july

[3] https://asiafoundation.org/slideshow/bangladeshs-garment-workers

[4] Yosi Glick, Izik Ben-Zaken, http://www.jinni.com/discovery/solutions.html

[5] Massimo Petrella, http://www.stereomood.com/en

[6] TasteDive team, https://tastedive.com/

[7] Gabriel de Souza Pereira Moreira, Felipe Ferreira, and Adilson Marques da Cunha. 2018. News Session-Based Recommendations using Deep Neural Networks. In 3rd Workshop on Deep Learning for Recommender Systems.

[8] Shou-Hsia Cheng, Ming-Chin Yang and Tung-Liang Chiang."Patients satisfaction with the recommendation of hospital: effect of interpersonal and technical aspects of hospital care".International Journal of Quality and health care, 2003, Volume 15, PP. 345-355

[9] Zei-Mei Wang , Fan Yang. "An optimized location based mobile restaurant recommend and navigation system".WSEAS transaction on information science and application, Issue 5, Volume 6, May 2009.

[10] Keunho Choi, Donghee Yoo, Gunwoo Kim and Yongmoo Suh. "A hybrid online product recommendation system: Combining implicit rating-based collaborative filtering and sequential pattern analysis". Electronic commerce research and applications, ELSEVIER, Volume 11, Issue 4, 2012, page. 309-317

[11] Kai Zhang, Keqiang Wang, Xiaoling Wang, Cheqing Jin and Aoying Zhou. "Hotel Recommendation based on user preference Analysis". ICDE workshops, IEEE, 2015.

[12] Binita Gupta. "Real-time context aware recommendation engine based on a user internet of things environment". United States Patent Application Publication, January 15, 2015.

[13] Agostino Forestiero. " Multi-Agent Recommendation System in Internet of Things". 2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)

[14] Wuensch, Karl L. (October 4, 2005). "What is a Likert Scale? and How Do You Pronounce 'Likert?'". East Carolina University. Retrieved April 30, 2009.

[15] Lewis, C. H. (1982). *Using the "Thinking Aloud" Method In Cognitive Interface Design* (Technical report). IBM. RC-9265

Retinal Fundus Identification Utilizing Supervised and Unsupervised Nature of Deep Neural Network

Sanaullah Chowdhury†, Faozia Rashid Taimy†, Abdullah-Al Nahid†, Md Younus Ali†, Ferdous bin Ali††
†*Electronics & Communication Engineering Discipline*, Khulna University, Khulna, 9208, Bangladesh
††*Grey Tools Technology*, Bangladesh
sanaullahashfat, faozia.rashid.taimy, nahid.ece.ku}@gmail.com, younus.ali@ece.ku.ac.bd, hridoyferdous@yahoo.com

***Abstract*—Eye disease has become serious concern to people, specially who have been suffering from diabetes . Systematic eye diseases are manifest in around the retina of an eye. Digital photography of the retinal images provide a significant information about the disease. This kind of retinal images work as a powerful source for the doctors and ophthalmologist to take meaningful decision about the disease. This kind of Computer Aided Diagnosis(CAD) allow both the doctors and patient to take a second layer of satisfaction. Considering this issues, this paper segmented the retinal vessel images using the state-of-the-art Deep Neural Network(DNN) such as a combination of U-Net and Generative Adversarial Network (GAN) model. As the retinal images suffers due to the Contrast and Illumination problems, this paper also proposed two preprocessing algorithms to reduce this effect.**

***Index Terms*—Accuracy, Augmentation, CLAHE, GAN, Segmentation, U-Net**

I. Introduction

A complicated problem of an eye is Diabetic retinopathy that can affect eyes badly. The back of the eye blood vessels can be damaged by the light-sensitive tissue for this type of complication. Statistics show that this complications affect up to 80 percent [24] of those who is suffering from diabetes for 20 years or more. About 40 to 45 percent [26] Americans having diabetes who have faced diabetic retinopathy at some stages of their life but half of them are aware of it. It occurs when an enormous blood vessels leak blood and other part of the eyes. This causes the retinal tissue to distend, resulting in cloudy or distinct vision. A short areas of balloon-like lump in the retinas enormous blood vessels, called microaneurysms, happen at this first stage of the disease which is called mild nonproliferative retinopathy. When the new blood vessels distract with the normal flow of fluid out of the eye, then eye ball may build up by the pressure. This can harm the nerve that bearing images from eye to brain (optic nerve), turn into glaucoma. Fresh blood vessels may turn in the front part of eye and interfere with the normal eye it can pressurize the eye to increase (glaucoma). Gradually, glaucoma, diabetic retinopathy or both diseases can lead to complete the vision loss.

The risk of blindness can be reduced by early detection and treatment. People with this disease may need eye treatment more frequently. The number of eye specialist and ophthalmologist are not enough for the people who are suffering from these diabetic retinopathy diseases. That is why Computer Aided Diagnosis (CAD) system is needed for the people [22]. The investigation of Retinal images by the researches has a long history. In 1851 first ophthalmoscope was invented by Hermann Von Hemholtz and then in 1926 a fundus camera was invented which could take a 20 degree view of the retina. Therefor, a latest Scanning laser ophthalmoscopy (SLO) technology using laser light to create high resolution images in 1926 [18] . The first technique to detect and segment unusal structures was published in 1984 [27], when Baudoin et al. described a method on image analysis to find out microaneurysms, a characteristic loss of diabetic retinopathy. In 1910, Gullstrand improved the functionalities of the fundus camera, still this concept is used today [19]. Retinal imaging has improved quickly during the last few years and is now a main part of the clinical care and management of patients with retinal type diseases . In 2009 Retinopathy Online Challenge, it was difficult to compare the performance of retinal image analysis algorithms [19]. From this challenge there is a lot of methods and idea came for the future. There are so many proposals to solve those problem. One of the proposal which is consists of (1) optic disc detection, (2) optic disc segmentation and elimination, and (3) blood vessels detection and segmentation and the methods are i) Transformation of Color Channel into Grey Scale. ii) Morphological Operation that is based on Non-uniform Background Removal. iii) Stationary Wavelet Transform for visualization of tiny Low contrast Vessels. iv) Contrast Limited Adaptive Histogram Equalisation (CLAHE) for vessel enhancement [1] .

Different people and research groups are engaged for retinal image analysis and try to improve the effectiveness of the corresponding CAD system. In [2], blood vessels are segmented on retinal fundus images. In here, Gabor filter and morphological top-hat transformation are applied after pre-processing step in order to enhance blood vessels. For accurate understanding the diseases of retinal vessels therafter digonosis of those disease a novel unsupervised mechanism is proposed in [5]. Here author proposed vessels segmentation technique using unsupervised iteretive mechanism. In 2016 a paper is published to increase the performance of segmentation of retinal vessels images [6].

With the time, different advanced techniques and methods are introduced for the natural image classification. Among them the state-of-the-art DNN has gained enormous attention after the famous work proposed by Alex Krizhevesky, in his

model AlexNet in 2012. After that, a lot of variant of this model have been introduced for the image analysis and pattern recognition purposes. Initially this kind of mechanism is introduced for the natural image classification [7] [25]. However, with fine and advanced level of tuning, this models of DNN have been introduced for different objective defined image analysis problems, such as retinal image fundus segmentation. U-Net and Generative Adversarial Network (GAN) models can be considered as one of the variant of Alexnet. In U-Net model, an encoder and decoder are used, and this kind of model are mainly utilized for image segmentation. Where as a GAN model consists a generator and discriminator which is mainly introduced introduced for synthetic data generation. Recently, this kind of techniques are also utilized for retinal image analysis. Inspired by the above advancement of recent DNN, we have introduced a variant of DNN model in this paper which is combined of U-Net and GAN model for the retinal fundus segmentation.

In this paper section I provides introductory features about the objective of this research with some state-of-the-art literature in this area. Section II provide overall classification model and section III provides a structure description about the pre-processing work of this experiment. Section IV describes our proposed model along with schematic description about U-Net and GAN model.Finally in section V discusses our finding and describes the results. In the conclusion chapter, this paper concluded our work with the future direction of further research.

II. Classification Model

Main target of this work for finding out retinal vessel from the color retinal images, this work utilized the following steps to perform the overall segmentation operation as

1) Image Pre-Processing.
2) In classification model both U-Net(which is supervised in nature) and GAN(which is unsupervised in nature) have been combinely used.
3) Performance Evaluation.



Fig. 1: Classification Model

Figure 1 shows the overall classification model.

III. Pre-Processing Steps

Pre-processing is necessary to improve the visibility of macro structure of an image. Different pre-processing steps are necessary based on the image nature and their acquisition techniques. Among different image pre-processing steps, luminance being a major concerned to the researcher as it represents the apparent brightness of an image. Another issue is contrast, which represent a significance difference between two or more things which is unclear when we compare them. It is considered that retinal images are suffers due to low luminosity and contrast issues. Considering this two issues this work emphasis to improve luminicity and contrast enhancement. As Green channel of a retinal image contains more significant information, this work mainly emphasized on Green channel for pre-processing work. In this paper we have used following two algorithms to enhance the contrast and illuminance for retinal images

1) Algorithm-1: In Algorithm-1, image I has been taken which is RGB in nature. To reduce the intensity variation an average filter has been utilized with 2×2 kernel, then we have performed the following steps sequentially such as:

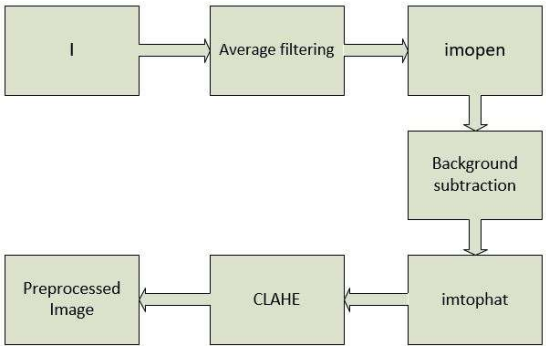


Fig. 2: Algorithm-1

- Morphological Opening: Performing erosion followed by a dilation operation is known as Morphological Opening, it can be represented as $I_{open} = (I \Theta s) \oplus d$
  - The erosion of a binary image I by a structuring elements with kernel size $e \times e$ (denoted $I \Theta s$) produces a new binary image g = $(I \Theta s)$. In the erosion method the targeted pixel $I(x, y)$ is considered to be 1 (or 0), if all the pixel of image I covered by the kernel is 1 (or 0).
  - The dilation of an image I by a structuring element d produces a new binary image g = $I \oplus d$ with ones if one pixel covered by the kernel is 1. Dilation has the opposite effect to erosion.
- Calculate the back-ground image.
- Subtraction of the image from the back-ground.
- Morphological Closing: Performing dilation followed by a erosion operation is known as Morphological Closing $(I \oplus d) \Theta s$.
- Finally we use CLAHE to enhance the contrast. Where the tiles number are 20 by 20 and clip limit is .00005.

2) Algorithm-2: Algorithm-2 performs the preprocessing operation in the LAB space instead of RGB space where the L lightness component and A and B represent the color component. Algorithm-2 is summarized as
   - Converts the image from RGB to LAB space.

- CLAHE operation is performed to improve contrast with NumTiles 8 by 8 and ClipLimit is 0.0001.
- Finally converts the LAB to RGB.

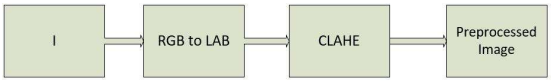


Fig. 3: Algorithm 2



(a) Original Image (b) Algorithm-1 (c) Algorithm-2

Fig. 4: Pre-processed images

Fig 4 a shows the original image where b and c represents the pre-processed image utilizing Algorithm-1 and Algorithm-2

IV. Our Proposed Model

A. Network Structure

In this experiment, we have utilized a DNN which is the combination of U-Net, GAN, Data augmentation technique .We have described the procedure as follow :

1) Generative Adversarial Network (GAN): GAN [15] is a recent advancement of Deep Neural Network, which is normally used for unsupervised machine learning. It consists of two neural networks one is generator and the other is discriminator.

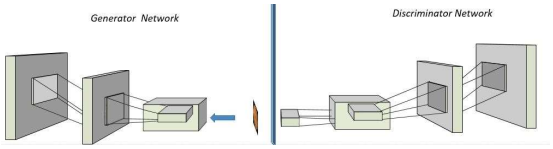


Fig. 5: Generative Adversarial Network Architecture

The generator tries to make new data following a particular probability distribution and passes this the data through the discriminator. The generator makes the data in such a way that it looks like a real data though those are fake or artificial. The target of the generator is to generate images that look like the real images. When the generate data pass through the discriminator the target of the discriminator to recognize those information that are not real. The objective function can be found as follows:

1) Objective Function:: The generator model is written as, $\mathscr{G}(z \rightarrow y; \mathbf{w_g})$, where a given noises distribution z is transformed to an image y. Normally for an image generation, the distribution is considered as Bernoulli distribution. The discriminator is represented as $\mathscr{D}(x \rightarrow p; \mathbf{w_d})$, here x represents the provided trained image and p represents its corresponding probability distribution which lies in the range from 0 to 1. Here $\mathbf{w_g}$ and $\mathbf{w_d}$ represent the weight for generator and discriminator. Discriminator normally works in a two-fold manner such as:

- Firstly it maximizes its weight such that input image x is classified as real data. That means, the objective is maximizing $\mathscr{D}(x)$ [8]
- Secondly, the discriminator tries to reduce the probability of miss-classification of any fake image to a real image. That means the objective is minimizing $\mathscr{D}(\mathscr{G}(z))$

Considering the above issues, the over all loss function [22], [23] $\mathcal{L}$ for GAN can be written as

$$\mathcal{L} = \min_{\mathscr{G}} \max_{\mathscr{D}} \{\mathbb{E}_{x \sim p_{data}}[\log(\mathscr{D})] + \mathbb{E}_{z \sim p_z}[\log(1 - \mathscr{D}(\mathscr{G}))]\}$$

In this experiment, the generator takes a fundus image and generates retinal segmented vessel as the same size of the input [12]. Then the discriminator classify the image weather the image is real (taken from the training set) or fake (created by the generator). The discriminator creates labels for each ingredient in the batch. The loss function are calculated risen from the labels. The images are directly pass through the discriminator because the fake image created by the generator is the part of the same computational graph. In the generator we use U-Net model for segment the vessel image [10].

2) U-Net: U-Net utilize encoder decoder concept for semantic segmentation of an image where encoder utilize downsample concept and decoder utilize up sample concept in a down sampling process few convolution blocks followed by maxpooling operation as given in Figure 5 .

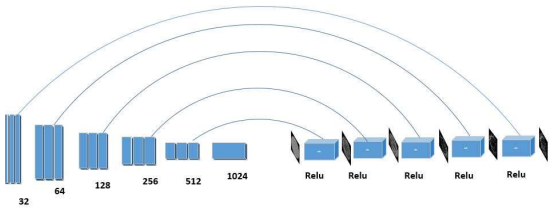


Fig. 6: U-Net model structure

U-Net model has two parts one is encoder and the other is decoder. In this experiment the first layer of the encoder take input image with size ($640 \times 640 \times 1$) and produce output of size ($320 \times 320 \times 64$) which is work as input of the next two parallel layer [9]. The two parallel

layers produce two outputs one is (160×160 × 64) and (80×80×128). Here first parallel output goes to the third layer as input. The 2nd parallel layer concat with the 4th layer and the concat output goes to the 5th layer. In the downsampling there use relu as activation function .

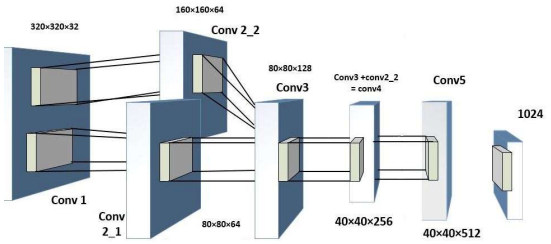


Fig. 7: GAN model generator downsample

After decoding the encoder the segmented image comes by using transpose convolution operation. In encoding process the original information come to in a compressed from. In here as a result of compression the original image loss his information. To increase the resolution with high localization properly an encoder model is concatated after the decoder part. The decoder part continuously decompressed the compressed information utilizing successive layers mention that in this particular procedure the pooling operation is replaced by upsampling operation.

B. Augmentation

This experiment utilizes public database DRIVE as this database contains very small amount of images this experiment created few synthetic images utilizing augmentation method such as :

- left right filp [11]
- 0 to 360 degree rotation with 3 degree incrementation
- z score

C. DATA set description

DRIVE [Digital Retinal Images for Vessel Extraction] [16] database was published for relative studies on segmentation on blood vessel in retinal images. This database is free for educational and research purposes. The images that set for DRIVE database were come through a diabetic retinopathy screening program in the Netherlands. In this program there were 400 people age between 25 to 29 for collecting images. From these images randomly 40 images were selected for the DRIVE database. These 40 images were divided into two group one is training and other is testing where both groups contained 20 images. For the training images a single manual segmentation and for the test two manual segmentations are available. In the case of test images one manual is gold standard and the other is the computer based segmentation. All of the images in the DRIVE database is used for making clinical diagnoses.

D. Package dependency

This experiment utilize a computer with with following configuration

- Processor Intel(R) Core(TM)i5-6200 CPU 2.3 GHz
- Memory 8 GB
- 64-bit operating system in this experiment training lasts for 35 hours.As a software platform here use Tensorflow

E. Result

There are some benchmark criteria to evaluate supervise machine learning technique. The Area Under Curve for Receiver Operating Characteristic (ROC AUC), [13] [28] Area Under Curve for Precision and Recall Curve (PR AUC) and dice coefficient and F1 measure are evaluated in this experiment.

- Receiver Operating Characteristic Curve ROC curve is a graphical visualization that represent the diagnostic ability of a binary classifier system. ROC is a probability curve and AUC is the amount of separability. AUC represent the model that how much it is suitable for distinguishing between classes. Depending on various threshold values a AUC-ROC curve is created to measure the performance of a classification problem. Sensitivity and Specificity are inversely proportional if sensitivity is increased, the specificity decreases. When the threshold decreases, more positive values produce for this reason the sensitivity increases and the specificity decrease.
  Every point on the curve represents a sensitivity/specificity pair. The area under the ROC curve (AUC) is a measurement of how the parameter can distinguish between sensitivity and specificity. In this experiment we utilize two algorithms where Algorithm-1 obtain 90.26 % ROC value and Algorithm-2 obtain 89.91 % ROC value. The sensitivity for the both algorithms 0.630343% and 0.632588% respectively. The specificity for the both first and second algorithms 0.965861% and 0.959156% respectively.

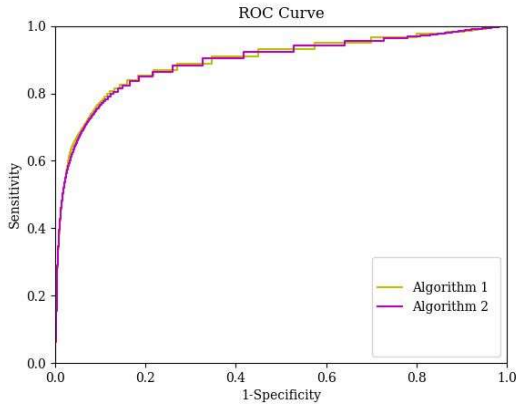


Fig. 8: Receiver Operating Characteristic Curve

- Precision-Recall Curve
  The precision-recall curve represent the balance between two desirable but incompatible features. Recall is the

measurement of the performance of the positive part of the information. Precision is the measurement of the performance of positive prediction. In the precision and recall calculation sensitivity is identical with recall and the positive prediction value is identical with precision. The precision-recall curve show that the performance of the model on different data of the image that is use in the experiment. In the experiment for Algorithm-1 PR value is 0.714784% and for Algorithm-2 the PR value is 0.701368%.

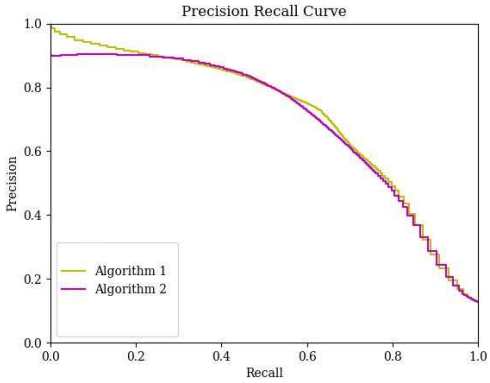


Fig. 9: Pecision-Recall Curve

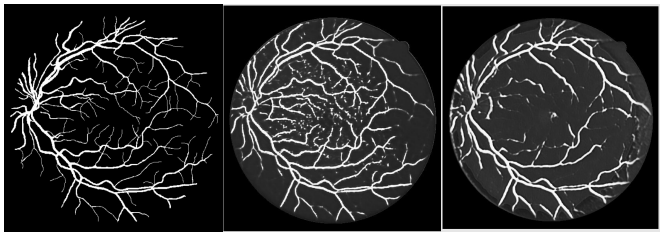
- Dice Score
  Dice score is mostly used to quantify the execution of image segmentation process. In the images there are ground truth region and from the images the dice algorithm calculate the dice score. This is the measurement of the similarity of the images. So, dice score is the size of overlapped of two segmented images. In the DRIVE data there are mask with ground truth.The dice algorithm generate mask/image which is binary that creates a mask for the segmentation . 67.61% and 66.14% dice score value for Algorithm-1 and Algorithm-2 respectively.
- F1 Score
  F1 is a function of precision and recall. F1 score brings the balance between precision and recall. In the case of accuracy the true negative takes for calculation but the F1 score takes the false negative and false positive both for calculation thus the F1 score may be better to bring balance between precision and recall. In the experiment Algorithm-1 has 64.50% F1 score and Algorithm-2 has 66.15% F1 score.
- Accuracy
  Accuracy is one parameter for evaluating classification models. Accuracy is the parameter for predicting that how much the model got right.

The following table summaries all of our findings

TABLE I: Result Summaries

Result	Algorithm 1	Algorithm 2
PR	0.7147	0.7013
ROC	0.9026	0.8991
Dice coefficient	0.6761	0.6614
Accuracy	0.9231	0.9175
Sensitivity	0.6303	0.6325
Specificity	0.9658	0.9591
F1 score	0.6450	0.6615

In Fig 10 a represents original segmented image, b and c represent segmented image for Algorithm-1 and Algorithm-2



(a) Original (b) Algorithm-1 (c) Algorithm-2

Fig. 10: Segmented Images

There are few works have been done on this dataset. M. Li et al. performed Multi-Scale Convolutional Neural Network and obtained 91.7% [29] accuracy. H.Fu et al. perform CNN and Conditional Random Field method and obtained 91 to 94% [30] accuracy. These two results and nearly comparable with our findings.

V. Conclusion

Retinal Fundus image identification is always being an active area of research. With the time being new methods have been introduced and adapted for the retinal image identification and classification. This paper has classified and segmented the fundus image using state-of-the-art DNN, specifically utilized both the supervised U-Net model and an unsupervised GAN model together for the retinal image segmentation. One of the main contribution of this paper is, in the decoder of the U-Net model, a parallel layer has been utilized. The overall achieved result of this model is moderated, however with the fine tuning and adjustment of the parameters, the overall performance can be improved.

VI. Acknowledgement

We have took help for creating the simulation environment from that site [17].

References

[1] Toufique A. Soomro, Junbin Gao, Mohammad A. U. Khan, Tariq M. Khan, Manoranjan Paul and Nighat Mir. "Role of Image Contrast Enhancement Technique for Ophthalmologist as a Diagnostic Tool for the Diabetic Retinopathy". 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA),Publisher: IEEE.

[2] Zafer Yavuz Bilgisayar Mhendislii Blm, Karadeniz Teknik niversitesi, Trabzon, Trkiye. "Blood vessel segmentation from retinal images based on enhencement methods" Cemal Kse. 2014 22nd Signal Processing and Communications Applications Conference (SIU),Publisher: IEEE.Date Added to IEEE Xplore: 12 June 2014

[3] Yitian Zhao, Yonghuai Liu, Xiangqian Wu, Simon P. Harding, Yalin ZhengRetinal. "Vessel Segmentation: An Efficient Graph Cut Approach with Retinex and Local Phase". April 1, 2015

[4] Wendeson S. Oliveira Joyce Vitor Teixeira, Tsang Ing Ren, George D. C. Cavalcanti, Jan Sijbers."Unsupervised Retinal Vessel Segmentation Using Combined Filters". February 26, 2016

[5] Sohini Roychowdhury ; Dara D. Koozekanani ; Keshab K. Parhi. "Iterative Vessel Segmentation of Fundus Images". IEEE Transactions on Biomedical Engineering (Volume: 62 , Issue: 7 , July 2015)

[6] These lecture notes follow Chapter 11 "Morphological image processing" of the textbook Nick Efford. Digital Image Processing: A Practical Introduction Using JavaTM. Pearson Education, 2000.

[7] Abdullah-Al Nahid, Mohamad Ali Mehrabi, and Yinan Kong, Histopathological Breast Cancer Image Classification by Deep Neural Network Techniques Guided by Local Clustering, BioMed Research International, vol. 2018, Article ID 2362108, 20 pages, 2018

[8] These lecture notes follow Chapter 11 "Morphological image processing" of the textbook Nick Efford. Digital Image Processing: A Practical Introduction Using JavaTM. Pearson Education, 2000. Alexnet[7] https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[9] Maayan Frid-Adar, Idit Diamant, Eyal Klang, Michal Amitai, Jacob Goldberger, Hayit Greenspan;GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in Liver Lesion Classification Neurocomputing (IF 3.241) Pub Date : 2018-09-21 , DOI: 10.1016/j.neucom.2018.09.013

[10] Liming Zhao, Jingdong Wang, Xi Li, Zhuowen Tu, Wenjun Zeng."Deep Convolutional Neural Networks with Merge-and-Run Mappings". (Submitted on 23 Nov 2016 (v1), last revised 19 Jul 2017 (this version, v2))

[11] Olaf Ronneberger, Philipp Fischer, Thomas Brox."U-Net: Convolutional Networks for Biomedical Image Segmentation". (Submitted on 18 May 2015)

[12] Jia Shijie ; Wang Ping ; Jia Peiyi ; Hu Siping."Research on data augmentation for image classification based on convolution neural networks". 2017 Chinese Automation Congress (CAC) Date of Conference: 20-22 Oct. 2017 Date Added to IEEE Xplore: 01 January 2018

[13] Abdullah-Al Nahid and Ferdous bin Ali and Yinan Kong, "Histopathological breast-image classification with image enhancement by convolutional neural network", 2017 20th International Conference of Computer and Information Technology (ICCIT)

[14] Jaemin Son, Sang Jun Park, Kyu-Hwan Jung."Retinal Vessel Segmentation in Fundoscopic Images with Generative Adversarial Networks" (Submitted on 28 Jun 2017)

[15] IanJ.Goodfellow, JeanPouget-Abadie, MehdiMirza, BingXu, DavidWarde-Farley, SherjilOzair, AaronCourville, YoshuaBengio Departement dinformatique et de recherche operationnelle Universite de Montreal Montreal, QC H3C 3J7."GenerativeAdversarialNets"

[16] "https://computervisiononline.com/dataset/1105138662"

[17] "https://github.com/ChengBinJin/V-GAN-tensorflow"

[18] A History of Retinal Imaging Posted on December 30, 2014 by Optos "https://blog.optos.com/index.php/a-history-of-retinal-imaging"

[19] Niemeijer M, van Ginneken B, Cree M, Mizutani A, Quellec G, Sanchez C, Zhang B, Hornero R, Lamard M, Muramatsu C, Wu X, Cazuguel G, You J, Mayo A, Li Q, Hatanaka Y, Cochener B, Roux C, Karray F, Garcia M, Fujita H, Abramoff MD.Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs. IEEE Trans Med Imaging. 2010 Jan;29(1):185195.[PubMed]." https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3131209/#R97"

[20] Gullstrand A. Neue methoden der reflexlosen ophthalmoskopie. Berichte Deutsche Ophthalmologische Gesellschaft. 1910;36. " https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3131209/#R32"

[21] Abdullah-Al Nahid and Yinan Kong, Involvement of Machine Learning for Breast Cancer Image Classification: A Survey, Computational and Mathematical Methods in Medicine, vol. 2017, Article ID 3781951, 29 pages, 2017

[22] Y. Liu, Q. Zhao and C. Jiang, "Conditional image generation using feature-matching GAN," 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Shanghai, 2017, pp. 1-5. doi: 10.1109/CISP-BMEI.2017.8302049 . http://ieeexplore.ieee.org/stamp/stamp.jsp?

[23] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros "Image-to-Image Translation with Conditional Adversarial Networks". (Submitted on 21 Nov 2016 (v1), last revised 22 Nov 2017 (this version, v2))

[24] Dr. Kuldeep Solanki, Dr. Jayant Kumar, Dr. ND Soni, Dr. Raguveer Chaudhary and Dr. Gautam Chand Sirvi. "Study the incidence of diabetic retinopathy with relation to duration of diabetes mellitus". http://www.allresearchjournal.com/archives/2016/vol2issue8/PartF/2-8-75-648.pdf

[25] Abdullah-Al Nahid and Yinan Kong, "Histopathological Breast-Image Classification Using Local and Frequency Domains by Convolutional Neural Network." Information 9, no. 1: 19, 2018.

[26] Dr Aisha Mohammed Bello Ahmed1 , Tarig Eltoum Yagoub2 , Hajir M..K. Alla3 , Nabbila Omer Musa4 & Muda A. Ali5 1 Assistant Professor, Head Department of Physiology, Faculty of Basic Medical Sciences, Omdurman Islamic University. 2 MBBS, MSc, MD. Dep. of physiol. Faculty of Medicine, Alzaiem Alazhari University, "Causes and Risk Factors Contributed and Responsible for Prevalence and Incidence of Retinopathy in Al- Walidain Hospital in Omdurman Sudan"

[27] Baudoin CE, Lay BJ, Klein JC. Automatic detection of microaneurysms in diabetic fluorescein angiography. Rev Epidemiol Sante Publique. 1984;32(34):254261. [Online]. Available: PM:0006522738. [PubMed]

[28] Abdullah-Al Nahid and Yinan Kong, "Local and Global Feature Utilization for Breast Image Classification by Convolutional Neural Network", 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA).

[29] Ming Li , Qingbo Yin, Mingyu Lu "Retinal Blood Vessel Segmentation Based on Multi-Scale Deep Learning", Computer Science and Information Systems pp. 117123 DOI: 10.15439/2018F127 ISSN 2300-5963 ACSIS, Vol. 15

[30] Huazhu Fu1(B), Yanwu Xu1, Stephen Lin2, Damon Wing Kee Wong1, and Jiang Liu1, "DeepVessel: Retinal Vessel Segmentation via Deep Learning and Conditional Random Field " , Springer International Publishing AG 2016 S. Ourselin et al. (Eds.): MICCAI 2016, Part II, LNCS 9901, pp. 132139, 2016. DOI: 10.1007/978-3-319-46723-8 16

Virtual Image from EEG to Recognize Appropriate Emotion using Convolutional Neural Network

Md. Rabiul Islam
Electrical and Electronic Engineering
Khulna University of Engineering & Technology
Khulna -9203, Bangladesh
rabiul00000@gmail.com

Mohiuddin Ahmad
Electrical and Electronic Engineering
Khulna University of Engineering & Technology
Khulna - 9203, Bangladesh
mohiuddin.ahmad@gmail.com

Abstract— **Electroencephalogram (EEG) that measures the electrical activity of the brain has been used extensively to recognize emotion. Normally feature based emotion recognition requires a strong effort to design the perfect feature or feature set related to the classification of emotion. To curtail the manual human effort we designed a model by using a virtual image from EEG with Convolutional Neural Network (CNN). Initially, we calculated Pearson's correlation coefficients form different sub-bands of EEG to formulate a virtual image. Later, this virtual image was fed into a CNN architecture to classify emotion. We made two distinct protocols; between these, protocol-1 was to classify positive and negative emotion and protocol-2 was to classify four distinct emotions. An overall maximum accuracy of 81.51% on valence and 79.42% on arousal was obtained by using internationally authorized DEAP dataset. Our proposed method is helpful in recognizing emotions efficiently.**

Keywords—EEG, Convolutional Neural Network, emotion, correlation, virtual image.

I. INTRODUCTION

Simply, emotion means the instant mental state and interaction. The degree of pleasure, displeasure, excitation, sadness, etc. can easily be extracted from emotion. However, emotion is not so simple always because it varies with respect to personality, motivation, mood, temperament, event and with the different environment of surroundings. In spite of having these type of wide variety and complexity, emotion has such unavoidable importance that attracts the researcher significantly from the last two decades.

Now a day's digital technology changes the environment very rapidly. But the improvement of different human interacted media depends largely with the appropriate recognition of the human feeling that means what he or she wants just right at that moment. To communicate computer with human, to detect present feelings of the patient, to understand customer interest for shopping, to realize interest of physically handicapped people or to detect the lie of a convict; emotion recognition is a fundamental prerequisite. Due to these inevitable importance's, emotion recognition has great importance in the field of neuroscience, computer science, cognitive science, and physiology.

Multitudinous approaches have been applied to extract emotion like facial expressions [1], gesture and speech signals, EEG signals, autonomous nervous signals, etc. For facial expression and speech signals, subjects need to express emotion explicitly. On the other hand, subjects have no way to control the automatically generated EEG signal. Therefore, EEG is the best and inimitable signal for emotion recognition.

Recently, deep learning has become an indispensable part in the field of artificial intelligence to process signals and information. Numerous deep learning architecture like Deep Belief Network (DBN), Convolutional Neural Network (CNN), Deep auto-encoder, Generative Adversarial Network (GAN) achieve outstanding result compared to shallow models like SVM, MLP, k-NN, CRF in many challenging fields like pattern recognition and image processing. In modern time for analyzing physiological signal (i.e. EEG, EOG, ECG, EMG) various deep learning-based architectures are applied successfully and attained comparable result [2].

In this paper, we used CNN to classify emotion. Firstly, we decomposed the EEG signal into 5 sub-bands namely delta ($1 \text{ Hz} < f < 4 \text{ Hz}$), theta ($4 \text{ Hz} < f < 8 \text{ Hz}$), alpha ($8 \text{ Hz} < f < 12 \text{ Hz}$), beta ($12 \text{ Hz} < f < 30 \text{ Hz}$) and gamma ($30 \text{ Hz} < f < 60 \text{ Hz}$). Since alpha, beta and gamma subbands are more relevant with emotion [3], we considered only these three sub-bands to formulate virtual image by calculating Person's Correlation Coefficients (PCC). Then the virtual images were set as the input of CNN based classification method.

The plan of the remaining paper is just like that. The background research is described in section II. Section III illustrates the methodology. The CNN model is explained in section IV. The result is reported in section V. At last, section VI terminates the paper.

II. RELEVANT RESEARCH

There have multiple proposed approaches in the field of emotion classification. Primarily the approaches can be classified into two types a) feature-based approaches, ii) deep architecture-based approaches. The feature-based approaches consist of the feature in the time domain, frequency domain, joint time-frequency domain, etc. The Hjorth index [4], simple statistics, non-stationary index, etc. are calculated for the time domain.

In the frequency domain, the popular methods are Power Spectral Density (PSD), Short Time Fourier Transform (STFT) [5], etc. The joint time-frequency domain consists of wavelet-based features [6].

Moreover, with these existing methods of emotion recognition, some scholars have started to use deep learning architecture. Martinez et al. used the skin conductance and blood volume pulse signal with an effective convolutional neural network to classify four emotional states [7]. Li et al. worked using differential entropy with DBN and compared with five baselines improvement of 11.5% and 24.4% [8].

Here, we used the CNN model upon a virtual image formed by Pearson's correlation coefficients and achieved the highest accuracy described in the result section.

III. METHODOLOGY

A. Database description

We made our experiment using DEAP dataset [3]. This dataset of emotion recognition consists of two parts, one is the ratings from an online self-assessment and other is the participant ratings, physiological recording and face video of experiments. The summary of the dataset is given in Table I.

TABLE I: THE SUMMARY OF DEAP DATASET

No. of participants	32 (16 female)
No. of video	40 (1 min long)
No. of distinct emotion	15
Rating scale	Valence, arousal, dominance, liking, familiarity
Rating value	1-9, except familiarity (1-5)
Recorded signals	32- channel 512 Hz EEG; physiological signals; Face video (no. 22)
EEG acquisition medium	Biosemi ActiveTwo

For collecting EEG data, they used the international 10/20 electrode positioning system shown in Fig. 1.

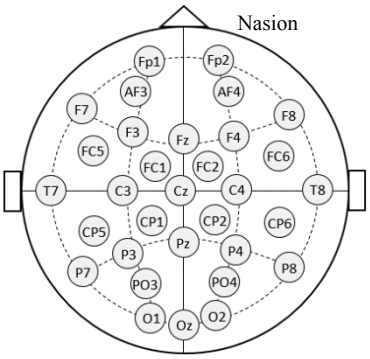


Fig. 1: International 10/20 electrode placement system [3] showing only these 32 electrodes used by 'DEAP'.

The name of 32 channels or electrode associated with 'DEAP' are Fp1, F3, FC1, AF3, F7, FC5, T7, C3, CP1, CP5, O1, Oz, O2, P7, P3, Pz, PO3, PO4, P4, P8, CP6, CP2, C4, T8, F4, F8, AF4, Fp2, FC6, FC2, Cz and Fz. The available data is 3d in the format like video×channel×data = 40×40×8064.

B. Overall experiment

The following block diagram of Fig. 2. represents the overall proposed method.

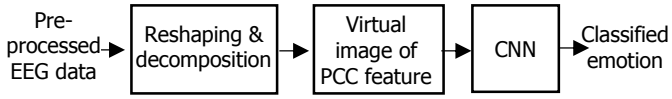


Fig. 2: The simplified block diagram of proposed method.

a) Pre-processing: The recorded raw EEG data of 512 Hz were down-sampled firstly at 128 Hz. Later to remove undesirable upper frequency a (4-45) Hz bandpass filter was applied. The EOG artifact also removed. The preprocessed data were downloaded from the DEAP website.

b) Reshaping and decomposition: Among the overall 40 channels data, we took only 32 channels of EEG data. Then we removed the first three-second pre-trial baseline data. Then the output EEG signal is decomposed into five frequency sub-bands.

c) Formation of a virtual image: EEG is a one-dimensional non-stationary signal. Nevertheless, for classification by using Convolutional Neural Network architecture we need two-dimensional image type data. Here we calculated the PCC between every pair of 32 EEG channels. Since delta and theta sub-bands are related to deep sleep and subconscious mind, we considered the rest three sub-bands for calculating PCC. Thus, we formulated a 32×32×3 shaped three-dimensional data equivalent to a color image, which was termed as a virtual image. Then this image was taken as an input of the CNN model.

C. Pearson's Correlation Coefficient (PCC)

For emotion recognition, many scholars used numerous features like Power Spectral Density (PSD), Differential Entropy (DE), Differential Asymmetry (DASM), Rational Asymmetry (RASM), Wavelet Energy (WE), Relative Wavelet Energy (RWE), etc. We used Pearson's Correlation Coefficient (PCC) of a similar and consecutive point to build a matrix of cross-correlation.

Pearson's Correlation Coefficient or bivariate correlation is a measure of linear correlation between two one-dimensional variables. Mathematically, it is the ratio of covariance of two arrays to the product of their standard deviation. Its value ranges from -1 to +1. Where, +1 indicates 100% positive correlation, 0 indicates no correlation, -1 means 100% correlation but anti or negative correlation.

Suppose that, two different one-dimensional variables are x and y, where $x=(x_1, x_2, ..., x_n)$ and $y=(y_1, y_2, ..., y_n)$. Then the correlation between these two data sequences will be

$$\rho_{xy} = \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y} \tag{1}$$

Where, ρ_{xy} and $cov(x,y)$ represents the PCC and covariance respectively between x and y, σ indicates the standard deviation. However, In details, the equation can be written as follows.

$$\rho_{xy} = \frac{n\sum_{i=1}^{n}(x_i y_i) - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}} \tag{2}$$

D. Experiment Protocol

For a subject, for a single trial video and for a single channel there remain 8064 (63s ×128Hz) data points. As CNN required a large amount of dataset for training, to operate successfully we segmented the data to increase the quantity of virtual image. For doing this after removing the data of first 3 seconds we took only the last 60s data and segmented it into 20 samples each contains 384 data points of 3 seconds. Thus for only one participant, we got 800 samples. Thus we had 25,600 samples of the virtual image in total for all subject. In our work, we have made 2 different protocol.

a) Protocol-1: Here, we just classified two discrete emotions positive and negative. For each participant, we considered the valence and arousal rating. From the rating

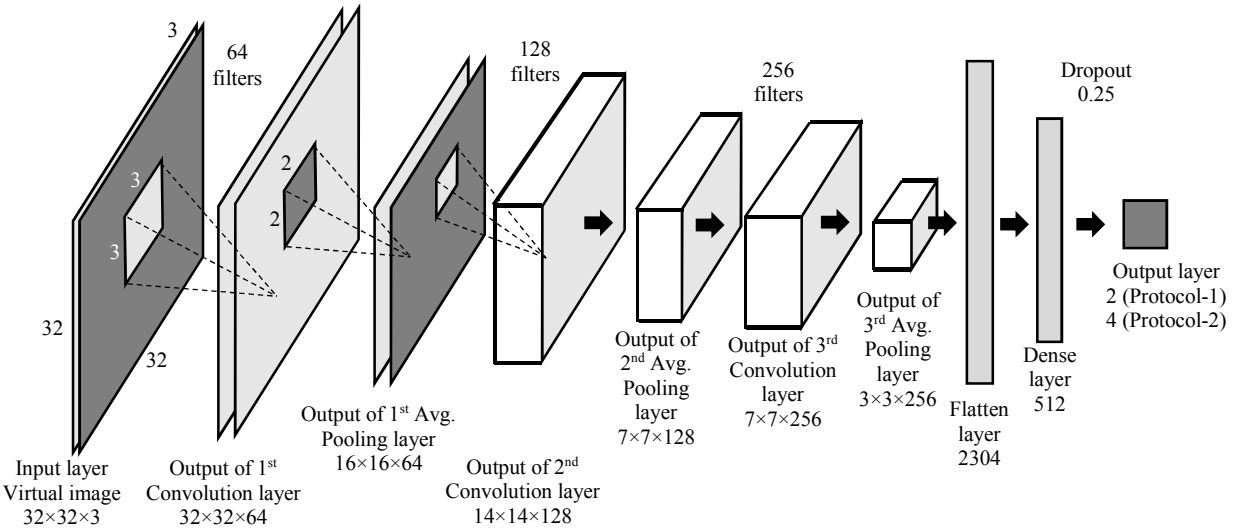


Fig. 3: Proposed CNN model

our model predicts 1 to less than 5 as low and from 5 to 9 as high. We classified as low as negative and high as a positive emotion.

b) Protocol-2: In this section, we classified 4 discrete emotion namely happy, angry, sad and relaxed by using a low and high value of valance and arousal. Here the high and low scaling is the same as the previous protocol. But here we considered low valence low arousal as sad, low valence high arousal as angry, high valence low arousal as relaxed and high valence high arousal as a happy emotional class using Russell's circumplex model.

IV. Convolutional Neural Network Model

Convolutional Neural Network or simply ConvNet is an effective model, which is able to assign proper importance (weights and biases) depending on numerous aspects of the input image and be able to differentiate one from other. A simple CNN architecture consists of a set of a convolutional layer, ReLu layer, and pooling layer. A complete CNN also contains Fully Connected (FC) layer like flatten and dense layer. Our overall model from the virtual image to classified emotion is shown in Fig. 3.

A. Convolution Layer

This layer is the core building block of CNN architecture, which make convolution operation between the input image and the applied filter. We used 64, 128 and 256 filters of size 3×3 and of stride 1 in the 1st, 2nd, and 3rd convolutional layer respectively. The output of CNN layer was then passed through an activation function. Simply, according to pure mathematics, the convolution operation in between two signals *h(t)* and *k(t)* can be expressed as follows.

$$(h * k)(t) = \int_{-\infty}^{\infty} h(t-\tau)k(\tau)d\tau \tag{3}$$

Here, for any single layer of convolution if the size of the input image is $n_H^{[l-1]} \times n_W^{[l-1]} \times n_C^{[l-1]}$ and the size of the filter is $f_H^{[l]} \times f_W^{[l]} \times n_C^{[l-1]}$ then the dimension of the output image after completing a single CNN layer will be $n_H^{[l]} \times n_W^{[l]} \times n_C^{[l]}$ where,

$$n_H = \frac{n_H^{[l-1]} + 2p^{[l]} - f^{[l]}}{S^{[l]}} + 1 \tag{4}$$

The calculation of n_W is similar to n_H. Where $p^{[l]}$ indicates the dimension of padding and $S^{[l]}$ indicates the value of stride at level *l*.

B. ReLu layer

The activation function decides if the neuron would fire or not. There remains different activation function like Sigmoid, hyperbolic tangent (thah), Rectified Linear Unit (ReLu), Maxout, Exponential Linear Unit (ELU), softmax, softsign, Scaled Exponential Linear Unit (SELU) and softplus. Here we used ReLu as activation function which is the best fit with CNN architecture and faster.

$$F(x)=\max(0, x)$$

Lastly we used 'softmax' activation function to classify emotion.

C. Pooling layer

We used average pooling in spite of using normally used max pooling. This is because PCC based virtual image is a texture type image, not like an ordinary image. Here we used a 2×2 matrix with stride 1 for pooling.

TABLE II: The Summary of Applied CNN Model

Layer (type)	Output Shape	Parameter
batch_normalization_v1	32, 32, 3	12
conv2d (Conv2D)	32, 32, 64	1792
max_pooling2d(MaxPooling2D)	16, 16, 64	0
atch_normalization_v1_1	16, 16, 64	256
conv2d_1 (Conv2D)	14, 14, 128	73856
max_pooling2d_1(MaxPooling2D)	7, 7, 128	0
batch_normalization_v1_2	7, 7, 128	512
conv2d_2 (Conv2D)	7, 7, 256	295168
max_pooling2d_2 (MaxPooling2D	3, 3, 256	0
flatten (Flatten)	2304	0
dense (Dense)	512	1180160
activation (Activation)	512	0
dropout (Dropout)	512	0
dense_1 (Dense)	4	2052
activation_1 (Activation)	4	0

After using three sets of convolutions, ReLu, and pooling layers, we used two fully connected layers. The summary of our CNN architecture is given in Table II.

D. Optimizer and Loss function

In machine learning, the commonly implemented optimizers are Adamax, Nadam, Stochastic Gradient Descent (SGD), RMSprop, Adagrad, Adadelta, Adam. 'Adam' can be thought of as a combination of 'RMSprop' and 'SGD' with momentum. It has adaptive learning rates, which can compute the individual rate of learning for a different parameter. Hence, we used 'Adam' optimizer in our experiment. Cross entropy was applied as the loss function in this network.

The training and testing data were split as 80% and 20% respectively. The training epoch and batch size were 50 and 20. We chose the initial learning rate of 0.001 and dropout 0.25. The no of the total parameter is 1,553,808 and the trainable parameter is 1,553,418.

V. Results & Discussions

In our work, one of the main objectives was to convert the one-dimensional signal into a significant 2-dimension signal. By making a virtual image of PCC, we ensured that it is possible to use CNN to classify distinct cognitive state from the EEG signal. The complete experiment was done by maintaining two different protocol. The model output result is shown in Table III in terms of accuracy.

TABLE III: The classification accuracy of our model

Protocol	No. of emotion	Accuracy (Mean ± Std.)	
		Valence	Arousal
Protocol-1	2	81.51±10.7	79.42±11.8
Protocol-2	4	71.67±11.1	70.18±13.5

Hence, it can be notified that our CNN model outperforms in protocol-1 and reached in the state-of-the-art performance to classify positive and negative emotion.

Actually, the proper comparison in this type of research is not so meaningful due to a wide variety of input data, no. of EEG channel, no. of emotion classified. However, a comparison is presented in Table IV.

TABLE IV: The Comparison of result with accuracy

Method		No. of emotion	Accuracy (%)	
			Valance	Arousal
Bahari et al. 2013		3	58.05	64.56
Zhuang et al. 2014 [9]		4	70.90	67.10
Jirayucharoensak et al. 2014 [10]		9	49.52	46.03
Tripathi et al. 2017 [11]		2	81.41	73.36
Our Method	Protocol-1	2	81.51	79.42
	Protocol-2	4	71.67	70.18

The classification accuracy is largely dependent on the no. of output class. The more the no. of emotion have to be classified the less the classification accuracy would be found. However, in this work, our contribution is to propose a CNN based method to recognize emotion from multichannel EEG data, which reduce the trouble of human to manually extract the significant feature.

VI. Conclusions

In this research work, a Convolutional Neural Network was implemented on a virtual image data to classify emotion. The EEG data was firstly reshaped and decomposed and then PCC was calculated to build a featured virtual image. Our achievement was good, but further improvement will be possible by taking a large set of data for training the model. But undoubtedly, our experiment provides the important information that Convolutional Neural Network with significant input data is a very effective and potential network in the research domain associated with human brain signal.

References

[1] M. Soleymani, S. Asghari-Esfeden, Y. Fu and M. Pantic, "Analysis of EEG signals and facial expressions for continuous emotion detection," IEEE Transactions on Affective Computing, vol. 7, no. 1, pp.17-28, Jan. 2016.

[2] W. L. Zheng and B. L. Lu "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," IEEE Transactions on Autonomous Mental Development, vol. 7, no. 3, pp.162-175, 2015.

[3] S. Koelstra, C. Muhl, M. Soleymani, S. J. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt and I. Patras, "Deap: A database for emotion analysis; using physiological signals," IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18-31, 2012.

[4] S. H. Oh, Y. R. Lee and H. N. Kim, "A novel EEG feature extraction method using Hjorth parameter," International Journal of Electronics and Electrical Engineering, vol. 2, no. 2, pp. 106-110, 2014.

[5] G. Lee, M. Kwon, S.K. Sri and M. Lee, "Emotion recognition based on 3D fuzzy visual and EEG features in movie clips," Neurocomputing 144, pp. 560-568, 2014.

[6] M. R. Islam and M. Ahmad, "Wavelet Analysis Based Classification of Emotion from EEG Signal," In 2nd International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox's Bazar, Bangladesh, February 2019.

[7] H.P. Martinez, Y. Bengio and G.N. Yannakakis, "Learning deep physiological models of effect," IEEE Computational intelligence magazine, vol. 8, no. 2, pp.20-33, 2013.

[8] K. Li, X. Li, Y. Zhang and A. Zhang, " Affective state recognition from EEG with deep belief networks", In International Conference on Bioinformatics and Biomedicine, pp. 305-310, IEEE, December 2013.

[9] X. Zhuang, V. Rozgić and M. Crystal, "Compact unsupervised EEG response representation for emotion recognition," In IEEE-EMBS international conference on Biomedical and Health Informatics (BHI), pp. 736-739, IEEE, June 2014.

[10] S. Jirayucharoensak, S. Pan-Ngum, and P. Israsena, "EEG-based emotion recognition using deep learning network with principal component-based covariate shift adaptation," The Scientific World Journal, 2014.

[11] S. Tripathi, S. Acharya, R.D. Sharma, S. Mittal, and S. Bhattacharya, "Using Deep and Convolutional Neural Networks for Accurate Emotion Classification on DEAP Dataset" In Twenty-Ninth IAAI Conference, February 2011.

EduNet: A Deep Neural Network Approach for Predicting CGPA of Undergraduate Students

Md Rifatul Islam Rifat*, Abdullah Al Imran†, A. S. M. Badrudduza‡
*,‡Dept. of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh
†Dept. of Computer Science & Engineering
American International University-Bangladesh, Dhaka, Bangladesh
Email: {*irifat.ruet, †abdalimran, ‡asmb.kanon}@gmail.com

***Abstract*—Educational Data Mining (EDM) is an emerging research field concerned with the application of data mining, machine learning, and statistics in the discipline of education. Many researchers have already focused on EDM and exploring the educational data using several traditional data mining techniques to improve the educational performance of the students by extracting the concealed patterns and predicting the final outcome. In this study, we aim to propose a Deep Neural Network (DNN) based model to predict the final CGPA of the undergraduate business students with a minimal error than the traditional approaches. We have considered the performance of a decision tree model as the baseline performance. Experiments in this study have shown that our proposed DNN model can predict the CGPA with a significantly minimal error rate. To measure the performance of our model we have considered the three evaluation metrics namely Mean Squared Error (=0.008), Mean Absolute Error (=0.067), and Mean Absolute Percentage Error (=2.074). Our proposed model has successfully shown a promising prediction performance by reducing the MSE, MAE, and MAPE by 0.0146, 0.0431, and 6.043 respectively, compared to the baseline model.**

***Index Terms*—Deep Neural Networks, Regression, Educational Data Mining, CGPA Prediction.**

I. Introduction

In the modern era, education is the most important factor to change the world and thus many researchers have emphasized on exploring the educational data and trying to improve the quality of the education, performance of the students, and trying to build up an ideal educational system [1]. Moreover, the educational institutions that are providing higher education have already taken this as a challenge.

As an educational system generates large-scale data within a short period of time, the researchers have focused on the data mining techniques to analyze the educational data. Basically, data mining is an information extraction activity that identifies the concealed patterns in the data [2]. Nowadays, each institution aims at producing higher and exemplary education rates by employing various data mining methods.

The increasing emphasis of the researchers on data mining and education system has already made a new research field named Educational Data Mining (EDM). EDM is an emerging discipline concerned with developing methods for exploring educational data and using those methods to better understand students [3]. Various data mining techniques are engaged in EDM such as decision tree, k-nearest neighbors, association rule, neural networks, genetic algorithms, exploratory factor analysis, and stepwise regression.

In most of the Asian countries, many vicissitudes are observable in the performance of the students during their undergraduate period. Thus, it creates a big problem for both the teachers and students to track the educational performance and take necessary steps to avoid the educational failures. So, an early prediction of the final CGPA can solve the problem and help the teachers, course-advisors, and instructors to identify the students at risk and also give high confidence to students in their studies.

In this study, we have proposed a Deep Neural Network (DNN) model to predict the the final CGPA of the undergraduate business students with a minimal error rate based on the transcripts data of first four semesters. We have considered the performance of a decision tree model as the baseline performance. Our experimental results have shown that the proposed DNN model can predict the final CGPA with Mean Squared Error (=0.008), Mean Absolute Error (=0.067), and Mean Absolute Percentage Error (=2.074). However, the comparison between the proposed model and the baseline model have concluded that our proposed DNN model has successfully performed a promising prediction by reducing the MSE, MAE, and MAPE by 0.0146, 0.0431, and 6.043 respectively.

II. Related Work

Over the past few decades, several studies have been conducted in the discipline of education. In the recent years, most of the research in this discipline have emphasized the use of data mining techniques. Some authors have applied several types of machine learning methods to perform different types of studies including predicting the performance of the students, selecting the optimized features for prediction, and discovering the hidden pattern in the educational data. In this phase, we are going to briefly discuss some of the latest and relevant studies.

To predict students' final GPA using only the grades of the previous courses the authors Al-Barrak et al. [4] have applied decision tree. They have also identified some impactful relations between the courses and final GPA. On the

other hand, Sumitha et al. [5] have proposed a framework to analyze and classify students' detail based on academic records. In their proposed framework, they have applied K-mean Clustering algorithm on the most appropriate attributes selected by the ranking analysis. They have also applied five efficient classification algorithms namely Naive Bayes, Multilayer perceptron, SMO, J48, REP tree to predict the students' performance and J48 outperforms the others with the highest accuracy (=97%).

Selecting the optimized features significantly improve the learning performance of the model. Mueen et al. [6] have identified the appropriate attributes by applying the ranking algorithm and also used 10-fold cross-validation to train up the model. They have tested three classifiers namely Naive Bayes, Multilayer Perceptron, and C4.5 (J48) on 38 attributes to predict and analyze students' academic performance and Naive Bayes have yielded the best accuracy (=86%).

An empirical study on attribute selection has been conducted where the authors proposed a small feature set by using both filter and wrapper methods [7]. The proposed feature set contains 'internal grades', 'previous failure', 'internet usage hours', 'health condition', 'attendance percentage', and 'family relation quality' those are sufficient to make accurate predictions of student grades.

The prominent researchers Daud et al. [8] have applied the learning analytics (discriminative and generative classification models) to predict whether a student will be able to complete his degree or not. They have used information gain and gain ratio to select the best features. For predicting the performance of students they have used five classifiers, two generative models namely Bayes Network (BN) & Naive Bayes (NB) and three discriminative models namely SVM, C4.5, and CART. And, the SVM has been found as the most effective with F1-score 0.867.

As the aforecited paper [5], an identical study [9] have been conducted where the authors proposed an integrated system framework which mentioned six variable domains (psychological, cognitive, economical, personality, demographic and institutional domains) that have a great influence on student performance. Their architectural framework has contained three layers such as the user interface layer, the database system layer, and the execution system layer.

A comparative analysis has been performed between two neural network models, namely Multilayer Perceptron and Generalized Regression Neural Network, to identify the best model for predicting the students' academic performance based only on their academic result [10]. The authors have split the whole dataset into three segments (60:20:20) where 60% have been used as the training data, 20% as the validation data, and 20% as the testing data. They have used three performance metrics: Mean Square Error (MSE), Receiver Operating Characteristic (ROC), and accuracy to evaluate the performance. According to the comparison of these two models the GRNN, with spread parameter 0.45, has been found as the better model than MP with higher accuracy (=95%).

III. Problem Background

In most of the Asian countries, there exists no linear relationship among the performances of a student during the undergraduate period. Some students perform better in the commencing period but worse at the ending period, some perform worse in the commencing period but better at the ending period. Again, there exists a few students who always perform better or worse. Thus, it makes difficulties for teachers, course-advisers or instructors to identify the students who are at risk and to guide them properly.

The prediction of the final CGPA, indisputably, can help teachers, course-advisers or instructors to identify the students at risk and guide them properly. Another significant problem is that at the middle point of the graduation period some of the students become frustrated by thinking that if they eventually not be able to obtain their expected outcome and this frustration obviously influence their performance that increases the possibility of falling out. So, the early prediction of the final CGPA undoubtedly can help the teachers or instructors to take a proper decision for the improvement of students' performance and also carry away the frustration that reinforces their mind to perform better.

Though students' transcript data are available to every educational institutions, it is definitely a challenging task to predict the final CGPA of the students because of the rapid vicissitudes and nonlinierities in their educational performance. To handle the aforementioned problems we have aimed to predict the final CGPA using a DNN approach only by utilizing the students' transcript data of the first four semesters.

IV. Description of Dataset

In this research, we have prepared a real dataset by using the students' transcript data of business students collected from the Marketing department of a renowned university, situated in Bangladesh. The dataset consists of the students' data who already completed their graduation in 2013, 2014, 2015, and 2016. The collected raw data, that have contained 398 instances in total, have been manually processed and organized in a google spreadsheet. The dataset contains the attributes regarding each of the students are Student name, Student ID, Gender, All the courses including course' grade, Semester GPA, State of the students based on semester GPA, and final CGPA. The course codes are 101(c), 102(c), 103(c), 104(c), 105(c), 111(c),112(c), 113(c), 114(c), 115, 201, 202(c), 203(c), 204(c), 205, 211(c), 212, 213, 214, 215, 301(c), 302, 303(c), 304, 305, 311, 312, 313, 314, 315, 401, 402, 403, 404, 405, 411, 412, 413, 414, 415 and detail are available in [11]. However, we have used the courses that are only included in the first four semesters to predict the final CGPA. The courses that are included in the first four semesters ranges from 101 to 215.

V. Data Preprocessing

To solve a real-world problem using machine learning, it is indispensable to preprocess the raw data properly for achieving a better result from the applied model as it directly affects the

ability of the model to learn. Real-world data are generally incomplete, noisy, corrupted, inconsistent, and imbalanced and so it is essential to preprocess the data before feeding it into the model. Data preprocessing is an integral step such as data cleaning, data integration, feature encoding, instance selection, normalization, transformation, feature extraction, and selection. In this study, we have structured this phase according to the form of our dataset which includes missing value handling, feature engineering, feature scaling, feature encoding, data partitioning. All the tasks are discussed briefly in the subsections below.

A. Feature Engineering

Feature Engineering is the process of extracting a new feature subset by utilizing the existing features into formats that are suitable for the machine learning model [12]. We have performed feature engineering techniques to extract the gender of the students based on their name. Also, we have developed new feature columns that indicate the state of the students after completing each of the semesters on the basis of their obtained GPA. According to Table 1, we have used four terminologies namely '*Excellent*', '*Very Good*', '*Good*', and '*At Risk*' to imply the state of the students.

TABLE I
GPA RANGE FOR EVERY DEFINED STATE

GPA	Terminology
3.5 or above	Excellent
3.00 to less than 3.50	Very Good
2.50 to less than 3.00	Good
Less than 2.5	At Risk

B. Feature Encoding

Feature encoding is the technique that represents the categorical data into the numeric form. The neural networks can only learn from the numeric data that is why it is mandatory to encode the categorical data into numbers before feeding the data to the model. In this study, the categorical features that include in our dataset are the gender and the state of the students mentioned above. In order to encode these variables into a compatible numeric form, we have applied the encoding technique named One-hot encoding. This technique generates a binary column for each of the n categories of a categorical variable, where n is the number of unique values contained in that variable [13].

C. Feature Scaling

Feature scaling [15] is a crucial step in the data preprocessing pipeline that is used for data standardization. This includes a few techniques that scale the data into a fixed range to bring all the features to the same level of magnitudes, which suppresses the effect of outliers. It is essential for training a neural network based model as it helps the weights to converge more quickly. For feature scaling, we have applied min-max scaling to each feature column, where the new value X_{sc} of a sample X can be calculated as follows:

$$X_{sc} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Here, X_{min} is the smallest value in a feature column, and X_{max} is the largest value, respectively. This scaling brings the new value between 0 and 1.

D. Data Partitioning

To properly access the whole dataset, we have applied the training-validation-testing approach in our study. In a neural network model, it is a basic need to tune the hyperparameters and be cautious about the model overfitting and underfitting. A validation set is crucial for properly tuning the hyperparameters and optimize the best model. From the total number of 372 instances, we have used 68% (=261) instances for training, 20% (=66) for validation, and 12% (=45) for testing the model.

VI. Methodologies

In this study, we have maintained a proper workflow for conducting our experiment [14]. The entire workflow basically contains three consecutive steps such as designing the architecture of the DNN model, selecting the evaluation metrics, and defining a baseline performance score for the model. All the implementations of these three steps have been performed using the Python programing language.

A. The Architecture of the DNN Model

DNN models are basically comprised of three major layers: the input layer, hidden layers, and the output layer. The domain-specific architecture of this model can be defined by a few high-level artifacts such as the number of neurons in a layer, the number of hidden layers, the activation function, the optimizer, the learning rate of the architecture, the number of epochs, batch size, and so on. All of these parameters are collectively employed to design a network, and these parameters lead the learning process and the end performance of the model. Thus, it is indisputably a challenging task to select the value of these parameters as they need to be selected with experience and judgment. Fig. 1 represents the visual architecture of our proposed model.

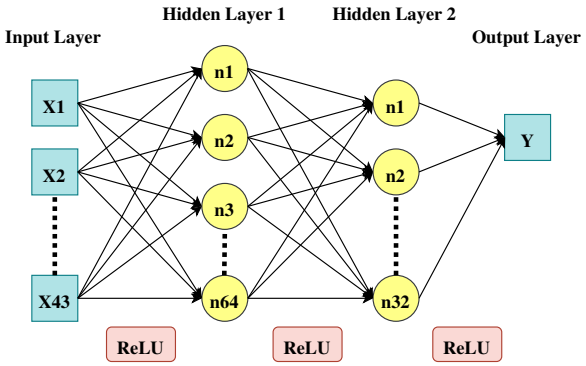


Fig. 1. DNN Architecture.

1) The Number of Neurons in the Input and Output Layers: A neural network architecture begins with configuring the hyperparameters which are the number of neurons in the input and output layer. The value of these hyperparameters depends on the shape of the training data. Our training data has a shape of (261, 43), which means there are 261 instances and 43 attributes. We have specified 43 neurons in the input layer as there are 43 input features in the training data. In contrast, we have specified only one neuron for the output layer as the problem is a regression problem and our expected outcome is a single numeric value.

2) The Number of Hidden Layers and Neurons: Determining the number of hidden layers and the number of neurons in each hidden layer is a crucial task as it characterizes the overall neural network architecture. A study conducted by Jeff Heaton [16] have concluded that networks with a single hidden layer can approximate any function that contains a continuous mapping from one finite space to another, networks with two hidden layers can represent an arbitrary decision boundary to arbitrary accuracy with rational activation functions and can approximate any smooth mapping to any accuracy, and networks with more than three hidden layers can learn complex representations and perform automatic feature engineering. As our training data is not very large, more than two hidden layers will make our model very complex, computationally inefficient, and may cause overfitting to the training data. So, we have designed our neural network architecture employing 2 hidden layers.

Again, to select the number of neurons in the hidden layer, we have applied a rule of thumb from [17]. The book has enunciated that the number of hidden neurons in the first hidden layer should be defined according to the number of input dimensions. If x be the final number of input dimension in a given training dataset, then we should use the nearest number to $2x$ in the power of 2. As our final input dimension is 43, we have employed 64 neurons for the first hidden layer which is nearest to $2x$ in the power of 2. Since the second hidden layer is the last hidden layer before the output, we have employed 32 neurons which the nearest number to $\frac{2x}{2}$ in the power of 2.

3) Activation Functions: Activation functions are an extremely important feature of the neural networks that perform functional mappings between the inputs and response variable. Basically, activation function is a non-linear transformation to the input signal making it capable to learn and perform more complex tasks. After computation, each hidden neuron of a hidden layer generates a single output which is then used as an input in the next layer in the stack. In this study, we have applied the activation function ReLU which is briefly discussed below.

Rectified Linear Unit (ReLU): ReLU [18] is the most widely used activation function while designing a DNN model. If the input is negative the function will convert it to zero and the neuron does not get activated, but for any positive value x, it returns that value back. It is defined as-

$$f(x) = max(0, x) \tag{2}$$

where x is the input to a neuron. So, this is the main advantage of using the ReLU function over other activation functions as it does not activate all the neurons at the same time. Activating a few neurons at a time making the network efficient and easy for computation.

4) Optimization Algorithm: In DNN, an optimization algorithm plays the most important role for obtaining a desirable result from the model. It shapes the neural network model into its most accurate possible form by adjusting the weights and train up model with an optimized cost function. Several optimization algorithms are used in neural networks such as Adam, Gradient Descent, Adagrad, RMSprop, and so on. However, for our experiment, we have applied the optimization algorithm Adam which is briefly described below.

Adaptive Moment Estimation (Adam): The Adam [19] optimization algorithm is an extension to stochastic gradient descent that updates network weights iterative based in training data and works well across a wide range of deep learning architectures. Indeed, this adaptive learning rate optimization algorithm is a combination of gradient descent with momentum and RMSprop algorithms. Adam has some advantages over other optimization algorithms such as it has relatively low memory requirements, invariant to diagonal rescale of the gradients, hyper-parameters have intuitive interpretation and typically require little tuning, and so on.

5) Evaluation Metrics: To evaluate the performance of our proposed model with measuring the error, we have applied three evaluation metrics namely Mean Squared Error, Mean Absolute Error, and Mean Absolute Percentage Error. All the metrics have been elected by considering their benefits and application. The details of these metrics have been discussed below.

Mean Squared Error (MSE): MSE is basically a measure of how close our fitted line is to data points. For each of the data points, it calculates the squared difference between the predicted value and the target value and then averages those differences. The smaller value of this calculation indicates the better fitting of the model. The following equation defines the mathematical formulation of MSE, that is:

$$\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{3}$$

Mean Absolute Error (MAE): MAE is extensively used for forecasting accuracy where the error is calculated as an average of absolute differences between the predictions and the target values. Basically, it is a linear score which means that all the individual differences are weighted equally in the average. One of the important advantages of this metric is that it penalizes huge errors that not as that badly as MSE does.

Thus, it is not that sensitive to outliers as mean squared error. MAE can be formulated as:

$$\frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{4}$$

Mean Absolute Percentage Error (MAPE): MAPE is a measure of the prediction accuracy that is used as a loss function for regression problems in machine learning. It is a non-scaled error metric and expressed as the relative error preference. For each data points, the absolute error is divided by the target value, yielding a relative error. The main advantage of this measure is its simplicity and it is easy to understand because it provides the error in terms of percentages. Thus, if the average range of the prediction is already known, we can easily estimate what the predictions are going to look like. Mathematically, it can be calculated by the following equation:

$$\frac{100\%}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{y_i} \tag{5}$$

6) Defining Baseline Performance: Defining the baseline performance is a crucial task to monitor and assess the effectiveness of the neural network model in the training phase. We have considered the performance of a Decision Tree [20] as the baseline for all of the test samples. We have conducted an experiment on our dataset by applying the Decision Tree and the baseline performance we have attained from the decision tree model are Mean Squared Error (=0.02390), Mean Absolute Error (=0.1114), Mean Absolute Percentage Error (=8.0047). To obtain a promising result, our model should offer a performance error less than the baseline error.

VII. Result and Analysis

In this study, we have measured the performance of our proposed neural network model using three evaluation metrics namely MSE, MAE, MAPE. Here, we have applied the MAE as the loss function and have trained our model up to 50 epochs. Fig. 2 presents the losses with respect to the epochs.

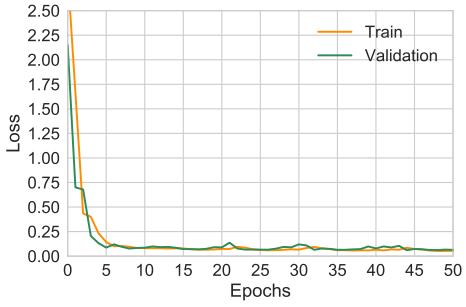


Fig. 2. Train vs Validaion loss across epochs.

The above visualization shows the learning rates of the model over the epochs. It can be apparent that the training and validation losses have been converged at 10th epochs. Here, the validation loss has been quite compatible in converging with training loss indicates a well-fitted model.

After the training and validation phase, we have applied the model on the test data to predict the final CGPA. The test data set have contained 45 instances. Table 2 presents the training, validation, and test result obtained from our model.

TABLE II
Performance of the Proposed Model

	MSE	MAE	MAPE
Training	0.13	0.119	3.536
Validation	0.071	0.116	3.517
Test	0.008	0.067	2.074

From Table 2 it can be observed that the validation result is very close to the training result, which indicates that our proposed model has learned well from the data without overfitting. Interestingly, in all the phases the results are quite compatible with very minimal variations.

In order to grasp the variations in actual and predicted CGPA more expressly, we have visualized all the 45 test cases into a line graph in the following Fig. 3.

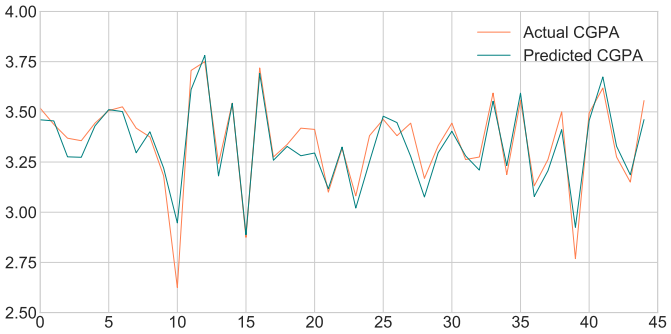


Fig. 3. Actual Productivity vs Predicted Productivity.

From the line graph in Fig. 3 it can be clearly observed that except few points, all the predicted values have a trivial variation with the actual values. For most of the samples, the model have fitted very well and predict the actual value with a minimal error. From all of the above analysis and error metrics we can concluded that our model has learned the underlying pattern from the educational data very well and can promisingly predict the actual CGPA of the undergraduate students.

The following Table 3 shows the comparison between the performance of the baseline model and the proposed model.

TABLE III
Comparison between the Baseline and the Proposed Model

	Baseline Performance (Decision Tree)	Proposed Model Performance (Deep Neural Network)
MSE	0.0226	0.008
MAE	0.1101	0.067
MAPE	8.1168	2.074

VIII. Conclusion

In this study, we have proposed a DNN based model in order to predict the final CGPA of the undergraduate business students with an error rate smaller than the traditional models. The dataset we have used was a real-world dataset collected from a renowned university in Bangladesh. We have performed a bunch of data preprocessing tasks such as feature engineering, scaling, encoding, and data partitioning to prepare our dataset in a suitable form for training the model. In the DNN architecture, we have selected the hidden layers, neurons, optimization algorithm and have tuned the hyperparameters by manual search and practical judgment. The performance of the model has been evaluated by using three evaluation metrics namely MSE, MAE, and MAPE. To examine the training and validation losses, we have used MAE as the loss metric. The training and validation performance of our proposed model indicates that the model has consistently learned from both training and validation data with very less deviation. In the test phase, our proposed model has also yielded very promising result such as 0.008 for MSE, 0.067 for MAE, and 2.074 for MAPE. And, our proposed model have significantly reduced the errors MSE, MAE, and MAPE by 0.0146, 0.0431, and 6.043 respectively than the baseline decision tree model. Most of the predicted points are very close to the true values and our proposed model performance have proved its consistent learning from the transcript data and showed a promising prediction performance.

In our future study, we aim to propose a framework that will be efficiently applicable for the prediction from this type of information. Furthermore, we will also incorporate more attributes such as age, study behavior, study time, family support, and disability for further analysis of the student performance.

References

[1] Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. Expert systems with applications, 41(4), 1432-1462.

[2] David, J. (2007). Hand. Principles of data mining. Drug safety, 30(7), 621-622.

[3] Sukhija, K., Jindal, M., & Aggarwal, N. (2016). Educational data mining towards knowledge engineering: a review state. International Journal of Management in Education, 10(1), 65-76.

[4] Al-Barrak, M. A., & Al-Razgan, M. (2016). Predicting students final GPA using decision trees: a case study. International Journal of Information and Education Technology, 6(7), 528.

[5] Sumitha, R., Vinothkumar, E. S., & Scholar, P. (2016). Prediction of Students Outcome Using Data Mining Techniques. International Journal of Scientific Engineering and Applied Sciences, 2(6).

[6] Mueen, A., Zafar, B., & Manzoor, U. (2016). Modeling and predicting students' academic performance using data mining techniques. International Journal of Modern Education and Computer Science, 8(11), 36.

[7] Chaudhury, P., & Tripathy, H. K. (2017). An empirical study on attribute selection of student performance prediction model. International Journal of Learning Technology, 12(3), 241-252.

[8] Daud, A., Aljohani, N. R., Abbasi, R. A., Lytras, M. D., Abbas, F., & Alowibdi, J. S. (2017, April). Predicting student performance using advanced learning analytics. In Proceedings of the 26th International Conference on World Wide Web Companion (pp. 415-421). International World Wide Web Conferences Steering Committee.

[9] Adejo, O., & Connolly, T. (2017). An integrated system framework for predicting students' academic performance in higher educational institutions. International Journal of Computer Science and Information Technology (IJCSIT), 9(3), 149-157.

[10] Iyanda, A. R., Ninan, O. D., Ajayi, A. O., & Anyabolu, O. G. (2018). Predicting Student Academic Performance in Computer Science Courses: A Comparison of Neural Network Models. International Journal of Modern Education and Computer Science, 10(6), 1.

[11] Department of Marketing. Website: http://dept.ru.ac.bd/marketing/index.php/programs/bba (Accessed on: 05 February 2019).

[12] Zheng, A., & Casari, A. (2018). Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists. " O'Reilly Media, Inc.".

[13] Marko. Mooi Sarstedt (Erik.). (2018). Concise Guide To Market Research: The Process, Data, And Methods Using Ibm Spss Statistics. Springer-verlag Berlin An.

[14] Al Imran, A., Amin, M. N., Rifat, M. R. I. & Mehreen, S. (2019). Deep Neural Network Approach for Predicting the Productivity of Garment Employees. In 2019 6th International Conference on Control, Decision and Information Technologies (CoDIT). IEEE.

[15] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of machine learning research, 12(Oct), 2825-2830.

[16] Heaton, J. (2008). Introduction to neural networks with Java. Heaton Research, Inc.

[17] Moolayil, J. (2019). Learn Keras for Deep Neural Networks. Apress. https://doi.org/10.1007/978-1-4842-4240-7

[18] Hahnloser, R. H. R., Sarpeshkar, R., Mahowald, M. A., Douglas, R. J., & Seung, H. S. (2000). Erratum: Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. Nature, 405(6789), 947–951. https://doi.org/10.1038/35016072

[19] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[20] Quinlan, J. R. (1986). Induction of decision trees. Machine learning, 1(1), 81-106.

Authorship Attribution in Bengali Literature using Convolutional Neural Networks with fastText's word embedding model

Hemayet Ahmed Chowdhury
Department of Computer Science and Engineering
Shahjalal University of Science and Technology
Sylhet, Bangladesh
Email: hemayetchoudhury@gmail.com

Md. Azizul Haque Imon
Department of Computer Science and Engineering
Shahjalal University of Science and Technology
Sylhet, Bangladesh
Email: azizulhaqueimon@gmail.com

Syed Md. Hasnayeen
Department of Computer Science and Engineering
Shahjalal University of Science and Technology
Sylhet, Bangladesh
Email: rummanhasnayeen94@gmail.com

Md. Saiful Islam
Department of Computer Science and Engineering
Shahjalal University of Science and Technology
Sylhet, Bangladesh
Email: saif.acm@gmail.com

***Abstract*—Authorship attribution (AA) is the process of attempting to identify the likely authorship of a given document by analyzing previous works of the authors in question. This paper proposes deep neural network-based continuous skip-gram models for authorship attribution in Bengali Literature. We present a data set of 2400 Bengali blog posts from 6 authors of current time and compare the performances of traditional lexical n-gram based models to our proposed approaches. We achieve a best accuracy of more than 92% on the held-out dataset with a deep convolutional neural network with skipgram word embeddings by fastText as the feature, which outperforms the other traditional models examined in this paper on Bengali Language. The results provide a clear indication that extracting features with the use of hidden layers in deep neural networks from continuous word embeddings work better as a feature set for authorship attribution systems on Bengali Literature than sparse lexical n-gram based features and shallow classifiers.**

***Keywords*—Authorship Attribution, Deep Learning, CNN, RNN, LSTM, skipgram, word embeddings, fastText.**

I. Introduction

One of the traditional classification problems is Authorship Attribution where an arrangement of texts with known authorship are utilized for preparing with an objective to decide the original author of the literary text [6]. The core application zones of authorship attribution include: identification of authors, detection of plagiarism, profiling authors, identification of inconsistencies in the writings of authors. As of late, authorship attribution is used in social media forensics as well to find falsified news sources, detect anonymous online threats and identify fake accounts [16]. Authorship Attribution in highly resrouced languages such as English received a great deal of consideration since the pioneering investigations of Mosteller and Wallace [1963] on the Federalist Papers. However, comparable work of such sorts are extremely uncommon in Bengali, an iconic language amongst the most broadly spoken South Asian languages. This absence of research advancement in Bengali Authorship Attribution could be credited to a lack of sufficient resources such as natural language processing tools and large data sets. [13].
In this paper, we approach the problem of Authorship Attribution with specialised deep neural networks based on word-embeddings. The paper explores some previously used methods and provides a comparative analysis of our approach with the traditional ones. After thorough investigations, we see that Deep Neural Networks, specifically the Convolutional Neural Network in our data set performs the best if implemented with word-embeddings from fastText, providing an accuracy rate of more than 92%. Our experiments give clear indications that word embeddings work as a much better feature than sparse and distinct lexical n-grams in capturing the stylometric patterns of different authors. Deep learning models provide game-changing results, outperforming the shallow classifiers, showing us that the hidden layers in the deep neural networks are capable of extracting the best possible features on their own from the word embedding representations.

II. Related Works

Authorship Attribution has been studied for a long period of time for languages which have the most resources to be worked with. In the initial stages of stylometric studies,

most notable work was done by Mosteller and Wallace. They worked with distribution of 30 function words, words like prepositions, conjunctions, and articles on various federalist papers to identify the original author of those papers[12]. Another notable work that was done by Bogdanova and Lazaridou was focused on cross-language Authorship Attribution. They used books from six different English authors and the Spanish translations of those books and proposed that a starting point to cross-language Authorship Attribution could be machine translation[1]. Zhao et al. [20] obtained impressive results on AP, Gutenberg, and Reuters-2158 corpora by using Kullback-Leibler divergence with Dirichlet smoothing, whereas, Nasir et al. [14] approached semi-supervised anomaly detection via multiple kernel learning for Authorship Attribution. For Authorship Attribution of short texts Sanderson and Guenter [2006] studied character and word sequences and for comparing the performances they used two Markov chain approaches[18]. Sequence kernels of different configuration were applied to a multi-topic data set of fifty authors and the performance was similar to the two Markovian approaches. The results were comparable with character sequence kernels and the results were better than word sequence kernels. The results also suggested that the size of training data had more impact on discrimination power than the amount of test data. It was revealed that if short texts were involved, the unmasking method would be less useful. In 2007 Jonathan H. Clark used synonym-based features for authorship attribution[4]. Bozkurt I. N. et Al. [2] took features like vocabulary diversity, frequency of words like articles, pronouns, conjunctions and bag of words to identify the stylometric patterns of five Milliyet columnist.

In highly resourced languages like German and English, research in the field of authorship attribution has come a long way in recent times. However, the same cannot be said about the progress in Bengali Literature. Up until 2015, only three threads of substantial research was note-worthy in Bengali. Das and Mitra worked studied authorship attribution on a data set of three authors consisting of 36 documents in total [8]. They used uni-gram and bi-gram features along with a probabilistic classification method, which is a variation of the Naive Bayes classification method, where uni-gram delivered an accuracy of 90% and bi-gram, a solid 100%. However, the validity of their experiments come into question due to the fact that the data set was small and the authors had very different styles of writing, making the classification task easier by a large margin. Chakraborty performed a ten-fold cross-validation on three classes showing that Support Vector Machine classifiers can provide best accuracy of up to 84% [3]. Jana looked into Sister Nivedita's influence on Jagadish Chandra Bose's writings which, as intriguing as it is, does not directly concern the topic in question, which is attribution of authors. [10].Since then, Shanta Phani also attempted attribution on three authors using machine learning techniques, much similar to Suprabhat Das's work [15]. P. Das, R. Tasmim and S. Ismail have also investigated the field on four Bangladeshi authors of current time and used lexical features like word frequency, word and sentence length, type-token ratio, number of conjunction and pronoun, etc [7]. Hossain and Rahman developed a voting system with multiple features on an SVM and achieved an accuracy of 90.67% based on lexical and stylometric features. [9]. Taking all the previous works into account, we can clearly see two patterns in the studies that were performed. Most of the approaches were based on shallow classifiers and sparse, discrete features like n-grams. What is also notable is that very little work has been done using deep learning models and the use of word embeddings as text representation is non-existent in the field of authorship attribution in Bengali literature. Our approach consists precisely of the two techniques aforementioned, where we use specialised deep neural networks, such as the RNN with LSTM gates and networks with hidden convolutional layers based on word embeddings as the text representation, based on the consensus that the deep layers of the neural network models will be able to extract the stylometric features held by the word embedding representation.

III. Corpus

A web crawler was developed using python, which collected 2400 articles of six different authors from three online public blogs. To keep the data set as balanced as possible, 350 documents were used per author. To maintain privacy and copyright protection, the authors are code-named FE, HM, IJ, MT, RN, and RG for this paper.

The data set was not normalized or cleaned, leaving it comparatively noisy. The blog texts are also collected from blogs of similar topic and time, making the data set more realistic and difficult for classification. 2100 articles were used for training , while taking 300 documents from the corpus randomly as the testing set.

Author	Total Words
FE	2382241
HM	3372688
MT	2515428
IJ	2401315
RN	1418598
RG	2454574

TABLE I
Corpus Size Per Author

IV. Methodology

The methodology described in this paper is divided into three sections. The first part explains the feature extraction and selection procedures, while the latter describes the classifiers used in the experiments. The third section consists of the performance analysis of our proposed systems on the different feature sets, how they compare with the traditional methods and why they perform the way they do.

A. Feature Selection

1) Shallow Features: With regard to the choice of shallow features, it has been observed that analysis of high frequency words (mostly function, or closed class, words) and/or n-grams provides the most consistently reliable results in stylometry dilemmas.[11] Our data set was consistent with the observations due to it's sparsity and moderate size, providing best results on uni-grams and bi-grams. Thereby, our choice of shallow features for this paper were frequency based embeddings or more specifically,

- count vectors of uni-grams
- count vectors of uni-grams and bi-grams

2) Prediction Based Word Embeddings: Our proposed feature set for this paper was the skipgram model, a type of prediction based word embedding. Given a set of sentences, the model loops on the words of each sentence and tries to use the current word to predict its neighbours, or on from an abstract perspective, its context. Words with similar context are are placed closer in a vector space. The feature was selected based on the consensus that word embeddings that hold context will capture the stylometric features of an author better than independent sparse matrices of uni-grams and bi-grams.

Word2Vec, Glove and fastText are three of the more popular and established algorithms for generating word embeddings. However, in this case of Bengali Literature, a low resource language, our choice tilted more towards fastText due to its ability to represent rare words or words not present in the training set better than the other two. Word2Vec and Glove assign random vector spaces or completely ignore the words that are not in the training data set for representation, which affects performance in Bengali Literature due to the complex structure of the language and unique vocabulary of different authors. While Word2Vec and Glove feed individual words into their own neural networks they use for creating the word embeddings, fastText breaks words into several n-grams(sub words) before putting them into the embedding process. This breakdown of words into numerous n-grams, 3 in our case, means that rare Bengali words, or words not in the training set will now be represented better, since it is highly like that some of their n-grams appeared in some of the words that do occur in the training set. The procedure may not be one hundred percent accurate but shows much better results than just assigning null values or random vector matrices to words.

The continuous word embedding model is mainly extracted to be fed into the proposed specialized deep neural networks, where the dense layers will extract the stylometric features on their own during training.

B. Experimental Setup

1) Proposed Approach: In our approach, we proceed with skip-gram word level embeddings by fastText. Due to a lack of available pre-trained vectors based on a large corpus of words in Bengali we literature, we generate our own vector model based on the held out training corpus. In generating the word vectors, the following parameters were used:

- Vector Dimensions of length 100.
- Context window of 5.
- 5 learning iterations.

The CNN architecture we proposed is a modification of the one used by Collobert et al. [5], who used a CNN with a non-static word embedding channel where the vectors are modified during training using backpropagation. Our approach varies in the sense that the word embedding channel is kept static and the vectors that are pre-trained by fastText's embedding algorithm are not modified during model training.

Initially, the articles of each author are padded to a length of 2000 words, which seemed optimum for the held out data set. The word embedding representations of the padded texts created by fastText on the training data are fed into the input layer. Two more hidden layers are used over the convolutional layer to extract the important features and a feature vector is obtained. A Max-over-time pooling layer [5] in turn condenses this feature vector to its most important feature by taking its maximum value and naturally deals with variable input lengths. A final fully connected layer with softmax activation outputs the probability distribution of the authors present in the data set. The entire model is unrolled for 10 epochs. Hidden states are initialized to 0. The batch size is kept 32. The number of activation per layer is chosen to be 200. An initial learning rate of 1 is used.

The tools and environment setup employed during the implementation are as follows:

- Python 3.5
- Keras
- fastText
- Tensorflow
- NLTK
- Google Colab

2) Comparison Methods: For comparison purposes, we analyze our approach against some methods, most of which are traditionally used.

1) Shallow classifiers with lexical n-gram features
   - a Support Vector Machine with linear kernel and C=1 with n-grams of range 2 as features.
   - A Naive Bayes Classifier with n-grams of range 2 as features.
   - A multi-layered perceptron with n-grams of range 2 as features.
2) Another deep learning architecture
   - A recurrent neural network(RNN) where the word embedding representations of the texts are used as the input to an LSTM layer. Two more hidden layers are used to capture important features, and a softmax output layer delivers the probability distribution of the authors, just like the proposed CNN architecture. All other parameters, such as the static word vectors that remain unmodified during training are kept the same as the CNN approach.

C. Results and Analysis

To analyze the performance of the classification model, the testing data set of 50 articles per author was used. The training and testing data set have no articles in common.

We investigated how traditional classifiers like the Support Vector Machine(SVM) , the Naive Bayes Classifier (NB) and the flat Neural Network(NN) performed on our data set with bi-grams as their input, which showed comparatively decent performances as a feature set on the field of authorship attribution.

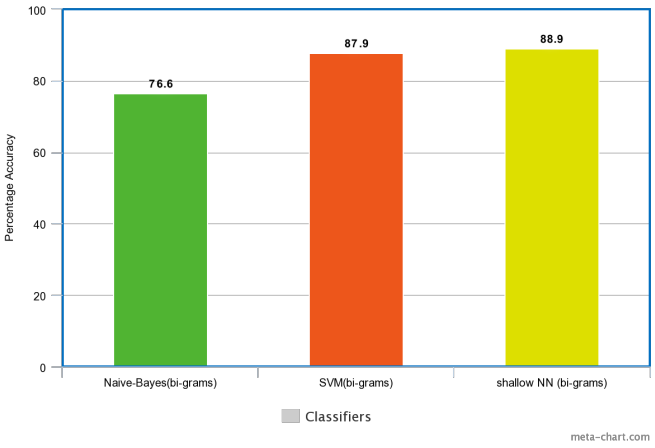


Fig. 1. Accuracy rates of different classifiers on n-gram features

Classifier	Feature	Accuracy
Naive Bayes	n-gram[1,2]	76.6%
Support Vector Machine	n-gram[1,2]	87.9%
Flat Neural Network	n-gram[1,2]	88.9%

TABLE II
PERFORMANCE ANALYSIS OF CLASSIFIERS ON N-GRAM FEATURES

Figure 1 along with Table 2 shows the performance of different classifiers on the shallow lexical n-gram based features.

From the diagrammatic representations, we can see the flat neural network clearly outperforms the other classifiers on the bi-gram feature set, with an accuracy of 88.9%, allowing us to pick neural networks as our choice of classifier for further studies.

Our original approach was to combine deep learning models with word embeddings as features, for which reason we chose to move forward with deep neural networks as they can extract features from the word embeddings very well.For our investigations, we performed separate experiments with a deep Recurrent Neural Network with LSTM gates (LSTM) and a deep Neural Network with a convolutional layer (CNN).

Classifier	Feature	Accuracy
CNN	skip-gram	92.9%
RNN with LSTM	skip-gram	89.6%
Flat Neural Network	n-gram[1,2]	88.9%

TABLE III
PERFORMANCE ANALYSIS WORD EMBEDDINGS AND LEXICAL N-GRAMS AS FEATURES

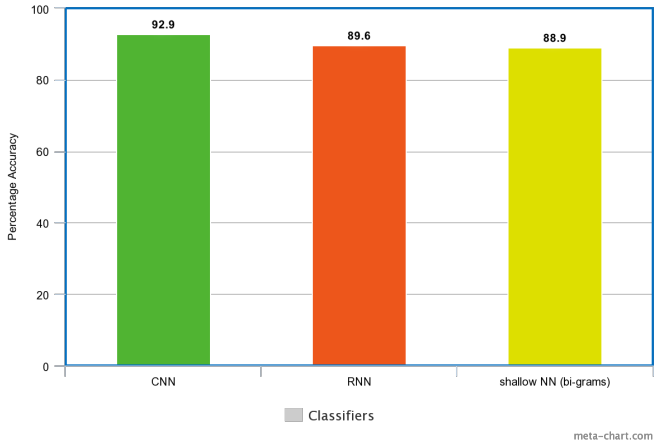


Fig. 2. Accuracy rate comparison of continuous word embeddings and lexical n-gram features with neural networks

Figure 2 along with Table 3 shows the effects of specialized neural networks on word embeddings. The RNN with LSTM gives us an accuracy of 89.67% on the data set, while the CNN outperforms all the previous models with an accuracy of 92.9% on the skip-gram word embeddings.

The RNN with LSTM is one of the very few classifiers that can hold memory, as it takes its previous output as one of the inputs [19]. The LSTM gates help to filter out the less important features, helping the dense layers to extract the features that define the author's styles the best [19].

CNNs are mainly used to process images due to their ability to detect patterns [17]. Sentences also seem to have a dimension, the sequence of words. The CNN in this case seems to capture the context and style of the different authors from the sequence of words very well in this scenario. On the held out dataset, the CNN seems to outperform all the other models , with an accuracy of around 92%.

We can clearly see using deep neural networks to extract stylometric features provides better results than using independent lexical frequency based features with shallow classifiers, as both the deep learning architectures outperform the best accuracy of 88.9% achieved by a flat neural network on bi-grams. The performance boost may be due to the fact the features extracted by well employed deep learning architectures, are at the very least in this case, better than the ones manually extracted. This just goes to show the potential deep learning possesses in the field of authorship attribution due to its ability to capture some of the most relevant and granular patterns in the writing styles of authors that may otherwise be a difficult task to complete manually.

V. CONCLUSION

On the held-out dataset, we have shown that word embeddings clearly outperform the lexical n-gram based features. Best accuracy of more than 92% was achieved with a convolutional neural network based on continuous skip gram word embeddings on our dataset. We believe as the data set size will get bigger, the margin between the accuracy score of

word embeddings and lexical n-gram features will increase. As of future work, we would like to investigate the impact of stacking multiple deep learning classifiers on large data sets. We believe ensembling deep learning models that are different in nature, that is, learn from different types of features, will provide best results.

REFERENCES

[1] Dasha Bogdanova and Angeliki Lazaridou. "Cross-language authorship attribution". In: *Proceedings of the 9th International Conference on Language Resources and Evaluation* (2014).

[2] I. N. Bozkurt, O. Baglioglu, and E. Uyar. "Authorship attribution". In: *2007 22nd international symposium on computer and information sciences* (Nov. 2007), pp. 1–5.

[3] Tanmoy Chakraborty. "Authorship Identification Using Stylometry Analysis in Bengali Literature". In: *CoRR* (2012).

[4] Jonathan H. Clark and Charles J. Hannon. "A Classifier System for Author Recognition Using Synonym-Based Features". In: *MICAI 2007: Advances in Artificial Intelligence*. 2007, pp. 839–849.

[5] Ronan Collobert et al. "Natural Language Processing (Almost) from Scratch". In: *J. Mach. Learn. Res.* 12 (Nov. 2011), pp. 2493–2537. ISSN: 1532-4435.

[6] R. M. Coyotl-Morales et al. "Authorship attribution using word sequences". In: *Progress in Pattern Recognition, Image Analysis and Applications* (2006).

[7] P. Das, R. Tasmim, and S. Ismail. "An experimental study of stylometry in bangla literature". In: *Electrical Information and Communication Technology (EICT), 2015 2nd International Conference* (2015).

[8] Suprabhat Das and Pabitra Mitra. "Author identification in Bengali literary works". In: *Pattern Recognition and Machine Intelligence,* (2011).

[9] M. Tahmid Hossain et al. "A stylometric analysis on Bengali literature for authorship attribution". In: *Computer and Information Technology (ICCIT), 2017 20th International Conference of IEEE* (2017).

[10] Siladitya Jana. "Sister Nivedita's influence on J. C. Bose's writings". In: (2015).

[11] Matthew L Jockers and Daniela M. Witten. "A comparative study of machine learning methods for authorship attribution". In: *Literary and Linguistic Computing (Lit Ling Comput)* (2010).

[12] P. Juola. "Authorship attribution". In: *Foundations and Trends in Information Retrieval* (2008).

[13] Sibansu Mukhopadhyay, Tirthankar Dasgupta, and Anupam Basu. "Development of an online repository of Bangla literary texts and its ontological representation for advance search options". In: *Proceedings of the Workshop on Indian Language and Data: Resources and Evaluation Workshop Programme* (2016).

[14] A. Jamal Nasir, Nico Gornitz, and Ulf Brefeld. "An off-the-shelf approach to authorship attribution". In: (2014).

[15] Shanta Phani, Shibamouli Lahiri, and Arindam Biswas. "Authorship Attribution in Bengali Language". In: *ICON*. 2015.

[16] Anderson Rocha et al. "Authorship Attribution for Social Media Forensics". In: *IEEE Transactions on Information Forensics and Security* (2016).

[17] Sebastian Ruder, Parsa Ghaffari, and John G. Breslin. "Character-level and Multi-channel Convolutional Neural Networks for Large-scale Authorship Attribution". In: *CoRR* (2016).

[18] Conrad Sanderson and Simon Guenter. "Short text authorship attribution via sequence kernels, Markov chains and author unmasking: An investigation". In: (2006).

[19] Wojciech Zaremba, Ilya Sutskever, and Oriol Vinyals. "Recurrent Neural Network Regularization". In: *CoRR* (2014).

[20] Ying Zhao, Justin Zobel, and Phil Vines. "Using Relative Entropy for Authorship Attribution". In: *Information Retrieval Technology*. 2006, pp. 92–105.

BHCDR: Real-Time Bangla Handwritten Characters and Digits Recognition using Adopted Convolutional Neural Network

Muhammad Aminur Rahaman*, Md. Mahin†, Md. Haider Ali‡, Md. Hasanuzzaman§
Dept. of Computer Science and Engineering
Green University of Bangladesh * †, *University of Dhaka* ‡ §
Dhaka, Bangladesh
aminur@cse.green.edu.bd*, mahin@cse.green.edu.bd†, haider@du.ac.bd‡, hzaman@cse.univdhaka.edu§

***Abstract*—Machine learning algorithm suffers to recognize the Bangla handwriting from images because of the complex design, diversities among different writers and similarity between characters and digits. In recent times, deep learning is becoming very popular among the researchers for Bangla Handwriting Recgnition (BHR) because of its high efficiency in terms of memory, time complexity and robust feature extraction. This research aims at improving the performance of baseline Convolutional Neural Network (CNN) by increasing the recognition accuracy with minimizing the computational overhead; this paper presents a real-time Bangla Handwritten Characters and Digits Recognition (BHCDR) system using adopted CNN. Our proposed preprocessing technique, data augmentation and incorporating dropout filters in the baseline CNN architecture have achieved the goal. The proposed eight layered architecture has used two convolutional layers followed by two Maxpooling layers with 25% dropout filters from one layer to another and two fully connected layers with 50% dropout followed by a softmax classifier. The proposed model is trained and tested using 118,698 images of Bangla lekha-isolated dataset and 21000 images of CMATERdb dataset for Bangla hand-written characters and digits maintaining the ratio of 4:1 respectively. The proposed model has achieved the mean accuracy of 97.43% for classification with the average computational costs of 44.95 ms/f.**

***Index Terms*—Convolutional Neural Network (CNN), Bangla Handwriting Recognition (BHR), Data Augmentation, Dropout Filters.**

I. Introduction

Bangla Handwriting Recognition (BHR) has been an area of research interest for a few decades because it has the prospect of saving time and efforts in digitizing important documents like medical prescriptions, official documents. Bangla language or Bengali which is the national language of the people's republic of Bangladesh is one of the most highly used languages as speaking and writing medium throughout the world. Bangla, which is used by 242 million people, is the sixth most used language all over the world [1]. In addition to being the official language of Bangladesh, it is also used as an official language in many states of India where it is the second most used language [2].

Bangla character set is one of the most complex character set because of it's designed complexity. It consists of fifty normal characters. It also consists of 334 compound characters [3] and the ten digits. In our research, we have used 50 characters and 10 digits only. Bangla characters have a complex convoluted structure with high symmetry to one another. There is also high symmetry between many characters and digits of bangla language. Because of this high symmetrical nature and complex convoluted structure of this character set, recognizing the hand-written Bangla characters and digits is a hard task. The difficulty increases multiply when it comes to handwritten character recognition as individual writing habit from each person also add separate complexity.

Among the very early researches on Bangla character recognition, the work of Ray et. al. [4], of Bangla online characters-collected from digitized pen and mouse is mentionable. Later a number of other studies have worked in the same domain and a number of them have adopted the neural network-based approach. In [5] the scholars have normalized the character images at the beginning and then used CNN for the classification task without any feature extraction method achieving accuracy for the training and test sets respectively above 93% and 85%. The work of Roy et. al. [3] have proposed a novel approach of greedy layer-wise supervised training method for deep convoluted neural network and have achieved an error rate of 9.67%. But most of these method suffers from high computational costs. Alom et. al. [6] have proposed a 11 layered CNN model for BHR and have achieved satisfactory results. Their evaluations were compared with a number of methods like Deep Belief Network, CNN with and without dropout, with Gaussian and Gabor filters respectively proved that their system performed better. A very recent research [2] has tried to improve the BHR by modifying Convolutional Neural Network based the ResNet-18 architecture. With an additional dropout layer placed after the second convolutional layer, the researchers are able to produce accuracy up to 94.6% with a boost of nearly 0.5%. More number of layers of the systems [2], [6] have made them more complex in terms of computational costs, feature extraction and recognition accuracy.

This research aims at improving the baseline CNN architecture [6] furthermore by adding additional improvements to increase the classification accuracy with reduced computational

cost. Finally, a comparison with the baseline architecture [6] for Bangla hand-writing recognition is presented. The main contributions of this paper are summarized as follows:

- Modifying the baseline CNN architecture by integrating dropout filters along with Batch Normalization (BN) in our proposed BHCDR to reduce the overfitting for producing more generalized output with regularization which reduces the computational costs.
- The performance of the proposed BHCDR model is improved by including Data Augmentation technique.
- We have added preprocessing module including segmentation, morphological dilation and erosion, Gaussian smoothing filter and finally clipping and normalization to increase the classification accuracy of the proposed adopted CNN model.

The rest of the paper is organized in the following manner: the proposed methodology is described in section II. experiment setup is presented in section III, result analysis is presented in section IV and in section VI conclusion is presented.

II. Proposed Methodology

The proposed model for Bangla hand-written characters and digits recognition (BHCDR) is shown in Fig.1. The proposed model is designed with two phases: First phase is designed for preprocessing and the Second phase is designed for classification using adopted CNN model. The proposed adopted CNN architecture is shown in Fig.2.

A. Baseline Convolutional Neural Network

Convolutional Neural Networks consist of the convolutional layer, pooling layers, and fully connected layer which are stacked entities with identity loops. Each entity or module consists of multiple convolutional layers which correlate its' inputs and map them into the next layer of the CNN including different kernels. The outputs of the kernels pass through the linear or non-linear activation functions ReLU to generate the output of the convolutional layer which is known as feature maps, mathematically represented by Eq.1.

$$In_j^l = f\left(\sum_{i \epsilon Im_j} In_i^{l-1} * K_{i,j}^l B_j^l\right) \quad (1)$$

Where In_j^l represents the outputs of each present layer, In_i^{l-1} is the output of the previous layer, $K_{i,j}^l$ represents the kernel (weight matrix) of the current layer, and B_j^l represents the bias of the present layer and Im_j indicates input map selection. Here, $f(.)$ represents a Rectified Linear Units (ReLU) which is a basically nonlinear activation function. The symbol * is represented as convolutional operator [7]. The MaxPooling function is used to compute the high-level features to reduce the size of features which minimizes the computational cost of the model [8]. The baseline model parameters of CNN consisting of 11 layers for Bangla handwritten isolated digits recognition system developed are listed in TABLE I which was developed by Alom et. al. [6]. Besides, the proposed adopted CNN model is configured as 8 layers are described in TABLE II. The baseline CNN used four convoluted layers, no dropout and Global Average Pooling (GAP) was used as classifier whereas the proposed model has used only two convoluted layers with three dropout filters and two fully connected layers instead of GAP.

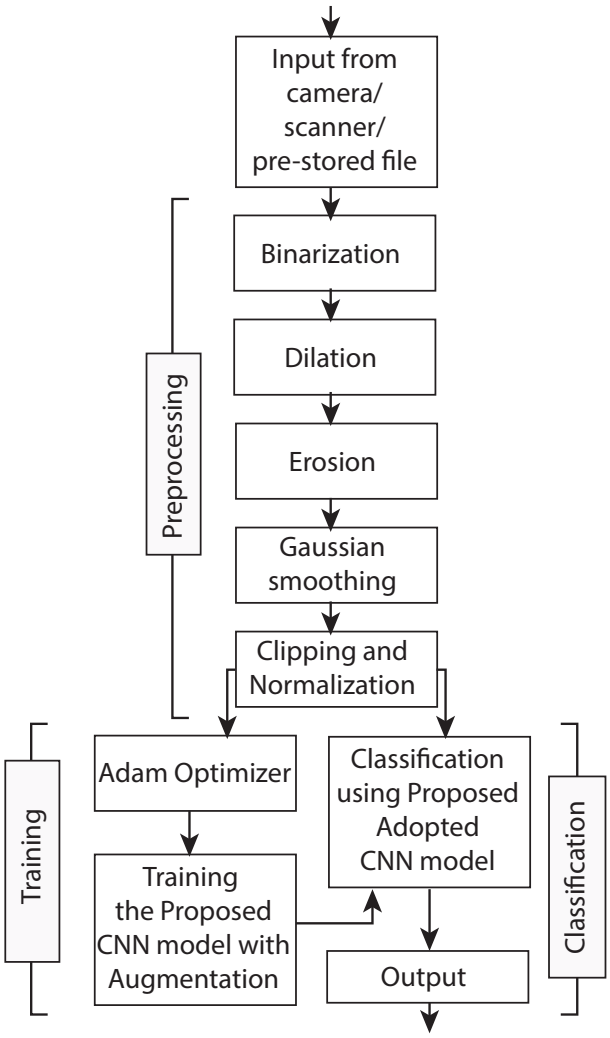


Fig. 1. Block diagram of the proposed system.

TABLE I
Baseline CNN configuration

Symbol	Descriptions	Output	Configurations
In	input	32×32	
Con1	Convolution-1	30×30	128@3×3, ReLU
Con2	Convolution-2	28×28	128@3×3, ReLU
MP1	Max-pooling-1	14×14	128@ ,2×2, stride 2
Con3	Convolution-3	12×12	256@3×3, ReLU
Con4	Convolution-4	10×10	256@,3×3, ReLU
MP2	Max-pooling-2	5×5	256@2×2, stride 2
Con5	Convolution-5	3×3	512@ 3×3, ReLU
Con6	Convolution-6	3×3	512@1×1, ReLU
GAP	Global Average Pooling	3×3	512 filters
Outputs	SoftMax	1×1	60 classes

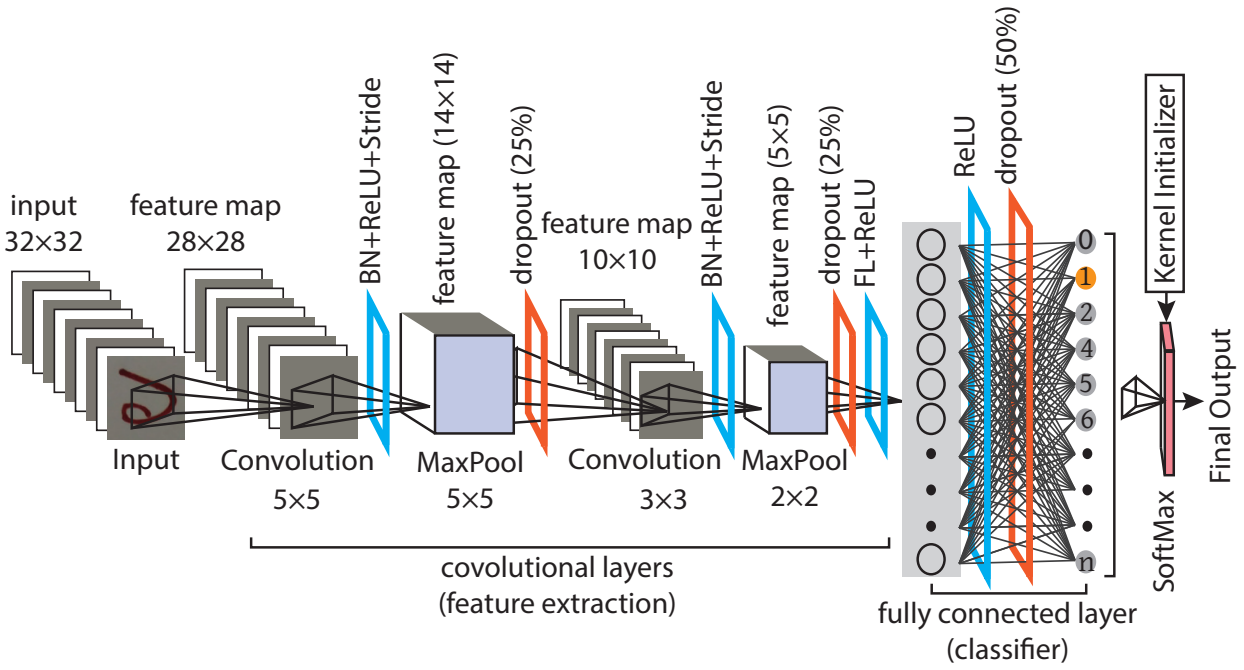


Fig. 2. Proposed adopted Convolutional Neural Network (CNN) model.

TABLE II
PROPOSED ADOPTED CNN CONFIGURATION

Symbol	Descriptions	Output	Configurations
In	input	32×32	
Con1	Convolution-1	28×28	32@5×5,BN+ReLU+Stride
MP1	Max-pooling-1	14×14	32@3×3,25% dropout
Con2	Convolution-2	10×10	64@2×2,BN+ReLU+Stride
MP2	Max-pooling-2	5×5	64@2×2,25% dropout,FL+ReLU
FC1	Fully Connected	1×1	256 Neurons,ReLU,50% dropout
FC2	Fully Connected	1×1	128 Neurons
SM	SoftMax	1×1	60 classes

B. Preprocessing

The proposed system acquires isolated Bangla handwritten characters and digits images from camera or scanner or pre-stored file in real-time. The system converts these acquired images into binary images. We have used two datasets to train and test the system. One dataset is the BanglaLekha-Isolated dataset [9] and another dataset is CMATERdb [10] (CMATERdb 3.1.1 and CMATERdb 3.1.2 for Bangla hand-written digits and characters respectively). The BanglaLekha-Isolated dataset which is already preprocessed as binary images where characters are converted to white and its background is converted to black. Another dataset CMATERdb is also preprocessed but its characters are black and the background is white. So, the system needs to invert the images of CMATERdb dataset to make as similar to the BanglaLekha-Isolated dataset. After that, morphological dilation and erosion operations are applied followed by Gaussian smoothing filter to make it noise free and robust dataset [11], [12]. All dataset are already resized in a square shape with appropriate padding [2]. We have used clipping and normalization technique [11] on the resized data to convert these into a fixed size 32×32.

Algorithm 1: Proposed algorithm for adopted CNN model execution for BHCDR

Data: 32×32 image.
Result: 60 classes of Bangla hand-written characters and digits.

```
begin
    Apply ADAM (δ = LearningRate = .002);
    for n iterations in all batch do
        Con1 (Filter, Kernel, BN, ReLU, Stride);
        MP1 (PoolSize);
        DR1 (DropoutRate=25%);
        Con2 (Filter, Kernel, BN, ReLU, Stride);
        MP2 (PoolSize);
        DR2 (DropoutRate=25%);
        FL (ReLU);
        FC1 (Neurons, ReLU, KernelSize, Initializer);
        DR3 (DropoutRate=50%);
        FC2 (Neurons, ReLU, KernelSize, Initializer);
        SoftMax (Classes, KernelInitializer);
    end
end
```

C. Proposed Adopted CNN model

The proposed adopted 8 layered CNN architecture is shown in Fig.2. Each entity or module of the proposed model takes input In and generates $F(In)$ output pairs which are then added to the input In of next convolutional layers. The resulted new input is represented by Eq.(2).

$$F(In) = F(In) + In \tag{2}$$

In our proposed model, dropout [13] layers are added after the first MaxPooling (MP1) and second MaxPooling (MP2) layers. Another dropout layer is also added in front of the second fully connected layer (FC2). Many researchers have used dropout to prevent feature adaptation and overfitting as a substitution of batch normalization (BN) [14] in their proposed model having a large number of parameters. Although another model [2] has used dropout along with BN which also performs better in generalization. Motivated by their approaches, we have applied the dropout along with BN in our proposed model. All dropout layers are added to reduce the overfitting for producing more generalized output with regularization. As a result, the proposed system ensures to minimize the overall computational costs.

The first layer of our proposed adopted CNN model is input layer which is connected with the first convolutional layer (Con1) containing a 5×5 kernel which has 32 filters and 1×1 stride size with a BN and ReLU activation function according to the Eq.3.

$$ReLu(In) = Max(0, In) \tag{3}$$

The first convolutional layer takes input as 32×32 pixels images. Next Max-Pooling layer (MP1) containing 2×2 kernel with 32 filters that uses maximum value to reduce significantly the number of parameters and computational costs in the network. This MP1 layer also controls the overfitting. Then the dropout layer (DR1) is configured with 25% (based on fine-tuning the model for different values) of neurons are excluded to reduce overfitting. The next hidden layer of the proposed model is another convolutional layer (Con2) containing a 3×3 kernel that has 64 filters and 1×1 stride size with a BN and ReLU activation function as like as Eq.(3). The second convolutional layer (Con2) is followed by another Max-Pooling (MP2) and 25% (based on fine-tuning the model for different values) dropout layer (DR2). A Flatten layer (FL) is used in this stage to convert the two-dimensional matrix data to a vector followed by a fully connected layer (FC1) with ReLU activation function and 50% (based on fine-tuning the model for different values) dropout layer (DR3) to obtain the final output to be processed next layers. The first fully connected layer (FC1) contains 256 neurons. The second fully connected layer receives the previous 256-dimensional output of the first fully connected layer (FC1) containing 64 neurons with the ReLu activation function which has 128 neurons. The final part of the proposed adopted CNN model is the Softmax activation function which is activated by Eq.(4) for final output layer containing 60 neurons for 60 classes of Bangla hand-written characters and digits.

$$f(In_i, W) = WIn_i \tag{4}$$

Where, f is defined as a mapping function that takes an input dataset In_i and produce them to the output classes as WIn_i which represents the dot product of the input dataset In_i and weight matrix W.

Inspired by [15], data augmentation is applied on both online and offline datasets to improve the model generalization and overfitting during the training phase. The system augments input datasets during the training phase by applying only rotations and translations operation. By using this rotation operation, the training datasets are increased 10 times in which the system rotates and translates the dataset 1% and 0.5% in each step respectively. All of these operations are applied randomly inside the mini-batches from scratch that are fed into the network. By training the shape and appearance of stream networks, the proposed model helps to converge faster. Incorporating the augmentation technique and dropout filters to the proposed model ensures that the proposed system is able to minimize computational costs by keeping higher classification accuracy.

Finally, the proposed model is executed according to the configuration listed in TABLE II using Algorithm 1. We have experimented using Adam [16] optimizer in the proposed Algorithm 1 with a learning rate of $\delta = 0.002$ (based on fine-tuning the model for different values of δ) to find the global minima of the categorical cross entropy loss function. Adam optimizer is memory efficient and computationally faster. More specifically, according to the experimental result of the system [2], Adam optimizer performs better than both Root Mean Square Propagation (RMSProp) and Stochastic Gradient Descent (SGD) optimizers.

III. Experimental Results

In our experiment, the proposed model for Bangla hand-written characters and digits recognition uses an ASUS UX305C model Notebook PC. The configuration of the experimental platform is specified as Intel Core m7 CPU and 2GB RAM. We have conducted an experiment on Microsoft visual studio 2010 and EmguCV (C and OpenCV wrapper) in Windwos10.

A. Datasets preparation

The proposed model is trained and tested using two datasets named BanglaLekha-Isolated dataset [9] and CMATERdb [10] (CMATERdb 3.1.1 and CMATERdb 3.1.2 for Bangla hand-written digits and characters respectively) dataset. The BanglaLekha-Isolated dataset contains 84 classes where 50 classes Bangla alphabet and 10 classes are Bangla digits (0-9) and 24 classes are joint letters. The CMATERdb dataset contains 231 classes of Bangla hand-written alphabets and digits images where 50 classes are alphabets (vowel and consonants) and 10 classes are digits and the other 172 classes are joint letters. In our experiment, we have used only 50 alphabet classes data and 10 Bangla digit classes data from both datasets.

Fig.3 shows the example training dataset- where class 01 to 50 are representing Bangla hand-written characters dataset including Bangla vowels (class 01-11) and consonants (class 12-50) and class 51 to 60 are represent-

ing Bangla digits. The BanglaLekha-Isolated dataset contains total (98,950+19,748)=118,698 images of Bangla alphabet and digits, and CMATERdb dataset contains total (15000+6000)=21000 images of CMATERdb 3.1.2 for Bangla hand-written 15000 characters and CMATERdb 3.1.1 for Bangla 6000 digits. We have used randomly selected about 4:1 of these both datasets for training and testing. According to the ratio, 94,958 images have been used as training dataset and 23,740 images have been used as testing dataset among 118,698 images of Bangla lekha-isolated dataset. Similarly, 16200 images have been used as training dataset and 4200 images have been used as training dataset among 21000 images of CMATERdb dataset.

B. Perfomance Mesurement Parameters

To evaluate the performance of the proposed Adopted CNN model and the baseline CNN model [6] we have used two parameters such as Accuracy and computational costs. Accuracy is calculated using Eq.(5).

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

Where TP is the true positive, TN is the true negative, FP is the false positive and FN is the false negative which are measured from the confusion matrices that are not presented here due to saving space and simplicity.

The computational cost for the real-time system is calculated based on the time it takes to preprocess, extract features using convoluted layers and classify a single test image of Bangla hand-written characters and digits against the training dataset in milliseconds per frame (ms/f).

C. Performance Evaluation

Fig.4 shows the accuracy for different 60 classes of Bangla hand-written characters and digits recognition using proposed



Fig. 3. Example training dataset.

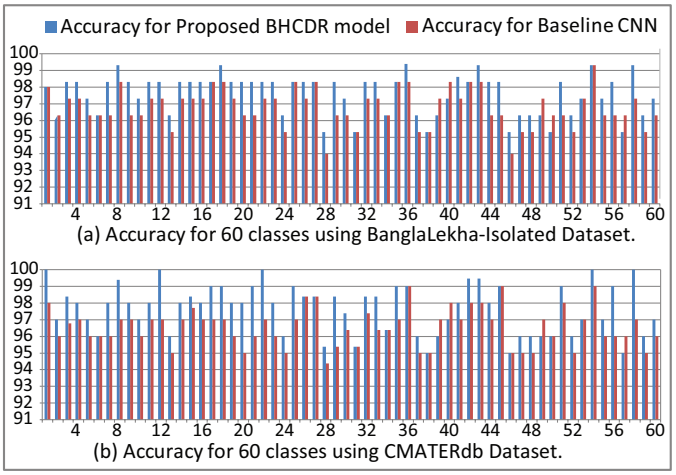


Fig. 4. Accuracy rate for different 60 classes of Bangla hand-written characters and digits recognition using proposed BHCDR model for (a) BanglaLekha-Isolated dataset and (b) CMATERdb Dataset.

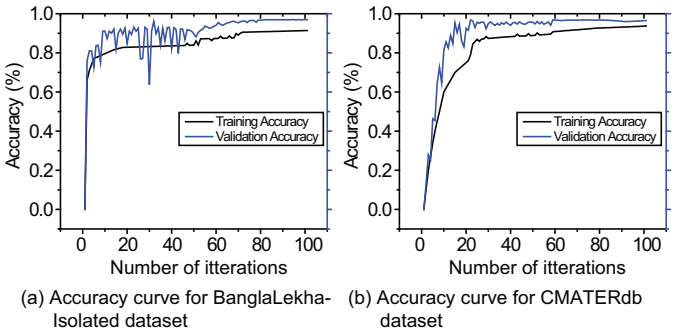
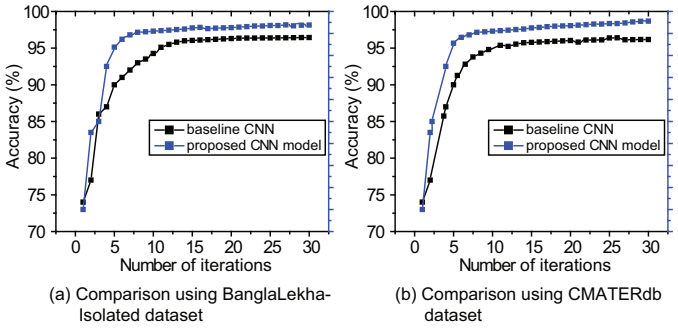


Fig. 5. Training and testing validation accuracy versus iterations curves for performance analysis of the proposed BHCDR model using (a) BanglaLekha-Isolated Dataset and (b) CMATERdb dataset.

BHCDR model comparing with the baseline CNN architecture [6] for BanglaLekha-Isolated dataset and CMATERdb Dataset. It is evident that the proposed system performs correctly and more accurately than the baseline model [6].

Fig.5 shows the training and validation accuracy curves for the proposed BHCDR model using both datasets. From this analysis, it is shown that the training and validation performances are close to each other for different iterations using both mentioned datasets which proves that the proposed BHCDR model is not overfitting the data.

The TABLE III shows the comparison of classification accuracy of the proposed model and baseline model. We have evaluated mean accuracy and computational costs for the proposed model and baseline model using both datasets of BanglaLekha-Isolated dataset [9] and CMATERdb [10]. The baseline CNN model [6] achieves 96.51% mean classification accuracy with the average computational costs of 89.27 ms/f. Our proposed model of CNN achieves 97.43% mean classification accuracy with the average computational costs of 44.95 ms/f which is relatively better than the baseline model [6]. The baseline model developed by Alom et.al. [6] have achieved



(a) Comparison using BanglaLekha-Isolated dataset

(b) Comparison using CMATERdb dataset

Fig. 6. Comparison analysis graph for the proposed BHCDR model with baseline CNN model using (a) BanglaLekha-Isolated dataset and (b) CMATERdb dataset.

TABLE III
COMPARISON ANALYSIS WITH RESPECT TO ACCURACY AND COMPUTATIONAL COSTS BETWEEN PROPOSED BHCDR MODEL AND BASELINE CNN ARCHITECTURE USING DATASET-1 (BANGLALEKHA-ISOLATED DATASET) AND DATASET-2 (CMATERDB DATASET)

dataset	accuracy (%)		cmputational cost (ms/f)	
	BHCDR	basline CNN	BHCDR	basline CNN
dataset-1	97.28	96.5	44.98	89.28
dataset-2	97.58	96.52	44.92	89.26
mean	97.43	96.51	44.95	89.27

99.13% average accuracy by experimenting their system for Bangla hand-written characters and digits dataset separately. But we have used both characters and digits datasets in a single experiment by assigning the class labels 01 to 60 which increases the confusions and reduces the accuracy due to similarity among characters and digits datasets.

Fig.6 shows the comparisons of classification accuracy between the proposed model and baseline model considering 30 iterations. Although the comparison graph has reached the optimal accuracy from 10 iterations, we have evaluated 20 more iterations for better compatibility. The comparison results prove that the proposed model performs two times faster than the baseline model with keeping higher classification accuracy.

The implementation of data augmentation technique in baseline CNN architecture improves the performance of the proposed BHCDR to achieve much higher recognition accuracy than the baseline CNN model without overtraining.

IV. CONCLUSION

This paper proposes real-time Bangla handwritten characters and digits recognition (BHCDR) using adopted convolutional neural network. The datasets are used in these experiments vary in different angles, sizes, contrasts, noises, and translation. Data augmentation technique is done on the dataset that generates transformed images of datasets and increases the training dataset instead of overtraining that ensures the robustness of the proposed model. Our preprocessing steps and introducing the dropout filters in the modified CNN model enhanced the performance of the proposed BHCDR by eliminating the data overfitting. The proposed model is trained and tested using 94,958 and 23,740 images of 118,698 BanglaLekha-Isolated dataset and 16200 and 4200 images of 21000 CMATERdb datasets respectively. The system has achieved 97.43% mean accuracy with the computational costs of 44.95 ms/f which is comparatively two times faster than baseline CNN model [6] with keeping higher classification accuracy. With this improvement, the proposed BHCDR results described that the proposed model is an effective CNN model for Bangla handwriting recognition. Future work of this research will extend and evaluate the proposed adopted CNN model with other compound characters datasets. The proposed system can be applied for automatic vehicle number plate recognition, postal code recognition, bank check recognition, any hand-written document digitization and so on.

REFERENCES

[1] T. Hassan, H. A. Khan, Handwritten bangla numeral recognition using local binary pattern, in: 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), IEEE, 2015, pp. 1–4.

[2] M. A. R. Alif, S. Ahmed, M. A. Hasan, Isolated bangla handwritten character recognition with convolutional neural network, in: 2017 20th International Conference of Computer and Information Technology (ICCIT), IEEE, 2017, pp. 1–6.

[3] S. Roy, N. Das, M. Kundu, M. Nasipuri, Handwritten isolated bangla compound character recognition: A new benchmark using a novel deep learning approach, Pattern Recognition Letters 90 (2017) 15–21.

[4] A. K. Ray, B. Chatterjee, Design of a nearest neighbour classifier system for bengali character recognition, IETE Journal of Research 30 (6) (1984) 226–229.

[5] M. M. Rahman, M. Akhand, S. Islam, P. C. Shill, M. Rahman, et al., Bangla handwritten character recognition using convolutional neural network, International Journal of Image, Graphics and Signal Processing (IJIGSP) 7 (8) (2015) 42–49.

[6] M. Z. Alom, P. Sidike, T. M. Taha, V. K. Asari, Handwritten bangla character recognition using the state-of-the-art deep convolutional neural networks, Computational Intelligence Neuroscience 2018 (Article ID 6747098) (2018) 1–13.

[7] J. Yang, Y. Zhao, J. C.-W. Chan, C. Yi, Hyperspectral image classification using two-channel deep convolutional neural network, in: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, 2016, pp. 5079–5082.

[8] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, pp. 1097–1105.

[9] M. Biswas, R. Islam, G. K. Shom, M. Shopon, N. Mohammed, S. Momen, M. A. Abedin, Banglalekha-isolated: a comprehensive bangla handwritten character dataset, arXiv preprint arXiv:1703.10661.

[10] CMATERdb, Hand written handwritten bangla numeral+character database, https://code.google.com/archive/p/cmaterdb/.

[11] M. A. Rahaman, M. Jasim, T. Zhang, M. H. Ali, M. Hasanuzzaman, A real-time hand-signs segmentation and classification system using fuzzy rule based RGB model and grid-pattern analysis, Frontiers of Computer Science 12 (6) (2018) 1258–1260.

[12] M. A. Rahaman, M. Jasim, M. H. Ali, M. Hasanuzzaman, Bangla language modeling algorithm for automatic recognition of hand-sign-spelled bangla sign language, Frontiers of Computer Science, (in press).

[13] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, The Journal of Machine Learning Research 15 (1) (2014) 1929–1958.

[14] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, arXiv preprint arXiv:1502.03167.

[15] S. C. Wong, A. Gatt, V. Stamatescu, M. D. McDonnell, Understanding data augmentation for classification: when to warp?, in: 2016 international conference on digital image computing: techniques and applications (DICTA), IEEE, 2016, pp. 1–6.

[16] D. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980.

Rice Disease Identification and Classification by Combining Deep Convolutional Neural Network with Support Vector Machine

Md. Jahid Hasan
Dept. of Electrical and Electronic Engineering
Hajee Mohammad Danesh Science and Technology University
Dinajpur, Bangladesh
jahidnoyon36@gmail.com

Shamim Mahbub
Dept. of Electrical and Electronic Engineering
Hajee Mohammad Danesh Science and Technology University
Dinajpur, Bangladesh
shamimmahbub230@gmail.com

Md. Shahin Alom
Dept. of Electrical and Electronic Engineering
Hajee Mohammad Danesh Science and Technology University
Dinajpur, Bangladesh
ashahin200@gmail.com

Md. Abu Nasim
Dept. of Electrical and Electronic Engineering
Hajee Mohammad Danesh Science and Technology University
Dinajpur, Bangladesh
abunasim373@gmail.com

***Abstract*—Rice is one of the most extensively consuming staple foods over the world. But due to various diseases of rice, farmers can't yield rice according to their expectation. Therefore, the quality, quantity and production of rice are being interrupted. It is difficult for the farmers to diagnose diseases, because of the lack of knowledge and unavailability of professional specialists. It also requires much time in manually identifying and classifying rice diseases. Therefore, an automated and accurate identification system has become essential to overcome this problem. A unique and fruitful approach has been presented to detect rice diseases with the implementation of an artificial intelligence (AI) and computer vision technologies. Here an AI has been developed by combining a Deep Convolutional Neural Network (DCNN) with Support Vector Machine (SVM). The Inception V3 DCNN architecture model has been modified using transfer-learning method. Then DCNN has been retrained with more than 1080 images of nine different rice diseases which are 80% of the image dataset. The features, extracted by the retrained DCNN was then used to train the SVM classifier. The network efficiently classified the images of nine diseases with an accuracy-level of around 97.5%.**

Keywords—Rice disease identification, Computer Vision, Artificial Intelligence, Deep Convolutional Neural Network (DCNN), Inception V3 model, Transfer learning, Support Vector Machine (SVM).

I. INTRODUCTION

In Bangladesh, rice is the major food of about 135 million peoples. Rice cultivation provides about 48% of rural development, it also contributes two-third of total calorie supply. An individual person can gain about one-half of the total protein requirement from rice. Over the past three decades rice is cultivated on about 10.5 million hectares which has remained almost stable in the meantime. About 75% of the total cropped area and over 80% of the total irrigated area is cultivated with rice [1]. Thus, rice plays a vital role in the livelihood of the people of Bangladesh as well as in the Asian countries. But the quality as well as the quantity of the rice production may be degraded due to plant diseases. As having diseases in plants are quite natural. If a proper step is not taken in this regard then it causes serious effects on rice plants which affects respectively product quality, quantity and productivity. These diseases are caused by pathogens such as fungi, bacteria, viroid, nematode and viruses. Typically, these pathogens damage the photosynthetic apparatus and affects the growth of plant [2]. Therefore, it is the major task to identify and classify rice plant disease in early stage [3].

At present, farmers use their own experience to detect diseases. Without recognizing plant diseases, farmers use pesticides in excessive quantities, which can't help in the prevention of disease, but can have malignant effects on plants. Though some rice diseases can create similar lesions or spots area and also different spots could be created from same diseases because of different rice diversities and local conditions or weather [4]. Thus, sometimes their misclassification creates bad impact on rice cultivation. So, they need an advice from rice disease specialist. In remote or rural areas, rice disease specialist can't able to give quick remedies or advice to the farmers in the right time and they also require expensive equipment and large amount of time for manually identifying and classifying rice diseases. On the other hand, for manual data processing requires more eyes to check and double check data for its accuracy. While an automated system identifies and classify diseases affected images more accurately than manually identifying process.

To overcome this problem, the process could be automated with the implementation of artificial intelligence (AI) and computer-vision technologies. For this reason, a hybrid network has been proposed by combining deep convolutional neural network (DCNN) with support vector machine (SVM) to classify rice diseases. This process could be made faster and easier when both time and human-efforts are saved significantly. To reduce loss function, a gradient based algorithm has been used by multi-layer architecture such as

DCNN [5]. In recent years, Convolutional Neural Network (CNN) is also used to analyze diseases of tea [6], apple [7], tomato [8], grapevine, peach, and pear [9]. It has also shown phenomenal accuracy in classifying and detecting that diseases. Deep convolution neural network is also used for achieving higher accuracy [10]. On the other hand, SVM is a classifier that learns a training dataset using properly chosen kernel parameters [11]. Combination of CNN and SVM has also been implemented recently for high dimensional data processing and classification such as visual pattern recognition [12] and image classification [13]. Thus, we were influenced to develop an AI by combining DCNN and SVM for the classification of rice diseases.

The proposed method represented a system to classify the disease of rice plants and assists the farmers to take exact calculation and also aid to increase the production. In this paper, for the classification of images of rice disease, we developed a hybrid network by combining both DCNN and SVM. A pertained Inception V3 DCNN model [14] was first modified and then retrained with a train-dataset containing more than 1080 images of nine rice diseases using transfer learning method. Then the retrained DCNN model automatically extracts one-dimensional features from the images that are used to train SVM classifier. The complete network is then tested on a test-dataset of 270 images.

The paper has been divided into six sections. In section II, various detection techniques of diseases of rice crops are reviewed. Section III presents the dataset preparation stage. Section IV explained the proposed modified Inception V3 DCNN architecture with SVM classifier to extract feature and classification respectively along with the results and discussion are illustrated in section V. Conclusions are summarized in section VI.

II. Related Work

A number of approaches have been already employed to accurately identify or classify plant diseases from images. Most of them used general image processing techniques, SVM classifier, K-means clustering, Neural Network (NN) and so on.

A pattern recognition based identification system of rice disease has been investigated by Phadikar et al. [15]. They used Self Organizing Map (SOM) neural network for classifying rice diseases. To normalize the size of disease spot, fractional zooming algorithm has been employed & images are segmented by Hue Intensity Saturation (HIS) model. In this paper, the authors considered two rice diseases named leaf blast and brown spot.

Another work has been done in [16] where they proposed an automated system based on the morphological changes for the classification of rice disease using Bayes & SVM classifier. They proposed two approaches to classify rice leaf diseases: (1) between infected and healthy leaves and (2) among infected leaves. They acquired 79.5% and 68.1% accuracies for Bayes and SVM classifier based system respectively.

In 2009 Yao at el. [4] developed an automated rice disease detecting system by presenting image processing techniques and SVM. They operated Otsu Method for the segmentation of disease spots. From Gray level co-occurrence matrix (GLCM), texture features were adopted to classify three (rice bacterial leaf blight, rice sheath blight and rice blast) disease of rice plants and SVM based function namely radial basis kernel function was selected for identifying these diseases. They earned overall 97.2% accuracy.

A computer vision based approach has been suggested by Ghyar et al. [17] for the detection of rice leaf disease. They considered only two disease of rice which are leaf blast and brown spot. Area of disease, GLCM and color moments are those extracted features. Artificial NN and SVM are used for classification of two diseases as well as healthy leaves. For getting higher accuracy, they employed K-means clustering for image segmentation.

To classify rice disease images, in this paper [18], non-linear SVM classifier using RBF kernel has been applied. Leave One-Out cross validation method was considered for the assurance of what training data could be learned about the applied method. Here they used Principal Component Analysis (PCA) technique for reducing pixel data from 3-dimension to two-dimension.

Recently, a deep convolutional neural networks (DCNN) based work was proposed by Lu et al. [19] to identify rice diseases by considering 10 common diseases for identification. A CNN model has been used to obtain a higher recognition accuracy and for the improvement of the convergence speed of the trained parameter. The older deep learning frameworks namely LeNet-5 and AlexNet were used as CNN architecture model for detecting diseases. They used around 500 image samples. Though, this small dataset was not quite enough for the classification of 10 rice diseases when it comes to deep convolutional neural network. Nevertheless, the author claims to achieve 95.48% accuracy.

Another DCNN approach has been taken in [20], where the authors developed a system that can recognize rice leaves by its health status. They used three classes in total which included normal, unhealthy and snail infested rice plants. For the classification they employed an old AlexNet architecture model of DCNN and claim about 91.23% accuracy. Although, they didn't mention any specific diseases or classes of rice for the classification. The proposed model only can tell that it is affected by diseases or not. Considering all these issues, we have presented nine most commonly affected rice diseases image classification by combining DCNN with SVM classifier which gave higher accuracy.

III. Dataset Preparation

The task of preparing dataset consists of two steps: assortment of disease images and pre-processing.

A. Assortment of disease images

We collected images of disease of rice plants from our varsity rice research field under the supervision of a professor and also some online resources named IRRI [21], BRRI [22], BRKB [23], Plantix [24] and so on.

Here we considered nine different diseases of rice which are most commonly affected diseases mainly in Bangladesh as well as other countries. These are- Bacterial Leaf Blight, Rice Blast, Brown spot, False Smut, Leaf Smut, Red Stripe, Leaf Scald, Sheath Blight and Tungro. Generally, the color, shape & size of Rice blast disease depending on varietal resistance, environmental conditions and the age of the lesions [25]. Rice blast mostly attacks in these two organs of the rice plants and

those organs are- Leaf and Collar, and Neck and Node. Though we put these two affected organs in same class but we trained them separately. A catalogue is presented in Table I that includes the name of those diseases and disease area for which they are liable.

TABLE. I

SELECTED DISEASES FOR CLASSIFICATION

Disease Image	Scientific Name	Disease Name	Affected Areas
	*Cochliobolus miyabeanus*	Brown Spot	Leaf blades and glumes
	*Ustilaginoida virens*	False Smut	Rice grain
	*Xanthomonas oryzae*	Bacterial Leaf Blight	Leaf
	*Microdochium oryzae*	Leaf Scald	Leaf
	*Entyloma oryzae*	Leaf Smut	Entire leaf surface
	*Microbacterium sp.*	Red Stripe	Leaf
	*Rhizoctonia solani*	Sheath Blight	Leaf sheath
	*Rice tungro bacilliform virus*	Tungro Virus	Leaf
	*Magnaporthe oryzae*	Blast	Leaf and collar or neck and node

B. *Pre-processing:*

For the classification process, image pre-processing task plays a vital role. The main purpose of this section is to improve the quality and quantity of images of the dataset. Feature extraction performance of CNN mostly depends on the quality, diversity and numbers of image samples. Therefore, the purpose of pre-processing task is to escalate the number of image-samples without reducing image quality or resolution. Therefore, we manually crop those large images containing numerous samples and raise the number of images.

Then all images are rescaled to 299px square size, which is required by Inception V3. Finally, the images go through an augmentation process that involves random flipping and vertical or horizontal translation up to 10 pixels. Thus, the number and variety of samples are increased and the unnecessary edges are removed as well. An example of manual-cropping process implemented on a disease image samples is represented in Fig. 1.



(a) (b) (c) (d)

Fig. 1. Manual cropping process of a sample image.

IV. METHODOLOGY

The biggest advantage of using DCNN is the perfection in extracting invariant features. But it is not always perfect as a good image classifier due to most of the trainable parameters are tuned in the hidden layer. On the contrast, SVM tune very few parameters to provide high classification accuracy for perfectly extracted one-dimensional feature vector, though it fails to learn very complicated features due to nits fixed kernel parameters [26]. Therefore, we have used Inception-V3 DCNN model with a non-linear SVM classifier to create a hybrid network in order to achieve high prediction accuracy for classifying rice disease images. The complete procedure of predicting rice diseases sample is therefore separated into two sections: feature-extraction and classification.

A. *Feature-extraction:*

Feature is an important and unique property of an image. To find out these relevant features from an image using a suitable algorithm means feature-extraction. Accurate feature selection approach enhances the classification speed, accuracy and reveal to train a classifier with large amount of data. In feature extraction, we can obtain the most relevant information from the original data and represent that information in a lower dimensional space. So Feature selection is an important post processing step, since a classifier will not be able to recognize images from poorly selected features.

Here the principal advantage of CNN is that we don't need to manually extract features from images. CNNs are able to automatically learn to extract complex hierarchical features from an image by using different kinds of filters during training time.

One of the deepest and complex DCNN architecture is Inception-V3. To perform, classification of each image Inception-V3 model uses nearly 25 million parameters and 5 billion multiply-add and floating-point operations [12]. But it is not generally affordable for one individual to train such a huge network efficiently due to it requires massive amounts of data, more time and multiple GPUs and TPUs. Therefore, the transfer learning [27] method comes into play, which is the recycle of a pre-trained model that permits us to partially use its weights on our network. Thus, it reduces the required training-time and decreases computational complexity as well. The pre-trained Inception-v3 model trained with more than 1 million images to achieves state-of-the-art accuracy for recognizing general objects with 1000 classes [28]. The Inception V3 of the model has 316 layers where we kept frozen initial 42 layers (up to 'mixed0' layer) to prevent them

from updating weights. The rest of the layers were retrained with disease images of rice plants. The number of iterations performed by the retraining process is calculated from equation (i).

$$\text{Number of Iteration, N}i = (Ns/Bs)*Ne \ldots\ldots (i)$$

Here Ns represents the number of samples, Ne represents the number of epochs and Bs is for batch-size. The batch-wise training accuracy is increased through back-propagation process that reduces error-rate significantly after several iterations. The complete model architecture is shown in Fig. 2.



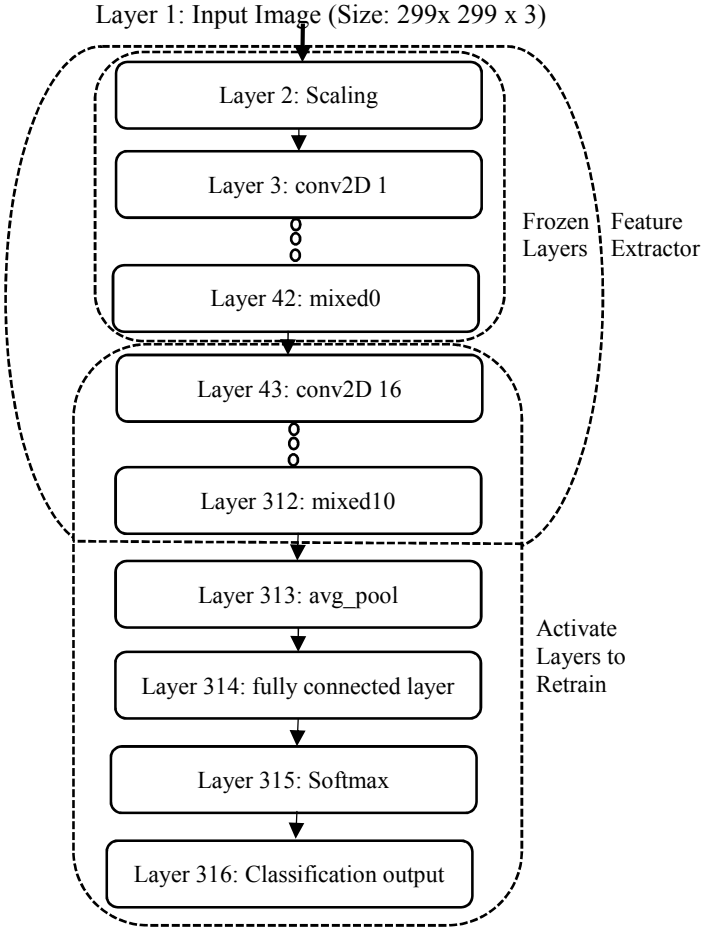


Fig. 2. Architecture of the modified Inception-V3 DCNN model.

Once the DCNN model is retrained, we eliminated its last 4 layers containing average-pool, fully-connected layer, softmax classifier and output layer. The rest of the network was then used for extracting features from rice diseases images.

B. *Classification:*

In classification stage, all extracted features from DCNN model are flattened by using 'global average-pooling' function. For future use the one-fifth features are separated. Then we train Support Vector Classifier (SVC) as a multi-class classifier with all rest of the features using one-VS-one method and it was implemented with RBF kernel. The hyper parameters of SVC classifier such as 'gamma', cost parameter 'C' were tuned using grid-search algorithm. Finally, the classifier was used to predict the previously separated features and calculated the validation accuracy. The complete process is represented as a block-diagram in Fig. 3.

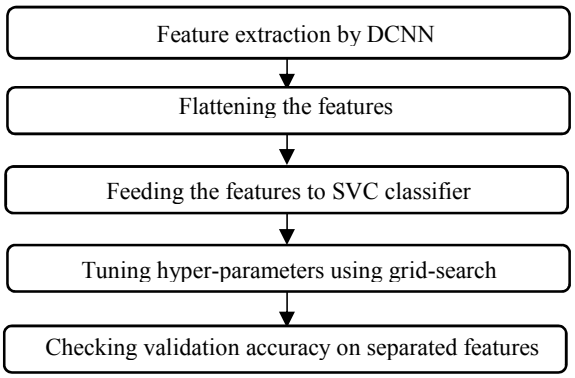


Fig. 3. Steps of training SVM classifier

Thus both DCNN and SVM models get trained for feature extraction and classification respectively. Now the hybrid network is ready to predict image of rice diseases.

V. Result and Discussion

This paper is experimented in two phases: Training Phase and Testing Phase. The complete process of training and testing was rendered on a personal computer (CPU: Intel core i7 @ 3.6GHz 64-bit, RAM: 8.00 GB and free-space on SSD: 100 GB) with 'Windows 10' operating system. We collected a dataset of images of rice plant disease including nine most commonly affected diseases. For training phase, we used 80% samples or images of the dataset to train the hybrid network and which is nearly 1080 images of disease. In order to validate the DCNN model, nearly 216 images are separated which is 20% of the training dataset. As, our target is to achieve high accuracy in extracting invariant features of rice diseases by using DCNN model. Therefore, in feature extractor section the value of active-layer's learning-rate was set to very low at 0.0001. Ultimately, the train and validation accuracy of DCNN model is tested for separate combinations of epochs and batch-size, which is shown in Fig. 4.

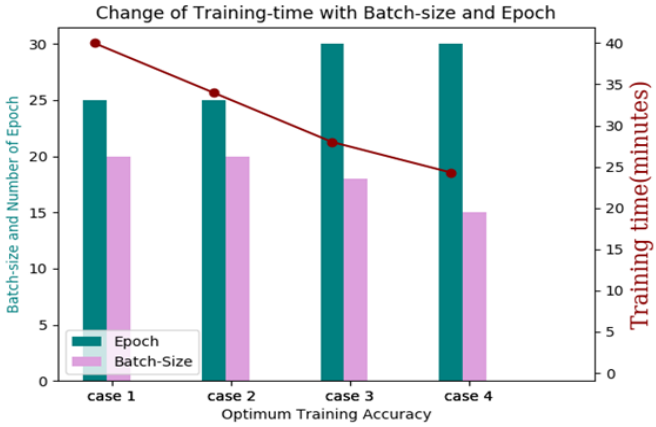


Fig. 4. Change of batch-size and epochs to reduce training-time

In Fig. 5, there are four cases representing four different combinations of batch-size and epochs. Validation accuracy for each of these cases were above 93%. Here case 4 combination (Batch-size: 30 and epoch: 15) required the least

training-time of 24 minutes 30 seconds. Thus, the training-time was reduced by around 18 minutes from case 1 to case 4, while the validation accuracy was reduced by only 2%. Now, the final values of all training-parameters of DCNN model, and hyper-parameters of SVM model are represented in Table II.

TABLE II

FINAL VALUES OF ALL USER-DEFINED PARAMETERS

DCNN		SVM	
Parameter	***Value***	***Parameter***	***Value***
Learning-rate of active layers	0.0001	Kernel	RBF
Batch-size	30	C	100
Number of epoch	15	Gamma	0.0001
Number of frozen layers	42	_	_

The graphical presentation of the training process was recorded while the DCNN model was retrained with the fixed user-defined parameters. It shows training and validation accuracy as well as training and validation loss, which are shown in Fig. 5 and Fig. 6 respectively.

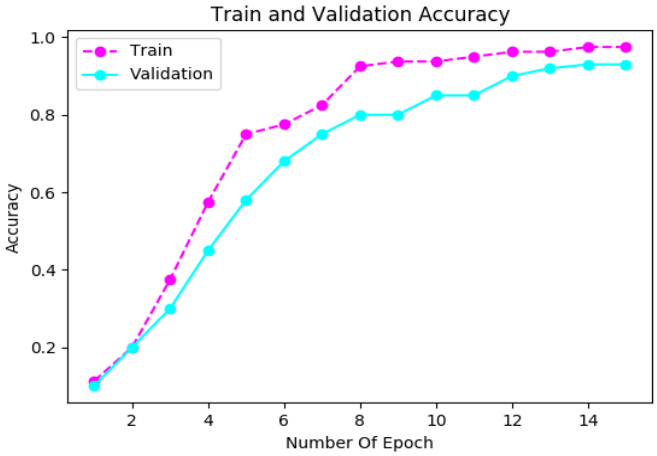


Fig. 5. Train and validation accuracy with respect to the number of epochs.

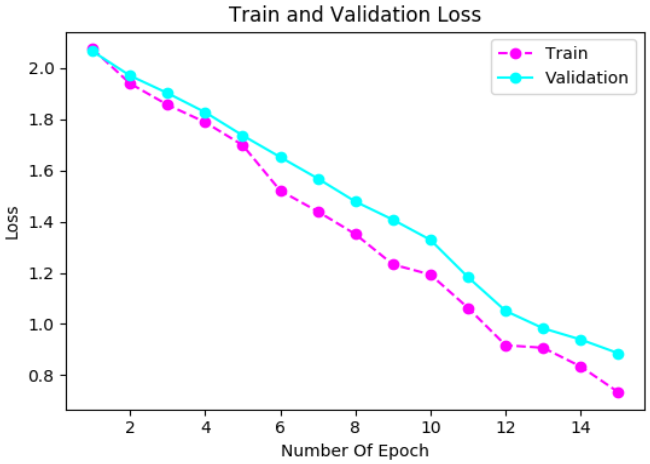


Fig. 6. Train and validation loss with respect to the number of epochs.

From Fig. 5 and Fig. 6, it has been observed that, during the last 4 epochs, both train and validation accuracy were above 93% and both train and validation losses were reduced below 1 respectively. The hybrid network with tuned SVM parameters was finally tested on the test-dataset of 270 samples, where it was achieved prediction accuracy of 97.5% and 94.6% in CNN-Softmax combination. In terms of CNN-Softmax combination, it needs massive amount of data to properly adjust model weights through back-propagation. But the combination of DCNN-SVM works well on small dataset for perfectly classifying an image due to its kernel trick. Therefore, the combination of DCNN-SVM has been used in this proposed model. Several classification results are shown in Fig. 7.

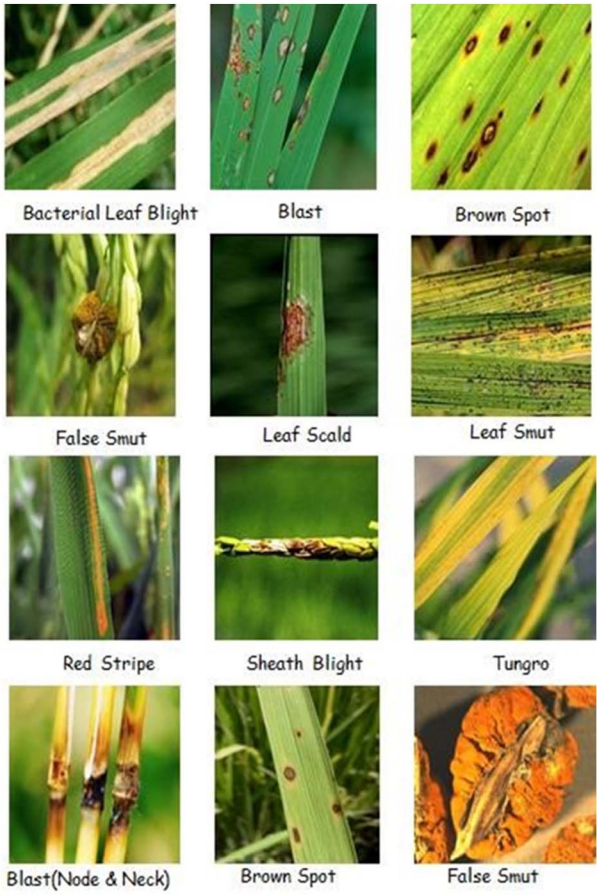


Fig. 7. A few results of classifying test-images of rice disease.

VI. CONCLUSIONS

Identification & classification of rice disease is quite impossible for farmers by naked-eye. It requires a significant amount of time, human effort and professional expertise. Hence, an effective approach has been taken to automate the process of classifying and identifying disease from images of rice plants. In this paper, the proposed method was aimed to develop an automated system to classify images of rice plant disease with the implementation of artificial intelligence and computer vision technologies. We represented a hybrid model by combining DCNN and SVM for the classification of rice diseases. The Inception V3 DCNN model was used for the automatic feature extraction. Supervised learning classifier SVM with RBF kernels was used for the classification of images of rice disease. We considered nine diseases of rice plants which are most affected diseases. Then

we built an image dataset of rice diseases. 80% of the dataset was used for training purpose and we achieved 97.5% prediction accuracy in training phase. This proposed model successfully classified a test-dataset of 270 images of rice plant diseases. Therefore, this proposed method has been made completely automated for the classification of rice diseases.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Md Mamunur Rashid, Professor at Department of Plant Pathology, Hajee Mohammad Danesh Science & Technology University (HSTU) for providing valuable information and image database and also full support in order to prepare this paper.

REFERENCES

[1] Rice in Bangladesh. [Online]. Available: http://www.knowledgebank-brri.org/riceinban.php. [Accessed: 16-Feb-2019].

[2] Hartmut K.Lichtenthaler1, "Vegetation Stress: an Introduction to the Stress Concept in Plants", Journal of Plant Physiology, Volume 148, Issues 1–2, pp 4-14, 1996.

[3] K. Muthukannan, P. Latha, R. Pon Selvi and P. Nisha, "CLASSIFICATION OF DISEASED PLANT LEAVES USING NEURAL NETWORK ALGORITHMS", ARPN Journal of Engineering and Applied Sciences, pp. 1913-1919, 2015.

[4] Q. Yao et al., "Application of support vector machine for detecting rice diseases using shape and color texture features", International Conference on Engineering and Computation, pp. 79-83, 2009.

[5] LeCun, P. Haffner, L. Bottou, and Y. Bengio. "Gradient Based Learning Applied to Document Recognition", Proceedings of the IEEE, Vol. 86, Issue. 11, Nov 1998.

[6] B. C. Karmokar, M. S. Ullah, M. K. Siddiquee, K. M. R. Alam, "Tea Leaf Diseases Recognition using Neural Network Ensemble", International Journal of Computer Applications, Volume 114 – No. 17, March 2015.

[7] G. Wang, Y. Sun, J. Wang, "Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning", Computational Intelligence and Neuroscience, Volume 2017.

[8] A. Fuentes, S. Yoon, S. C. Kim, D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition", Sensors 17 (9) (2017) 2022.

[9] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, D. Stefanovic, " Deep neural networks based recognition of plant diseases by leaf image classification", Computational intelligence and neuroscience 2016.

[10] X. Zhang, Y. Qiao, F. Meng, C. Fan, M. Zhang, "Identification of maize leaf diseases using improved deep convolutional neural networks", IEEE Access, vol. 6, pp. 30370-30377, 2018.

[11] C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, 1998.

[12] J. Nagi, G. A. D. Caro, A. Giusti, F. Nagi and L. M. Gambardella, "Convolutional Neural Support Vector Machines: Hybrid Visual Pattern Classifiers for Multi-robot Systems", 11th International Conference on Machine Learning and Applications, pp. 27-32, 2012.

[13] Abien Fred M. Agarap, "An Architecture Combining Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for Image Classification", arXiv:1712.03541 [cs], Dec 2017.

[14] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, pp. 2818-2826, 2016.

[15] S.Phadikar and J. Sil, "Rice disease identification using pattern recognition techniques", Computer and Information Technology,11th International Conference on, Khulna, pp. 420-423, 2008.

[16] S.Phadikar, J. Sil, and A. K. Das, "Classification of Rice Leaf Diseases Based on Morphological Changes", International Journal of Information and Electronics Engineering vol. 2, no. 3, pp. 460, 2012.

[17] Bhagyashri S. Ghyar ; Gajanan K. Birajdar, "Computer vision based approach to detect rice leaf diseases using texture and color descriptors", 2017 International Conference on Inventive Computing and Informatics (ICICI), India, pp. 1074-1078, 2017.

[18] Takuya Kodama, Yutaka Hata, "Development of Classification System of Rice Disease Using Artificial Intelligence", Japan, pp. 3699 - 3702, 2018.

[19] Y. Lu, S. Yi, N. Zeng, Y. Liu, Y. Zhang, "Identification of rice diseases using deep convolutional neural networks", Neurocomputing 267, pp 378-384, 2017.

[20] Ronnel R. Atole, Daechul Park, "A Multiclass Deep Convolutional Neural Network Classifier for Detection of Common Rice Plant Anomalies", International Journal of Advanced Computer Science and Applications, pp. 67-70, 2018.

[21] "Rice blast disease," Agriculture and Food. [Online]. Available: https://www.agric.wa.gov.au/rice/rice-blast-disease. [Accessed: 16-Feb-2019].

[22] Irri.org. (2019). Home | International Rice Research Institute. [online] Available at: https://irri.org/ [Accessed 18 Feb. 2019].

[23] Brri.gov.bd. (2019). Bangladesh Rice Research Institute-Peoples Republic of Banglaedsh. [online] Available at: http://www.brri.gov.bd/ [Accessed 18 Feb. 2019].

[24] Knowledgebank-brri.org. (2019). Bangladesh Rice Knowledge Bank | Bangladesh Rice Knowledge Bank. [online] Available at: http://knowledgebank-brri.org/ [Accessed 18 Feb. 2019].

[25] Plantix.net. (2019). [online] Available at: https://plantix.net/en [Accessed 18 Feb. 2019].

[26] F. J. Huang and Y. LeCun, "Large-scale Learning with SVM and Convolutional for Generic Object Categorization," 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, NY, USA, pp. 284-291, 2006.

[27] X. Li, T. Pang, B. Xiong, W. Liu, P. Liang, T. Wang, "Convolutional Neural Networks Based Transfer Learning for Diabetic Retinopathy Fundus Image Classification," 10th International Congress on Image and Signal Processing, Biomedical Engineering and Informatics, 2017.

[28] Keras.io. (2019). Applications - Keras Documentation. [online] Available at: https://keras.io/applications/#inceptionv3 [Accessed 18 Feb. 2019].

Human Activity Recognition Using Smartphone Sensors: A Dense Neural Network Approach

Abdul Kadar Muhammad Masum
Dept. of Computer Science and Engineering
International Islamic University
Chittagong, Bangladesh
akmmasum@yahoo.com

Sadia Jannat
Dept. of Computer Science and Engineering
International Islamic University
Chittagong, Bangladesh
sadiasuhe@gmail.com

Erfanul Hoque Bahadur
Dept. of Computer Science and Engineering
International Islamic University
Chittagong, Bangladesh
ehb1729@gmail.com

Md. Golam Rabiul Alam
Dept. of Computer Science and Engineering, BRAC University, Bangladesh
rabiul.alam@bracu.ac.bd

Shahidul Islam Khan
Dept. of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh
nayeemkh@gmail.com

Mohammad Robiul Alam
Dept. of Computer Science and Engineering
International Islamic University
Chittagong, Bangladesh
robi3443@gmail.com

Abstract*—At present, mobile apparatuses are becoming increasingly sophisticated and the latest generation of Smartphone that incorporates many diverse and powerful sensors such as linear or tri-axial accelerometer, linear or tri-axial gyroscope, Global Position System a.k.a. GPS etc. The availability of these sensors in mass-marketed communication devices creates new exciting opportunities for data mining and machine learning approach; besides creates a new research arena for Human Activity Recognition a.k.a. HAR. Wireless sensor based HAR has a significant impact on monitoring diseases related to ambulatory assessment, physical activities, elderly care, sports science, physical therapies and so on. The aim of this exploration is to recognize ten human activities stated as walking, walking upstairs, walking downstairs, sitting, standing, lying, using toilet, jogging, writing and typing, positioning the Smartphone on hand. We collected labeled tri-axial accelerometer and tri-axial gyroscope sensors data from both male and female subject while they had performed listed activities. We developed an android application to attain required sensors data implying frequency of 1 Hz. After that, we applied a Dense Neural Network a.k.a. DNN model to predict the activities from collected accelerometer and gyroscope sensors data. We also practiced Support Vector Machine, Decision Tree, K-Nearest Neighbor, Random Forest, Naïve Bayes and Logistic Regression along with DNN for operative study. Our main goal was to scrutinize a comparative study on both male and female subjects' data with statistical and deep classifiers. We also corroborated that, data amassed in lower frequency may perform good classification too. We achieved utmost 94.38% and 93.35% classification accuracy on the male and the female dataset respectively using DNN model.

Keywords—Human Activity Recognition; Accelerometer Sensor; Gyroscope Sensor; Smartphone Sensor; Dense Neural Network

I. INTRODUCTION

Human Activity Recognition a.k.a. HAR intends to recognize and identify the actions and goals of one or more individuals from a series of espials over a period of time-based on the person's actions and atmospherically situation [1]. The sensing of the activities is the fundamental step for establishing an activity recognition system. It plays a vital role in establishing intercommunication between persons and helps them to develop their interpersonal relations. Due to the availability of enough and important sensors in Smartphone, it simplifies a better user experience, enhanced information by providing user-friendly and exciting applications about the world around the Smartphone and provide robust and increased battery life. It can be considered as the best and most used communication device as more than 2.5 billion people are using Smartphone worldwide.

The sensors which are included in the Smartphone produce a different type of data in the variation of the movement of the Smartphone. And, amassed data by the Smartphone sensors on the basis of the movement of the Smartphone can be used in identifying the types of human physical movement. After identifying the physical movements, reports related to the health issues can be generated and then those reports can be utilized for the diagnosis of human activity related diseases in a timely basis.

Deep learning allows building complex concepts out of simpler concepts. Here input data is passed through a series of non-linearity or non-linear transformations before it becomes output. In contrast, modern machine learning algorithms are considered shallow because, the input can only pass a few levels manual identification of features in data and instead relies on whatever training process it has in order to discover the useful patterns in the input example. Known transformations in Multi-Layer Perceptron's a.k.a. MLP while back-propagation speed up learning methods. Nonlinear functions having the range from -1 to 1, centering all factors to zero mean in gradient-descent based optimizing and this leads to a significant speedup in learning. In this disquisition, gradient factor centering were done instead of the transform of nonlinearities makes the easier to generalize contexts. Second order optimization methods are useful to decrease the number of iterations to pruning but they are heavy in cost in the computation of high dimensional model with large matrices.

In our exploration, we have collected tri-axial accelerometer and tri-axial gyroscope sensors data from both male and female subjects, placing the Smartphone on their hand. We developed an android application to attain required sensors data implying frequency of 1 Hz. After the data collection phase, since we have used an android application, we had an intense fear to experience faulty data due to

hardware or software error. Thus, we eradicated null values, outliers, inconsistent instances, and standardization approaches. And, we also performed Principle Component Analysis a.k.a. PCA to select required features. Then we feed them to the DNN model along with other baseline models like Support Vector Machine a.k.a. SVM, Decision Tree a.k.a. DT, K-Nearest Neighbor a.k.a. KNN, Random Forest a.k.a. RF, Naïve Bayes a.k.a. NB and Logistic Regression a.k.a. LR to recognize the activities in consonance with walking, walking upstairs, walking downstairs, sitting, standing, lying, using toilet, jogging, writing and typing. We achieved utmost 94.38% and 93.35% classification accuracy on the male and the female dataset respectively using DNN model.

In the rest of this disquisition, a precise literature review is described in section II. We have incorporated activity recognition methods in section III, experimental evaluation in section IV and section V summarizes about our findings and future scopes.

II. Related Work

There exist many studies for recognizing simple, complex or combination of both human activities where an accelerometer or gyroscope sensor was either used alone or in a combination of the accelerometer with gyroscope sensor [2]. Besides, with these two sensors, other sensors such as a magnetometer, temperature sensor, proximity sensor and so on were also used in the research field of human activity recognition.

To briefly go over the history of human activity recognition system, many researchers [3, 4] determined physical activities using multiple wearable sensors placed on body parts like knee, waist, arms and ankles. They developed a classification model for classifying activities which are sitting, walking and standing. Again, wearable sensors like blood volume pulse sensor, galvanic skin response sensor, an electromyogram sensor and a respiration sensor were used for the identification of physiological signals, but these wearable sensors are quite expensive as well as, maintaining those sensors needs regular efforts. Moreover, the wearable sensors should be lightweight and capable of providing high performance [5]. Other than wearable and ambient sensor, some works were done using Smartphone sensors. Bayat, et al. [6] and R. Kwapisz, et al. [7] determined six activities using the accelerometer sensor of Smartphone in different works. Recently activity recognition using sensor-based data became a hot research topic for a wide range of applications. Several works have been done related to human activity recognition. In the past, most of the works were conducted using ambient sensors or sensors attached to the body. A system was created using a combination of twenty different sensors to identify activities like walking, standing, laying etc. [8]. Similarly, Maurer, et al. [9] used an e-Watch to recognize six activities in consonance with descending, running, sitting, standing, ascending and walking. Moreover, Tapia, et al. [10] conducted research on home-usable sensors only. Later in another research, five accelerometers were placed on different locations of 21 gymnasts' body Tapia, et al. [11]. Activities namely walking, climbing downstairs, sitting down, standing up were being recognized by Brezmes, Gorricho and Cotrina [12]. For sitting, lying down, standing, walking, stair climbing, running and cycling Mannini and Sabatini [13] used five bi-axial accelerometers staying at the different body position such as hip, wrist, arm, ankle and thigh.

Some other simple activities like walking, running, upstairs, downstairs, sit-ups, brushing teeth, vacuuming recognized by Ravi et. al. [14]. They made a comparison of base-level and meta-level classifiers and resolved that combining classifiers using Plurality Voting is best for recognizing simple activities. For recognizing activities like walking forward, walking left, walking right, upstairs, downstairs, run forward, jumping, sit on a chair, standing Mi Zhang [15] proposed a Bag-of-Features approach which builds activity models using histograms of primitive symbols. Ten participants with activities in consonance with biking, stairs up, driving, lying, running, sitting, standing and walking using accelerometer sensor data was indicated by S. Dernbach et. al. [16]. Allocation of the Smartphone for collecting sensor data was user defined and sampling frequency was set to 80 Hz. Using Multi-layer Perceptron, those activities were classified acquiring accuracy of 93%. Besides, these complex activities such as, making coffee, making tea, toileting, ironing, drinking, making a meal, applying makeup, shaving and so on are recognized by Tao [17] by building activity models that mine a set of Emerging Patterns from a sequential activity trace. Activities that are required for preparing work, going shopping, doing housework such as office work, commuting, lunch, dinner are identified by Tam Huynh et. al. [18] using topic model. A combination of simple and complex activities such as cooking, cleaning, medication, sweeping, washing hands, and watering plants were recognized by Stefan et. al. [16]. They used fixed time window length with an overlap of one half of the time window length.

III. Human Activity Recognition Methods

In this section, we described the data generalization process, activities we are intending to identify and feature selection approaches.

A. Data Collection

We used a Xiomi Redmi 4A Smartphone that includes tri-axial accelerometer sensor and tri-axial gyroscope sensor to collect the user data. Both male and female data subject placed the Smartphone on their hand position, tied with a band while performing enlisted ten activities. They performed each activity for a specific period of time. Our experimental datasets contain twenty thousand instances as well as sensors data for male and female subject respectively. The android application that controlled the data collection process has a user interface to interact with the data subject and ask for permission to use sensor data so that no privacy violation occurs. In addition, the android application allowed data subjects to start or stop data collection. It also allowed data subjects to control the data collection for specific activities. We collected sensor data in 1 Hz of frequency for all activities.

B. Activities to be Identified

Human activities are mainly divided into two categories, they are, impulsive and steady activities. In our research, we considered both impulsive and steady activities. We intended to classify walking, walking upstairs, walking downstairs, jogging, typing and writing as impulsive activities and standing, sitting, lying, using toilet activities as steady activities in respect to the motion type of hand. These

activities are regularly performed by common people on daily basis. It is clear from their categories that, all the patterns are distinctive either on the periodic structure or in magnitude. Impulsive activities are distinguishable in periodic nature of oscillation and steady activities are distinguishable in the magnitude of the sensors data. Here, sensors data of tri-axial accelerometer and tri-axial gyroscope plotted for jogging as an instance of impulsive and sitting as an instance of steady activity for both male and female subjects.

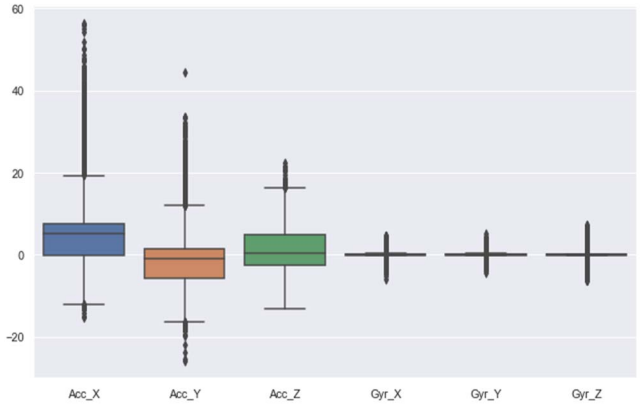


Fig 1: Frequency distribution of features of the male dataset in box-plot

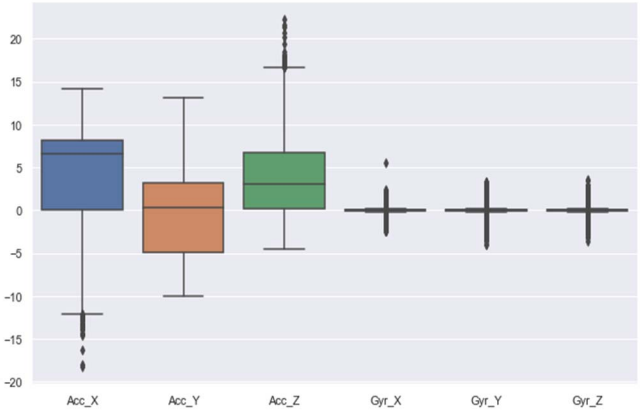


Fig 2: Frequency distribution of features of the female dataset in box-plot

C. *Pre-processing and Feature Selection*

In supervised learning, many candidate features are introduced to represent a better domain. But, many of these features are limited extent or utterly superfluous or irrelevant. Minimally sized feature subset drastically reduces the learning methods and running time of learning algorithm. As human activities are impulsive and non-impulsive as well as steady, we needed to select such features of sensors data that can differentiate between impulsive and steady activities. Since, we used an android application to amass data, we had intense fear to experience faulty data issue due to hardware or application error. Thus, we eradicated null values, outliers, and inconsistent instances. Since, tri-axial accelerometer and tri-axial gyroscope uses meter per square second and radian per second as unit respectively, thus we practiced the standardization approach in both dataset. And, Principle Component Analysis a.k.a. PCA is a technique to find features that are a subset of actual features to make a reasonable tradeoff to reduce or selection of features. Thus, we applied PCA technique on our dataset and we found out that, tri-axial accelerometer and tri-axial gyroscope sensors both have significant variations while performing orthogonal transformations, called rotations of the original features. We also found out that, all distributions put differently multivariate normal are sufficient enough to characterize the activities. Here is the plotting of variances of experimented features in a cumulative way after analyzing with PCA approach.

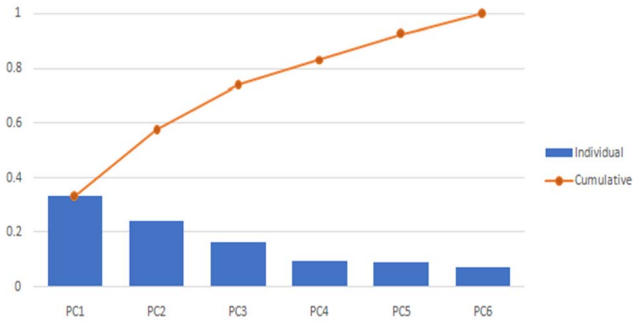


Fig 3: Individual and cumulative variance of principle components on male dataset

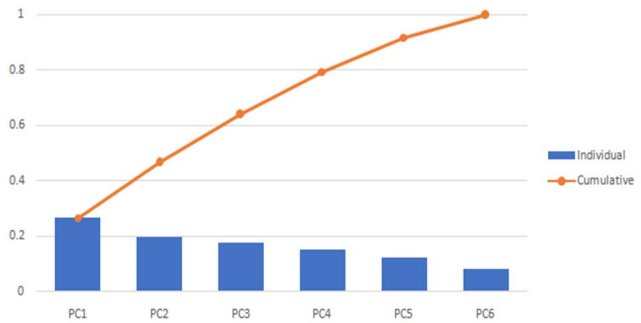


Fig 4: Individual and cumulative variance of principle components on female dataset

D. *Applied Machine Learning Algorithms*

To classify a data set, there are diverse machine learning algorithms. But, we applied Dense Neural Network a.k.a. DNN model and six base level classifiers naming Support Vector Machine a.k.a. SVM, Decision Tree a.k.a. DT, K-Nearest Neighbor a.k.a. KNN, Random Forest a.k.a. RF, Naïve Bayes a.k.a. NB and Logistic Regression a.k.a. LR.

Exercised DNN is composed of three or more layers (one input layer, one output layer and one or more hidden layers). Input layer receive the signal, output layer makes a prediction or decision about the input and the hidden layers are the true computational engine of the DNN. Formally, one-hidden-layer DNN is a function,

$$f: R^D \rightarrow R^D \quad (1)$$

Where, D is the size of input vector, X and L are the size of the output vector of f(x), such that, in matrix notation,

$$f(x) = G(b^{(2)} + W^{(2)} (S(b^{(1)} + W^{(1)}x))) \quad (2)$$

Where, bias vectors b(b(1),b(2)); weight matrices W(1), W(2) and activation functions G and S. The vector,

$$h(x) = \emptyset(x) = (s(b^{(1)} + W^{(1)}x))) \quad (3)$$

Constitutes the hidden layer $W^{(1)} \epsilon R^{D*D}$ is the weight matrix connecting the input vector to the hidden layer. Each column of $W^{(1)}$, i represents the weights from the input units to the i-th hidden unit. Typical choices for S includes hyperbolic tangent function and non-linear monotonic derivative of Rectified Linear Unit a.k.a. *ReLU* for multi-classification and logistic sigmoid function for binary classification respectively. These activation functions are mathemathically noted as,

$$tanh(z) = (e^{(z)} - e^{(z)}) / (e^{(z)} + e^{(z)}) \quad (4)$$

$$logit(z)= 1 / (1 + e^{-(z)}) \quad (5)$$

$$ReLU(z) = max(0, z) \quad (6)$$

All activations are scalar-to-scalar functions but their natural extension to vectors and tensors consists in applying them element-wise (e.g. separately on each element of the vector, yielding a same-size vector. The output vector is then obtained as,

$$O(x) = G(b^{(2)} + W^{(2)} h(x)) \quad (7)$$

Naïve Bayes classifier, structured on Bayes theorem, can be prescribed as a submissive but till now a robust data mining algorithm. The tentative feasibility i.e. feasibility of the manifestation of an instance given that, another instance has been taken place is illustrates by the Bayes theorem. Mathematical depiction of Bayes theorem pursue the below form:

$$P(X|Y) = P(Y|X)P(X) / P(Y) \quad (8)$$

Where, *P(X)* symbolize the preceding feasibility of manifestation of instance X, *P(Y)* denotes the preceding feasibility of manifestation of instance Y, *P(Y|X)* gives the succeeding feasibility regarding the manifestation of event Y given the instance X is true and *P(X|Y)* demonstrate the conditioned feasibility of instance X given Y is true. But in terms of categorization problem we can narrate Y as the data sample with unidentified class, X as the hypothesis of Y belongs to category C and *P (X|Y)* as the probability that for given data sample Y the hypothesis holds.

On the other hand, one of the most influential classifier named Random Forest, can be unfold as an ensemble learning mechanism that exaggerate the conjugation of decision tree. Initially a figure of decision trees are being proposed by Random Forest where, vindication nodes allude the predictor variables. Consecutively, the concluded effect of the performance are being affected by those achieved decision trees. In spite of the fact that, advancing of exalted dissonance may be boot up by generated sub-trees but less over-fitting can be accomplished with the collation of the learning rate. In our classifier structured, we chosen to create 100 decision trees for bringing out the finest outcome. In a nut shell Random Forest operation can be introduced as:

1. Dictatorialswatch are selected from a provided dataset.
2. Decision trees are initialized to conclude prediction outcomes by employing selected swatch.
3. A deciding vote is accomplished employing the predicted results. Reckoning the most vote finalization of predictions was accomplished.

Support Vector Machine a.k.a. SVM, another well-known classifier generates differentiating hyper-plane that simplifies classification regarding subjects. In terms of classification, for a conferred figure of categories, a proficient hyper-plane is assemble which collaborate in categorizing the imminent samples. To timbre the SVM model, selection of accurate kernels as well as regularization parameter and gamma is conspicuous. Here, the kernels are integrated to ascertain the similarity scope among perverse features and gamma value helps to timbre the SVM model effectively. In general, kernels literally embellish the simplification of non-linear categorization and since binary classification are basically introduced by SVM.

In our SVM model architecture, we exercised the “rbf” kernel and a gamma value of 0.01 for which we acquired the maximum performance outcome.

Besides those classifier, another extensively applied classification approach entitled as KNN, a learning algorithm which is non-parametric, uses a dataset with the data points divided into diverse groups with the objective of predicting the association of a new sample. As it is non-parametric, without fabricating any conjecture on the data allotment of given data, it delimits a model structure. Thus, it should be one of the first choices to classify when there is little or no prior knowledge about the allocated data. KNN specifies an integer k where k>0 for a new sample. When k=11, maximum accuracy can be achieve from our model structure. For normalization process we regenerated each attribute value by the following manner,

$$New_x = X_{mean} - Old_x \quad (9)$$

Where, Old_x is an old attribute value, X_{mean} is the average of old attribute values and New_x is the new attribute value. We did it to avoid the domination of a single feature. For the normalization value of the new instance we used the following formula,

$$Normalized\ X_i = |X_i - X_{mean}| / Standard\ Deviation \quad (10)$$

Where, X_i is a new sample. It approve us to render all features to fetch them in a new level to simplify comparison between them. For discrete value, the consuetude of discrete values are computed by measuring the dissimilarity among their frequencies. Finally, KNN uses the similarity measure between the training data sample and the new data sample for the prediction.

To train our networks, we used 80% instances of our datasets and 20% instances to test classification performance. We implemented DNN with machine learning framework scikit-learn APIs. In the classifier model, we used six hidden layers where the first, second and third hidden layers has the same number of perceptron numbers which is 500, for fourth hidden layer the number is 300, for fifth it is 200 and finally for sixth hidden layer the perceptron number is 100. In DNN, the more the hidden layers, the better is the accuracy, as a consequence by using six we achieved a good accuracy [19]. For activation function we used ‘*tanh*’ i.e. the hyperbolic tangent function. As solver we used ‘adam’ classifying method that refers to a stochastic gradient based optimizer. To control the step size in updating the weights we set the

learning rate at 0.0001 with maximum 2000 iterations to prune the model within global maxima.

IV. Evaluation of Experiments

We trained 80% of our total dataset using the proposed machine learning algorithm and tested the rest 20% for both datasets. Here, confusion matrices in heat-map, representing the performance of Dense Neural Network on both male and female dataset.

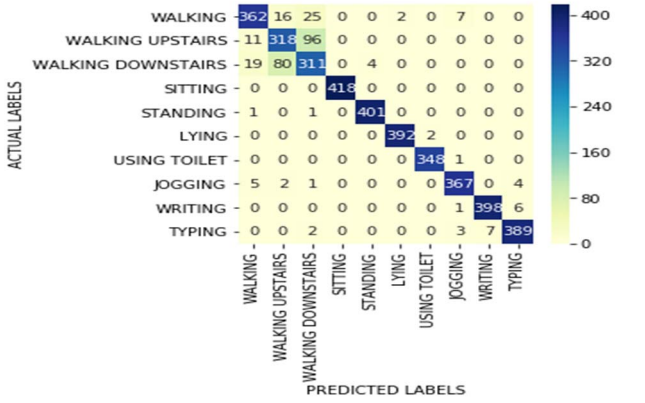


Fig 5: Performance of Dense Neural Network on male dataset

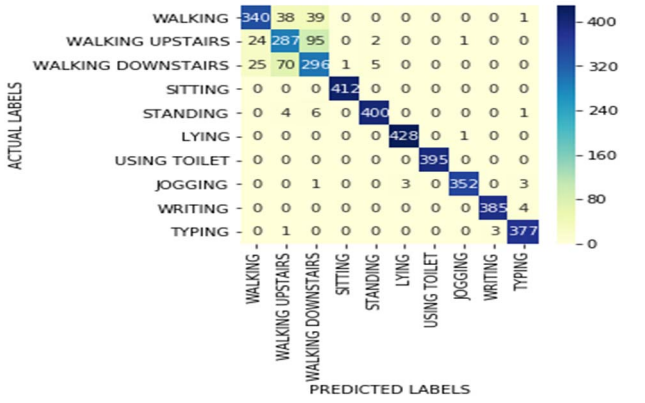


Fig 6: Performance of Dense Neural Network on female dataset

From the confusion matrix it is cleared that in case of male dataset for sitting, standing, lying and sitting in toilet all the instances were correctly classified. Great misclassification held for downstairs. Among 407 instances of walking downstairs 311 were correctly classified whereas 80 were misclassified as upstairs and 25 as walking, 1 as standing and 1 as sitting in toilet. Similarly, in case of jogging, writing and typing almost all instances has been correctly classified. From female dataset it is cleared that, all the instances of sitting and lying are correctly classified. For jogging among 418 instances only 1 instance were misclassified as typing. Greatest misclassification held for walking downstairs. Among 395 instances only 290 instances were correctly classified, and 39 were misclassified as walking, 70 as walking upstairs and 4 as standing. Though here we showed only the classification matrix for DNN as samples but we also obtained confusion matrix for other machine learning algorithm that is for SVM, DT, KNN, RF, NB and LR.

In case of SVM the accuracy we obtained for male dataset was 93.1% and for female dataset we obtained 92.25.For male dataset we found that from the 402 instances of standing, 399 were correctly classified where 2 were misclassified as downstairs and 1 as walking. The diagonal values of confusion matrix shows the number of instances that were correctly classified. On the other hand in case of female dataset from the 382 instances of standing 375 were correctly classified, 5 were misclassified as downstairs and 2 as jogging. For both dataset all instances of lying were correctly classified. For DT we obtained 89.52% accuracy for male dataset and 89.92% for female dataset. In KNN accuracy for male dataset was 93.42% and 92.52% for female dataset. We obtained highest accuracy for both male and female dataset in case of RF. It was 93.9% for male dataset and 92.85% for female dataset. From their confusion matrix it was cleared that, in male dataset for sitting all the instances were correctly classified. But for walking, upstairs and downstairs it shows greatest misclassification. And in female dataset all the instances of sitting and sitting in the toilet were correctly classified. We achieved lowest accuracy for LR. It was 77.42% for male dataset and for female dataset it was 76.97%. Biggest misclassifications were happened in recognizing walking, walking upstairs, walking downstairs and jogging for male dataset. And in case of female dataset misclassification found for walking, walking upstairs, walking downstairs and for standing. But for activities sitting in the toilet all instances were correctly classified. For NB accuracy obtained for male dataset was 85.32% and 85.45% for female dataset.

Table 1: Performance Estimation of Dense Neural Network and Base Level Classifiers

		Accuracy(%)	
		Male User Dataset	Female User Dataset
Dense Neural Network Model	DNN	**94.38**	**93.35**
Base Level Classifiers	SVM	93.1	92.25
	DT	89.52	89.92
	KNN	93.42	92.52
	RF	**93.9**	**92.85**
	NB	85.3	85.45
	LR	77.42	76.97

V. Conclusion and Future Work

In this paper, we determined, activity recognition task using two Smartphone sensors tri-axial accelerometer and tri-axial gyroscope. We predicted ten randomly performed human activities with higher accuracy for all practiced classification algorithms. We detected common activities as well as using toilet, writing and typing with Smartphone which are rare in previous works. We also showed comparison between two dataset which has not been done yet in previous studies. The data collection process was more reliable and the model developed showed better result than other Human Activity Recognition models. In future, we are planning to expand our work by predicting diseases from the predicted activities because pattern generated from daily physical activities have similarity in terms various diseases explanation. We would like increase both the number of sensors and human activities for further work. And, would also interested to extend our work using more splendid deep learning models whereas now we used only a Dense Neural Network model and some base level classifiers.

References

[1] A. K. M. Masum, A. Barua, E. H. Bahadur, M. R. Alam, M. A. U. Z. Chowdhury, and M. S. Alam, "Human Activity Recognition Using Multiple Smartphone Sensors," presented at the 2nd Int. Conf. on

Innovations in Science, Engineering and Technology (ICISET), Chittagong, Bangladesh, 2018.

[2] A. Barua, A. K. M. Masum, M. E. Hossain, E. H. Bahadur, and M. R. Alam "A study on Human Activity Recognition Using Gyroscope, Accelerometer, Temperature and Humidity data," presented at the 2nd International Conference on Electrical, Computer and Communication Engineering (ECCE 2019), Cox's Bazar, Bangladesh, 2019.

[3] S.-W. Lee, and K. Mase, "Activity and location recognition using wearable sensors," IEEE pervasive computing, vol. 1, no. 3, pp. 24-32, 2002.

[4] P. Lukowicz, H. Junker, M. Stäger, T. von Bueren, and G. Tröster, "WearNET: A distributed multi-sensor system for context aware wearables." pp. 361-370.

[5] S. C. Mukhopadhyay, "Wearable sensors for human activity monitoring: A review," IEEE sensors journal, vol. 15, no. 3, pp. 1321-1330, 2015.

[6] A. Bayat, M. Pomplun, and D. A. Tran, "A study on human activity recognition using accelerometer data from smartphones," Procedia Computer Science, vol. 34, pp. 450-457, 2014.

[7] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," ACM SigKDD Explorations Newsletter, vol. 12, no. 2, pp. 74-82, 2011.

[8] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, "Activity classification using realistic data from wearable sensors," IEEE Transactions on information technology in biomedicine, vol. 10, no. 1, pp. 119-128, 2006.

[9] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher, "Activity recognition and monitoring using multiple sensors on different body positions." pp. 4 pp.-116.

[10] E. M. Tapia, S. S. Intille, and K. Larson, "Activity recognition in the home using simple and ubiquitous sensors." pp. 158-175.

[11] E. M. Tapia, S. S. Intille, W. Haskell, K. Larson, J. Wright, A. King, and R. Friedman, "Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor." pp. 37-40.

[12] T. Brezmes, J.-L. Gorricho, and J. Cotrina, "Activity recognition from accelerometer data on a mobile phone." pp. 796-799.

[13] A. Mannini, S. S. Intille, M. Rosenberger, A. M. Sabatini, and W. Haskell, "Activity recognition using a single accelerometer placed at the wrist or ankle," Medicine and science in sports and exercise, vol. 45, no. 11, pp. 2193, 2013.

[14] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data." pp. 1541-1546.

[15] M. Zhang, and A. A. Sawchuk, "Motion primitive-based human activity recognition using a bag-of-features approach." pp. 631-640.

[16] S. Dernbach, B. Das, N. C. Krishnan, B. L. Thomas, and D. J. Cook, "Simple and complex activity recognition through smart phones." pp. 214-221.

[17] T. Gu, Z. Wu, X. Tao, H. K. Pung, and J. Lu, "epsicar: An emerging patterns based approach to sequential, interleaved and concurrent activity recognition." pp. 1-9.

[18] T. Huynh, M. Fritz, and B. Schiele, "Discovery of activity patterns using topic models." pp. 10-19.

[19] J. Mantyjarvi, J. Himberg, and T. Seppanen, "Recognizing human motion with multiple acceleration sensors." pp. 747-752.

GRU-based Attention Mechanism for Human Activity Recognition

Md. Nazmul Haque
Institute of Information Technology
University of Dhaka
Dhaka, Bangladesh
bsse0635@iit.du.ac.bd

M. Tanjid Hasan Tonmoy
Dept. of CSE
University of Dhaka
Dhaka, Bangladesh
2015-116-770@student.cse.du.ac.bd

Saif Mahmud
Dept. of CSE
University of Dhaka
Dhaka, Bangladesh
2015-116-815@student.cse.du.ac.bd

Amin Ahsan Ali
Dept. of CSE
Independent University of Bangladesh
Dhaka, Bangladesh
aminali@iub.edu.bd

Muhammad Asif Hossain Khan
Dept. of CSE
University of Dhaka
Dhaka, Bangladesh
asif@du.ac.bd

Mohammad Shoyaib
Institute of Information Technology
University of Dhaka
Dhaka, Bangladesh
shoyaib@du.ac.bd

***Abstract*—Sensor data based Human Activity Recognition (HAR) has gained interest due to its application in practical field. With increasing number of approaches incorporating feature learning of sequential time-series sensor data, in particular the deep learning based ones has performed reasonably in uniform labeled data distribution scenario. However, most of these methods do not capture properly the temporal context of time-steps in sequential time-series data. Moreover, the situation becomes worse for imbalanced class distribution which is a usual case for HAR using body-worn sensor devices. To solve this issues, we have integrated hierarchical attention mechanism with recurrent units of neural network in order to obtain temporal context within the time-steps of data sequence. The introduced model in this paper has achieved better performance with respect to the well-defined evaluation metrics in both uniform and imbalanced class distribution than the existing state-of-the-art deep learning based model.**

***Index Terms*—Human Activity Recognition, Attention Mechanism, Gated Recurrent Unit**

I. Introduction

Human Activity Recognition(HAR) is a domain of research aimed at recognizing human actions and movements from a series of observations. The increasing public adoption of smart devices with sensors such as accelerometer and gyroscope has created the opportunity to organize considerable amount of sensor data for classification of human activity. The research activities focused on HAR incorporates the compilation of sensor readings into sequential time-series data and develops models for recognition of activities by analyzing the acquired sequential sensor readings. HAR poses a variety of promising application domain which includes physical activity annotation in the field of medical data analysis [1], personal assistant system [2], augmented and virtual reality [3] and many others.

In the past years, the state-of-the-art solutions to HAR mainly based on the traditional machine learning techniques. This techniques mainly depend on heuristic based hand-crafted feature engineering that rely on low level representations. As the traditional machine learning models use low level representations, it lacks the characteristics of generalization [4], [5]. High level abstraction along with low level representations are necessary for a likely solution to this generalization problem.

Deep learning based methods deal with both low and high level representations of data. Therefore, recently two variants of deep learning methods namely convolutional [6] and recurrent [7] neural network models are become dominant over traditional methods in terms of performance. For example, to identify HAR, convolutional neural network(CNN) is used in [8] and [9] and recurrent neural network(RNN) is used in [7] and [10].

Although deep learning based methods show promising performance, conventional sliding window based approach for CNN is unable to fully capture the temporal context of the sensor reading [11] which is required for better classification of activities. For sequence data, RNN performs better than CNN in most cases as it captures sequence information [12]. However, RNN faces long term dependency problems [13] when the sequence is long enough. Note that, the sequence information found in HAR data is usually long. So, it is necessary to capture the long term dependency information for better classification.

Gated Recurrent Unit (GRU) is a variant of RNN which incorporate long term dependency information [14]. It is expected that GRU will perform better in case of HAR data as it is able to capture temporal context of sensor data. It is noteworthy to mention here that, all temporal context are not equally important for classification, some are more important than others. Hence, it is necessary to give more attention to the important temporal context than others. Moreover, during the acquisition of HAR data, usually the training data found for different activities are not equal which causes class imbalance problem. The emphasis on important temporal context also helps to solve this problem.

To capture the nature of different continuous movements

and to extract the salient features, in this paper, we propose an attention mechanism based GRU model architecture. This architecture plays a crucial role in capturing context of the sensor reading and extracts the class imbalance tolerance characteristics. The main contributions of this paper are as follows:

- We propose to use a hierarchical temporal attention with GRU for capturing important temporal contexts.
- The hierarchical model propose to use here is parallelizable.
- The model is able to handle class imbalance problem.

The paper describes related work in section II where different approaches for recognition of human activity are discussed. Section III has been used for describing the proposed methodology. Section IV contains the results as well as interpretations of the outcome of the proposed method. Section V concludes the paper.

II. Related Work

The most referred work proposed in [15] use fast Fourier transform algorithm for feature extraction, useful in recognizing different activities, that produce satisfactory results with numerous sensors set on distinctive parts of the body in conjunction with various data mining algorithms. Different approaches like K-nearest neighbors [16], decision trees [17], multi-class support vector machine [18] are used to classify human activities. All of these approaches require the use of hand-crafted features and show poor results for classifying in similar type of activities like walking down and walking up.

In the modern age, deep learning has become prominent in the area of learning models that represent features from low-level to high-level abstraction used in [4], [5] which allow to extract features automatically without hand-crafted feature engineering. A common form of neural network called fully connected neural network (FCNN) with Principal Component Analysis based feature technique is used in [8] and [9] for HAR and sensor data. But FCNN is very expensive in terms of memory (weights) and computation (connections). It also has a great chance of overfitting problem as every node is connected with every node in every layer. To extract additional features a new technique called Shift-invariant sparse coding [9] was proposed and used in combination with FCNN and handcrafted features. Convolutional neural network (CNN or convnet) [19] with dropout [20] for reducing overfitting is a recent breakthrough for feature extraction. It is used by [21] in gesture recognition that give state-of-the-art result. A hierarchical model using convnets is proposed in [22]. To recognize human activity for unlabeled as well as labeled data, [23] used semi-supervised convnet model to learn discriminative hidden features. Where convnet learns to recognize features of an object and combine these features to recognize larger object. Recurrent Neural Networks (RNN) based appoach proposed in [7] to recognize human activities and abnormal behaviour, shows some promise but leaves room for improvements. When the sequence is long, RNN faces log term dependency problems. To solve this problem, a combination of convnet and long short term memory [10] is used in [24] that outperforms other models on the KTH dataset. Gated Recurrent Units [25] are a variant of RNN that also addresses long term dependency issues. Adam optimizer [26] is a popular choice for training such neural networks.

Attention mechanism, introduced for sequence to sequence tasks tasks such as neural machine translation [27], speech recognition [28] has also been used for classification tasks in the domain of natural language processing [29]. The context vector computed by attention helps the network to learn where to focus on the representation generated by the encoder part for generating output sequence at each time step instead of compressing entire sequence to a fixed vector at once. Simplified form of attention mechanism [30] has been proposed for feed forward network which captures some long term dependencies.

The approaches described so far for activity recognition fail to capture the temporal context of sensor reading at different time steps of activity data which is required for better accuracy and generalization. Another approach proposed by [31] uses attention mechanism on top of a complex DeepConvLSTM architecture for finding relevant temporal context for activity recognition. In this work attention score is generated by applying attention after convolutional and pooling layers in DeepConvLSTM. This score does not reflect the hierarchy of simple features detected from raw sensor data and complex features detected from hidden state outputs in case of RNN (or from deeper layers in CNN). For finding relevant features for activity recognition, feature selection approaches used in [32], [33] can be applied. However, In this work, we propose an attention mechanism with GRU which distills more complex features that would be helpful for better classification.

III. Proposed Method

The proposed method combines several building blocks for constructing the network. We use Gated Recurrent Units, two different types of attention mechanism which are described in the following sections.

A. Gated Recurrent Unit

GRUs are a variant of Recurrent Networks that have been shown to be able to capture long term dependencies in temporal data while not suffering from similar vanishing gradient problem as regular RNN and requiring fewer parameters than LSTM. Hidden state $h_{<t>}$ calculation is based on (3) with the input vector $X_{<t>}$ and previous hidden state $h_{<t-1>}$ going through update and reset gates in (1) and (2).

$$Z_{<t>} = \sigma(W_{zx} \cdot X_{<t>} + W_{zh} \cdot h_{<t-1>} + b_z) \quad (1)$$

$$\Gamma_t = \sigma(W_{\Gamma x} \cdot X_{<t>} + W_{\Gamma h} \cdot h_{<t-1>} + b_\Gamma) \quad (2)$$

$$\begin{aligned} h_{<t>} = (1 - Z_{<t>}) \odot h_{<t-1>} + \\ Z_{<t>} \odot \tanh(W_{hx} \cdot X_{<t>} + \\ W_{hh} \cdot (h_{<t-1>} \odot \Gamma_t) + b_h) \end{aligned} \quad (3)$$

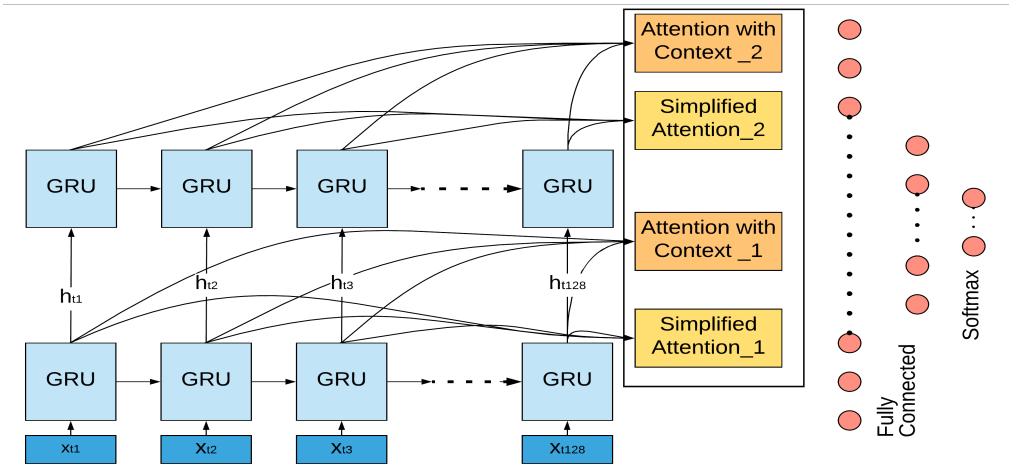


Fig. 1: Stacked 2 layer GRU model with simplified and context sensitive attention mechanism

Here, σ is used in (1) and (3) which refers to sigmoid function and $\odot$ in (3) represents element-wise multiplication.

B. Attention Mechanism for HAR

Context vector generated by attention allows a classification model to give varying levels of importance to the different temporal features generated by a recurrent network. Different variations of attention mechanisms such as simplified attention and hierarchical context based attention could be used for this type of tasks. These two types of attention could be used separately or a combination of their learned attention scores may be used. We take the latter approach.

1) Attention With Context: Because of susceptibility of sensor data to noise, context of a sensor value in relation to data on other time steps are helpful for creating a representation of data for classification. By using attention on both layers of stacked GRU, the model is able to exploit hierarchy of temporal features in the sensor data.

$$e_{<t_i>} = \tanh\left(W_{ac} \cdot h_{<t_i>} + b_{ac}\right) \tag{4}$$

$$\alpha_{<t_i>} = \frac{\exp(e^T_{<t_i>} \cdot e_s)}{\sum_t \exp(e_{<t_i>} e_s)} \tag{5}$$

$$c_{<i>} = \sum_t \alpha_{<t_i>} h_{<t_i>} \tag{6}$$

W_{ac} and b_{ac} in (4) are parameters to be learned. e_s in (5) allows the preservation of context information and is learned jointly when training the network. A summation of the relative weights of the time steps is generated as the context vector in (6).

2) Simplified Attention: Sensor data of human activities differ from natural language data in that words are context dependent in most cases where sensor data may not depend on the context. Simplified attention does not have to learn the e_s parameter.

$$e_{<t_i>} = \tanh\left(W_{as} \cdot h_{<t_i>} + b_{as}\right) \tag{7}$$

$$\alpha_{<t_i>} = \frac{\exp(e_{<t_i>})}{\sum_t \exp(e_{<t_i>})} \tag{8}$$

Context vector is obtained using (6) using the relative weight $\alpha_{<t_i>}$ obtained from (8)

C. Proposed Model Architecture

We propose a two layer stacked GRU architecture with attention applied to the hidden state outputs of each recurrent layer. The stacked layers help to learn more complex features from the sensor data. The attention scores from both layers are concatenated to create a hierarchy context vectors before feeding it to the densely connected layers. Both simplified and context-based attention scores are computed independently. Batch normalization is used before applying attention and after the concatenation of the the attention scores. There are three fully connected layers after attention module. The first two layers are used to learn respective weights for the different types of features obtained from the attention modules. Rectified Linear Unit (ReLU) is used for the activation of these layers. Dropout is also performed for regularization with probability d_1 and d_2. Dropout prevents a neural network from becoming heavily dependent on a specific weight of a single neuron by turning off a fraction of the neurons randomly during training. For classifying human activities, softmax activation is applied to the final layer. The model architecture is illustrated in Fig. 1. We train the model with learning rate α and decay factor λ.

Note that, a simpler model using a single GRU layer with same attention mechanism may also be used. Such a model requires fewer parameters to learn and resulting in less computational complexity. However, such simpler model may result in lower performance. In the proposed model architecture, sequential computation overhead may be reduced by computing the attention scores as well as the hidden states for the second GRU layer in parallel. In this case we can achieve better performance with much less time.

We use attention mechanism with GRU. Attention score may be calculated on derived features of CNN layers which may lead to the loss of context information. In comparison, using attention mechanism on the hidden states computed directly from sensor data facilitates better learning of the temporal contexts. In our proposed model, we use two types of hierarchy: firstly, hierarchy with two different types of attention (e.g. simplified and with context) and secondly, attentions from multiple layers of GRU. Using such a hierarchy of attention allows the incorporation of less complex features (low level information) with more complex features (high level representation of features) required for classification.

IV. Experiment Result and Description

A. Dataset Description

In this experiment, we use the benchmark HAR dataset [34] which provides time series information of six different activities(walking, walking_upstairs, walking_downstairs, sitting, standing and laying). Two types of sensor namely accelerometer and gyroscope are used to capture these information and randomly partitioned into 70% train and 30% test sets. Data has been preprocessed with noise-filters and sampled in fixed length sliding windows of 2.56 sec and 50% overlap which yields 128 readings per window. The training data is generated with a total of 7352 examples incorporating data of 21 randomly selected individuals. On the other hand, test set is composed of 2947 examples incorporating the remaining 9 individual subjects selected for this specific dataset.

B. Implementation Detail

During training, Adam optimizer is used with its default parameters (β_1=0.9 and β_2=0.999) for backpropagation. Moreover, the initial α is set to 0.001 which is presented in [35] with λ = 0.2 (decay based on validation loss). The dropout probability d_1 and d_2 are set to 0.25 and 0.1 respectively in the fully connected layers.

The standard accuracy metric defined in 9 is used to evaluate the methods used in this experiment.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

where, TP, TN, FP and FN are defined as True Positive, True Negative, False Positive and False Negative respectively.

To evaluate the performances of the methods in the class imbalance scenario, we first drop half of the training data and second drop two-third of the training data. In both cases, we keep unchanged test data (mentioned in the Dataset Description subsection). For measuring the performance of the methods in the class imbalance scenario, we use area under the curve (AUC) as an evaluating metric in this case.

We also performed five fold training (in each case, we consider 17 individuals information of all classes except the imbalanced class) and test (in each case, test data is fixed with 9 individuals as mentioned in the dataset description subsection) with 50 percent dropout of the training data of a specific class that is considered as an imbalanced class.

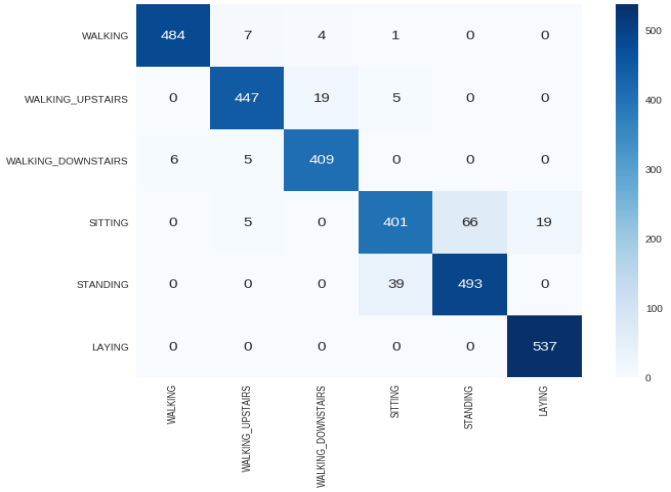


Fig. 2: Confusion Matrix when half of the training data for class 'Walking Downstairs' is dropped for stacked GRU model

C. Result and Discussion

In this experiment, we have compared our proposed model architectures with a state-of-the-art baseline CNN based method [6]. The column two of TABLE I presents the comparison according to the experimental setup described in the HAR dataset. It is observed from this table that in terms of accuracy, our proposed stacked GRU with attention performs better from both the proposed simplified model (with GRU + Attention) and the baseline CNN based method. Such an improvement is expected as we consider hierarchy in the attention model and two layer stacked GRU.

We have demonstrated the impact of class imbalance in TABLE I (Column 3-8) and TABLE II for different methods. In both cases (1/2 and 2/3 class drop), our stacked GRU with attention wins for all classes in terms of accuracy. Besides these, to test the generalization ability of the proposed method, we perform five fold training and test as mentioned in the implementation detail subsection. These results is demonstrated in TABLE III. In this case, we consider AUC as a performance metric as it performs better for imbalance dataset [36]. From this table, it is also observed that when imbalanced is injected, the proposed method wins for every classes individually which demonstrated the superiority of the proposed method.

Note that, the improvement of the proposed method with the baseline CNN model observed in the aforementioned tables is not that significant. The reason for such slight improvements can be explained by the simplicity of the dataset. In the dataset, the activities are separable by a large margin which is shown in Fig. 4. The figure is generated with dimension reduction using t-SNE [37]. So that when we create class imbalance,there is no significant drop in accuracies for both of the methods and the difference between these two methods in terms of accuracy also remains small. We believe, the improvement will be much larger for more challenging dataset.

We demonstrated the confusion matrix of the proposed method and baseline CNN in Fig. 2 and 3 respectively. The

TABLE I: 1/2 Drop of Specific Class Data in Training Set and Measurement using given Test Data in terms of accuracy (%)

Model Architecture	Performance (Accuracy %)	Class - 1	Class - 2	Class - 3	Class - 4	Class - 5	Class - 6
CNN	94 .022	89.37902952	88.49677638	90.02375297	90.6345436	92.33118426	92.09365456
GRU + Attention	93.79	90.90600611	92.50084832	90.77027486	92.195453	92.36511707	92.05972175
Stacked GRU (2 Layers) + Attention	94.16355	92.02578894	91.78825925	92.53478113	91.9239905	93.07770614	94.46895148

TABLE II: 2/3 Drop of Specific Class Data in Training Set and Measurement using given Test Data in terms of accuracy (%)

Model	Class - 1	Class - 2	Class - 3	Class - 4	Class - 5	Class - 6
Baseline (CNN)	90.56667798	89.68442484	89.41296233	89.88802172	91.44893112	91.58466237
2-Stacked GRU + Attention	90.73634204	91.38106549	92.26331863	92.02578894	92.56871395	93.31523583

TABLE III: Subject-wise Class Drop in K-Fold and Measurement using given Test Data in terms of AUC

Model Architecture	k - fold	Class - 1	Class - 2	Class - 3	Class - 4	Class - 5	Class - 6
Baseline (CNN)	0	0.989897104	0.988005576	0.989058376	0.990233114	0.986666755	0.992600584
	1	0.982893556	0.986482948	0.984618942	0.981161394	0.99192285	0.984878402
	2	0.980885725	0.987111175	0.98570861	0.987078789	0.99061635	0.983950708
	3	0.989421716	0.988371167	0.984282503	0.983464878	0.99116295	0.982648438
	4	0.990690272	0.990301677	0.99302656	0.986080276	0.987868021	0.989988681
	Average	**0.9867576746**	**0.9880545086**	**0.9873389982**	**0.9856036902**	**0.9896473852**	**0.9868133626**
Stacked GRU (2 Layers) + Attention	0	0.9891089821	0.9917239879	0.9937438938	0.9885901347	0.9892608509	0.9915665231
	1	0.9898042512	0.9845693577	0.9922561365	0.9879438003	0.9901914546	0.9906566782
	2	0.992904892	0.990711794	0.992136949	0.988169278	0.991100119	0.992741053
	3	0.988042117	0.992132516	0.990958787	0.987103942	0.989934521	0.993510194
	4	0.991783724	0.993914724	0.989005982	0.992895329	0.992081285	0.992756454
	Average	**0.9903287933**	**0.9906104759**	**0.9916203497**	**0.9889404968**	**0.9905136461**	**0.9922461805**

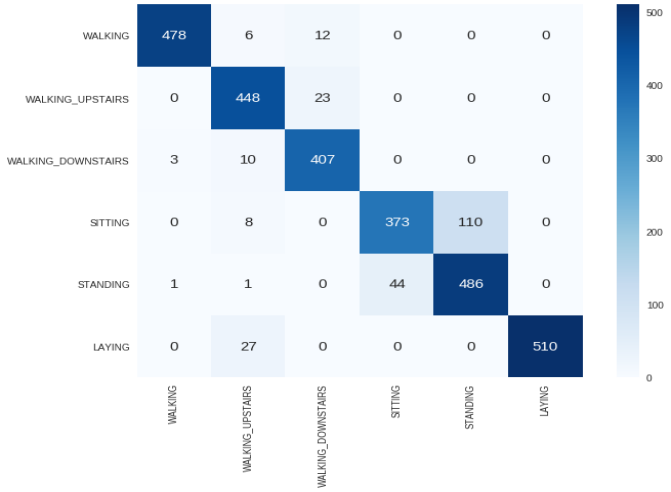


Fig. 3: Confusion Matrix when half of the training data for class 'Walking Downstairs' is dropped for CNN model

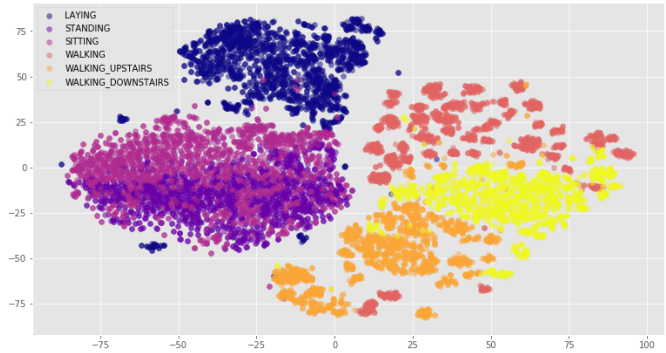


Fig. 4: Visualization of the separation of activities

confusion matrix indicates that misclassification occurs rarely and also within the vicinity of ground truth class even in imbalanced class scenario. However, from these two figures, it is observed that the proposed method has fewer misclasssification compare to the baseline CNN.

Based on the results in this experiment, it is evident that the proposed method is better in terms of accuracy and AUC metric. This method is able to produce such result by giving the attention to both simple and complex features which are learned from the proposed combination of two attention mechanisms.

V. Conclusion

The proposed method performs reasonably well for the recognition of human activity while not being sensitive to class imbalance in the training data as it learns temporal features using attention mechanism. More parrallelizable models could be developed in the future by constructing embedding from the temporal data. In this paper, we intend to show the effectiveness of the proposed mechanism using a benchmark HAR dataset. We believe the proposed method will also perform better for more complex datasets, collected from heterogeneous devices that have more varied class distribution in terms of activity duration and feature complexity which we will address in future.

ACKNOWLEDGMENT

This research is supported by the University Grants Commission, Bangladesh under the Dhaka University Teachers Research Grant No Reg/Admin-3/54292-94.

REFERENCES

[1] H. Song, D. Rajan, J. J. Thiagarajan, and A. Spanias, "Attend and diagnose: Clinical time series analysis using attention models," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[2] F. Montalto, C. Guerra, V. Bianchi, I. De Munari, and P. Ciampolini, "Musa: Wearable multi sensor assistant for human activity recognition and indoor localization," in *Ambient Assisted Living*, pp. 81–92, Springer, 2015.

[3] D. Van Krevelen and R. Poelman, "Augmented reality: Technologies, applications, and limitations," *Vrije Univ. Amsterdam, Dep. Comput. Sci*, 2007.

[4] Y. Bengio *et al.*, "Learning deep architectures for ai," *Foundations and trends® in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009.

[5] L. Deng, "A tutorial survey of architectures, algorithms, and applications for deep learning," *APSIPA Transactions on Signal and Information Processing*, vol. 3, 2014.

[6] C. A. Ronao and S.-B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," *Expert Syst. Appl.*, vol. 59, pp. 235–244, Oct. 2016.

[7] D. Arifoglu and A. Bouchachia, "Activity recognition and abnormal behaviour detection with recurrent neural networks," *Procedia Computer Science*, vol. 110, pp. 86–93, 2017.

[8] T. Plötz, N. Y. Hammerla, and P. L. Olivier, "Feature learning for activity recognition in ubiquitous computing," in *Twenty-Second International Joint Conference on Artificial Intelligence*, 2011.

[9] C. Vollmer, H.-M. Gross, and J. P. Eggert, "Learning features for activity recognition with shift-invariant sparse coding," in *International conference on artificial neural networks*, pp. 367–374, Springer, 2013.

[10] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, pp. 1735–1780, Nov. 1997.

[11] Y. Meng and A. Rumshisky, "Context-aware neural model for temporal information extraction," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2018.

[12] W. Yin, K. Kann, M. Yu, and H. Schütze, "Comparative study of cnn and rnn for natural language processing," *arXiv preprint arXiv:1702.01923*, 2017.

[13] Y. Bengio, P. Frasconi, and P. Simard, "The problem of learning long-term dependencies in recurrent networks," in *IEEE international conference on neural networks*, pp. 1183–1188, IEEE, 1993.

[14] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[15] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *International conference on pervasive computing*, pp. 1–17, Springer, 2004.

[16] W. Wu, S. Dasgupta, E. E. Ramirez, C. Peterson, and G. J. Norman, "Classification accuracies of physical activities using smartphone motion sensors," *Journal of medical Internet research*, vol. 14, no. 5, 2012.

[17] A. M. Khan, "Recognizing physical activities using wii remote," *International Journal of Information and Education Technology*, vol. 3, no. 1, p. 60, 2013.

[18] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine," in *International workshop on ambient assisted living*, pp. 216–223, Springer, 2012.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[20] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 06 2014.

[21] S. Duffner, S. Berlemont, G. Lefebvre, and C. Garcia, "3d gesture classification with convolutional neural networks," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5432–5436, IEEE, 2014.

[22] J. Yang, M. N. Nguyen, P. P. San, X. L. Li, and S. Krishnaswamy, "Deep convolutional neural networks on multichannel time series for human activity recognition," in *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[23] M. Zeng, T. Yu, X. Wang, L. T. Nguyen, O. J. Mengshoel, and I. Lane, "Semi-supervised convolutional neural networks for human activity recognition," in *2017 IEEE International Conference on Big Data (Big Data)*, pp. 522–529, IEEE, 2017.

[24] M. Baccouche, F. Mamalet, C. Wolf, C. Garcia, and A. Baskurt, "Sequential deep learning for human action recognition," in *International workshop on human behavior understanding*, pp. 29–39, Springer, 2011.

[25] K. Cho, B. van Merrienboer, Ç. Gülçehre, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *CoRR*, vol. abs/1406.1078, 2014.

[26] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2015.

[27] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *CoRR*, vol. abs/1409.0473, 2015.

[28] J. K. Chorowski, D. Bahdanau, D. Serdyuk, K. Cho, and Y. Bengio, "Attention-based models for speech recognition," in *Advances in neural information processing systems*, pp. 577–585, 2015.

[29] Z. Yang, D. Yang, C. Dyer, X. He, A. J. Smola, and E. H. Hovy, "Hierarchical attention networks for document classification," in *HLT-NAACL*, 2016.

[30] C. A. Raffel and D. P. W. Ellis, "Feed-forward networks with attention can solve some long-term memory problems," *CoRR*, vol. abs/1512.08756, 2015.

[31] V. S. Murahari and T. Plötz, "On attention models for human activity recognition," in *Proceedings of the 2018 ACM International Symposium on Wearable Computers*, pp. 100–103, ACM, 2018.

[32] N. X. Vinh, S. Zhou, J. Chan, and J. Bailey, "Can high-order dependencies improve mutual information based feature selection?," *Pattern Recognition*, vol. 53, pp. 46–58, 2016.

[33] S. Sharmin, M. Shoyaib, A. A. Ali, M. A. H. Khan, and O. Chae, "Simultaneous feature selection and discretization based on mutual information," *Pattern Recognition*, vol. 91, pp. 162–174, 2019.

[34] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," 01 2013.

[35] M. D. Zeiler, "Adadelta: an adaptive learning rate method," *arXiv preprint arXiv:1212.5701*, 2012.

[36] C. X. Ling, J. Huang, and H. Zhang, "Auc: a better measure than accuracy in comparing learning algorithms," in *Conference of the canadian society for computational studies of intelligence*, pp. 329–341, Springer, 2003.

[37] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.

Discretion way for Bleeding Detection in Wireless Capsule Endoscopy Images

Abdullah Al Mamun
Electrical and Electronic Engineering
Pabna Univrsity of Science and Technology
Pabna, Bangladesh
mamun130203@gmail.com

Md. Sohag Hossain
Electrical and Electronic Engineering
Pabna Univrsity of Science and Technology
Pabna, Bangladesh
msh22eee@gmail.com

Md. Galib Hasan
Electrical and Electronic Engineering
Pabna University of Science & Technology
Pabna, Bangladesh
eeemgh@gmail.com

Md. Motaher Hossain
Electrical and Electronic Engineering
Pabna University of Science & Technology
Pabna, Bangladesh
motaher_eee@pust.ac.bd

Abstract**—** In the case of modern diagnosis technology, WCE (Wireless Capsule Endoscopy) is one of the most efficient and effective technologies for the diagnosis of Gastrointestinal tract noninvasively. Bleeding in the GI tract is one of them. Since the duration of the WCE video is so long, it contains a large number of images. Consequently, it makes a huge burden for the physicians to recognize affected portion in the real required time and it is the main obstacle for using the WCE in the wider application. For releasing the extra burden of the clinicians, we have proposed a novel automatic computer-aided system for detecting bleeding in the GI tract. In our proposed method, we have applied some special preprocessing for obtaining the required informative portions of blood. A noble Weighted k-Nearest-Neighbor (WKNN) classifier has applied for separating distinct object with some statistical features in HSV color space. We have successfully achieved accuracy 98.8%, sensitivity 99%, specificity 99%, precision 95%, negative predicted value 99% and F1 score 97% after extensive experiment which will outperform some of the existing researches. More importantly, its computational time is very compared to others.

Keywords— Digestive tract, HSV, WKNN, Wireless Capsule Endoscopy, QSVM, Bleeding Detection

I. INTRODUCTION

Gastrointestinal tract diseases such as small intestine, large intestine, stomach etc. have become the largest problem in our modern era of life. A lot of peoples around the word are suffering seriously from this type of diseases including Bangladesh. In an observation, it has observed that approximately 27.8% peoples in Bangladesh are suffering from and die for bowel taint [1]. In another observation it is observed that 1.6 million peoples of American are suffering from and die for IBD (inflammatory bowel disease). This IBD has become epidemic since 2011. Diagnosis for the IBD has increased up to 70,000. Approximately 80,000 babies are suffered from CD (Crohn's disease) or UC (ulcerative colitis) [2] . The IBD has become the most common concern in the whole world [3]. So, the process of curing GI tact problem should be developed in a rapid way. The diagnosis and therapy of GI tact diseases are devolved with revolution due to the invention of fiber optic endoscopy to visualize lower and upper GI tract. But by using this procedure only small bowel, distal ileum and proximal duodenum can be visualized. For improving the endoscopy process, there are a lot of procedure are practiced like, endoscopy biopsies, X-ray, CT scan, ultrasound, endoscopy etc. but they failed to improve the process and give relief to the patients. The revolution is created in the arena of endoscopy process when CE is discovered in the year of 2000 by Given Imaging Company and employed in 2001 [4]. This procedure not only makes visible the small intestine of GI tract but also provides a new technology in which the images from GI tract are captured by a swallowed capsule without pain. In accordance with the acceptability, safety and value this technology has developed and this is certified by US FDA (Food and Drug Administration) [5]. After that CE is considered as the best possible procedure to investigate the whole small bowel for visualizing in a discretion way [6]. The use of CE is increasing day by day for visualizing the complete small vowel in a discretion way

In order to help the clinicians in diagnosis GI diseases, the process of automatic bleeding recognition from the WCE images plays a significant role [7]. Wireless transmitter, color camera, Light, image sensor and battery are the major components of WCE. To complete the whole GI tract, it takes 8 hours and captures approximately 57,000 images including 2 frames for every second. These captured images are sent to a computer with the help of a transmitter which is operated by radio-frequency [8]. Then the clinicians check out the diagnosis images for finding bleeding portion or some other disease indications which takes almost two hours to perform completely [9]. Very often it is too difficult for the clinicians to detect the symptom of the disease due to small size or unequal distribution of the objects in the image frames and also the oversight problem of the physicians. These problems courage the researchers to propose

a method for automatic bleeding detection as computer aided system[10].

A. Related Work

To release the burden of the clinicians, different types of bleeding detection technology have already employed from WCE. But they provide unsatisfactory result in specificity and sensitivity like SBI (Suspected blood indicator) technique [11], [12]. First of all a software was designed for detecting bleeding region whose sensitivity and specificity were only 21.5% and 41.8% [13]. According to [14] bleeding and ulcer has detected from of capsule endoscopy images which signifies chromaticity moment using HSI (Hue-Saturation-Intensity) color space and they got 91.50% accuracy, 94.56% specificity and 79.25 % sensitivity. After that a lot of research has been done by the researchers to develop in this arena. In [10] Probabilistic Neural Network (PNN) is proposed for bleeding detection and PNN classifier was developed for recognizing bleeding region. The result in accordance with specificity and sensitivity was 85.6% and 93.1% respectfully. In [14] they developed cluster-based sampling and create a method to study about large and disequilibrium data set in order to detecting bleeding image from CE.As multiple training data set was created in a random fashion, it reduced the loss of information. They also used SVM classifier for recognizing bleeding region which provided with better sensitivity and higher specificity. In [15] they developed a process which was able to reduce the complexity for computation by using super pixel segmentation process and also maintained higher degree of accuracy, sensitivity and specificity. The extracted the feature in the RGB space from the super pixel using red ratio and imposed into the SVM classifier. They also emphasized on the reduction of edge pixels. In [16] they have used different type of color feature like for CCV, RGB and HSV and got the accuracy and sensitivity in the purpose of bleeding detection 93.64% and 92.86% respectfully for CCV features. In [17] they have used the statistical features in the case of RGB color space for single order histogram provability. For distinguishing the bleeding image they have used neural network pattern and got 93% accuracy, 90% specificity and 96 % sensitivity. Again in [18] they have used frequency spectrum in the consideration of characteristics pattern. They have developed a technique which works on NGLCM (Normalized Gray Level Co-occurrence Matrix). They got a tremendous result in their proposed method. In [19] they have introduced two fold system. First of all they have taken color information to apply K-means clustering to obtain cluster Centre and observe the histogram. Then they have applied SVM classifier to differentiate the bleeding image and they got 95.75% accuracy in the case of YCbCr color space. In [20] they have emphasized on time consuming purpose. They have proposed feature extraction of holistic based on pixel. They have also introduced changeable color domain instead of using RGB color space directly. Statistical feature on the basis of lower and higher are used for the composite color space. Then they have applied SVM classifier to differentiate the bleeding image and they got satisfactory result.

This paper provides with a unique proposed method for bleeding recognition from WCE images in HSV color space. At first the WCE images are acquired by converting the video footage of WCE into specified format of images. In order to highlight the informative portion in the images, some prepressing is applied and then RGB color images are converted into HSV color images. After that we have extracted some statistical features and employed the data based into unique WKNN classifier. This unique proposed method is applied publicly obtainable wireless capsule endoscopy data base for testing.

II. OUR PROPOSED METHOD

Before We have used a unique method in order to detect the existence of blood in WCE images which are obtained from the WCE video footage. In this work we have applied some preprocessing steps for removing the unusual portion of images and highlighting the abnormal portion which is bleeding on this circumstance. The method that we have used is easy to calculate, time saving and give more accurate result of performance parameters. The layout of our proposed method is shown in Fig. 1.

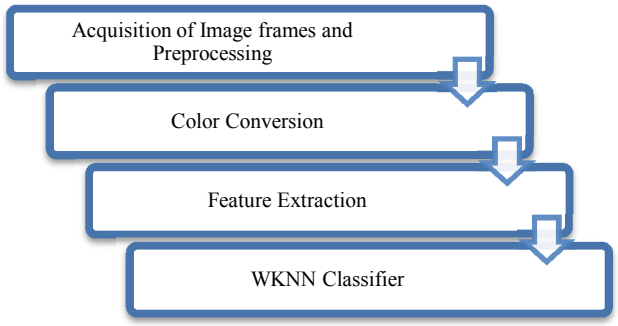


Fig. 1. Layout of our Proposed Method

A. Acquisition of Image Frame and Preprocessing

As discussed in the previous article that WCE is process of capturing the inner portion of our body but is in the video mode. So, in the initial stage of our work, we have to gather a strong database of images folder. This can be done by converting the video footage that obtained from WCE into image frame depending on the frame rate specified by the CE. This is done such that we can take necessary steps to perform any types of processing on the images, such as, preprocessing, segmentation, feature extraction classification and so on. On our work we have used PillCam®SB as CE whose frame rate is 2 and longevity is 8 hours which is sufficient to provide a enrich database for analyzing [21]. Among all of the we have used only 2393 images from the CE video to analysis which are available in [22]. After acquiring the image frames, we have to remove the black portions from the frames which don not contain any information. Firstly, for removing the black portion we took 67% of the image and cut out the remaining blackish portions from the frame which is regarded as preprocessing of the images.

B. Color Space Transform

HSV color model is associated with the three values of hue, saturation and value. HSV color model characterizes by the following three thinks [23].

- Color as hue
- On the basis of shade as saturation
- Last one is the brightness values

The ‘single hex cone model’ of color space is showing by the following Fig. 2 which describes the visualization of H, S and V parameters.

From the Fig. 2, the exterior side of the upper portion of the cone is referred to the original color and the angle of the circle is described by the H parameter. At the point on the axis of cone the value of s is 0 and it goes away from the axis if it increases the value of s. the color at the middle point of the circle is white. The value of V is like to the binary digit and 0 for black and 1 for white and goes from bottom to the center point of the circle [24].

Different statistical features of bleeding images are calculated by these corresponding values of three channel of HSV color space.

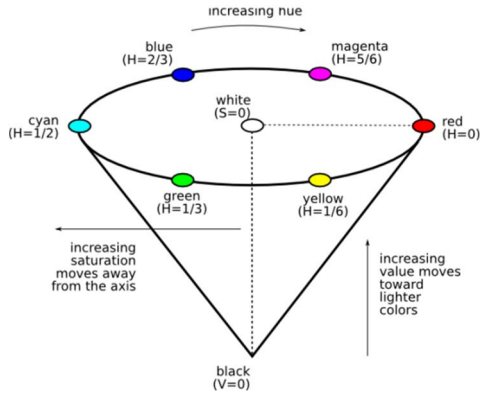


Fig. 2. Single Hex cone Model of Color Space (www.infohost.nmt.edu).

C. Bleeding Feature Selection

Feature extraction is the process of dimensional retrenchment which will represent the specified the information of a particular portion of image efficiently in a close-fitting feature vector. Over all, this vector contains some specified numerical values regarding the particular information of a determined sample. In order to extract features and calculations from the segmented image, we have used local standard deviation and mean of corresponding features of saturation, hue and value of the image as it provides with efficient results. These characterized features are applicable for the HSV color space since HSV color space is illustrated by saturation, hue and value. For that reason, we have converted the extracted image into the HSV color space and create the required feature victor. The values of hue, saturation and value of the HSV color image can be obtained by the following equations from the corresponding pixels values of three particular channel of RGB color space [25]. However, RGB is a color model which is described by three color channel such as red, green and blue. Simply RGB color space can be elaborated by every imaginable color which is formed from the above three channels. From the above notation every image has an intensity range of 0~255 of RGB elements.

$$H = arccos \frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2(R-B)(G-B)]^{\frac{1}{2}}} \quad (1)$$

$$S = \frac{\max(R,G,B) - \min(\mathrm{R},\mathrm{G},\mathrm{B})}{\max(R,G,B)} \quad (2)$$

$$V = \max(R,G,B) \quad (3)$$

For the purpose of classification, separate mean and local standard deviation is calculated from the specified values of hue, saturation and value. That's why we got a feature vector of total six specified feature. Sequentially, this vector is employed to the specified classifier in order to separate the bleeding information from the non-bleeding information.

D. WKNN Classifier

The Weighted K-Nearest Neighbor is the extended version of KNN and it is the observation of training data which is close enough to the new surveillance (b, a) and have the higher rate of weight which is apart from (b, a). Apart from the KNN the impact of the k is always analogous for all the neighbors though it may widely vary for the individual case. In order to fulfill this purpose, the distance has to transform into analogous measure which is referred to weights.

After this analogous measurement, each of latest point (b, a) will be categorized with the highest weighted added value as fellow,

$$\text{Max}_{\,m}\left(\sum_{i=1}^{k} k\left(D\left(a,a_{(i)}\right)\right) I\left(b_{(i)} = m\right)\right) \quad (4)$$

$$T = \{(a_i, b_i), i = 1, \ldots, n_M\} \quad (5)$$

Let equation (5) be the training data of a_i observation where a is the new surveillance and b have to be forecasted.

According to the function of distance $d\left(a, a_{(1)}\right)$, the value of k+1 nearest neighbor has to determine which is used to normalized the smallest k distance by the following equation (6)

$$D_{(i)} = D\left(a, a_{(i)}\right) = \frac{d\left(a,a_{(i)}\right)}{d\left(a,a_{(k+1)}\right)} \quad (6)$$

The normalized distance D(i) is transformed into weighted with kernel function as fellow

$$W_{(i)} = D(i) \quad (7)$$

Then the ultimate prediction will be

$$\hat{b} = \max_{m}\left(\sum_{i=1}^{k} w_{(i)} I\left(b_{(i)} = m\right)\right) \quad (8)$$

The main purpose of this method is to capture the degree of independency worst selection of k which results in a giant blunder classification error. Implicitly the quantity of neighbors is secluded in the weights. The value of k is adapted spontaneously if k is high enough. Neighbors with higher weight will influence the neighbors of lower weight even they have higher combination of neighbors [26].

The classification has been implemented having 10-fold validation techniques to split the samples into training and testing sets. After applying quadratic weighted k nearest neighbor classifier, we are able to separate the bleeding and non-bleeding images.

III. Result and Discussion

In the case of diagnosis like detecting bleeding or non-bleeding, the perfection of the work depends upon some major attributes. There are four possible events may be occurred at the time recognizing bleeding and non- bleeding images.

TABLE I
Extension of possible occurrence during classifying

Possible Occurrence	Extension
True bleeding detection (TP)	Detect the actual bleeding images

True non-bleeding detection (TN)	Detect the actual non-bleeding images
False bleeding detection (FP)	Detect as non-bleeding though it were bleeding
False non-bleeding detection (FN)	Detect as bleeding though it were non-bleeding

Accuracy is not the only reliable perimeter to justify the performance of work. Some other parameters like sensitivity, specificity, precision; F1score, negative predicted value etc. are the most reliable parameters to justify the performance of the work.

Sensitivity: Sensitivity detects positive disease. So, it describes the quantity or the probability of detecting the actual object what is actually needed to predict.

$$\text{Sensitivity}= \frac{Quantity\ of\ actula\ successful\ positive\ prediction}{Number\ of\ actula\ positives} \quad (9)$$

Sensitivity is the ratio of quantity of actual successful positive prediction to the number of actual positive [27].

Specificity: Specificity detects actually no disease. So, it describes the quantity or the probability of detecting the actual negative object. In short, it predicts negative when it actually negative [28].

$$\text{Specificity} = \frac{Quantity\ of\ actula\ negative\ prediction}{Number\ of\ actula\ negatives} \quad (10)$$

Specificity is the ratio of quantity of actual negative prediction to the quantity of actual negative [27].

Accuracy: Accuracy describes the measurement result of correct diagnosis test which finds out and shut out the provided condition. It can be calculated from sensitivity and specificity which is in the previous sections.

$$\text{Accuracy} = \frac{Quantity\ of\ actula\ prediction}{Total\ sample\ data} \quad (11)$$

Accuracy is the ratio of quantity of actual prediction to the quantity of total sample data [27].

Precision: precision can be defined as positive predicted value

$$Precision = \frac{Quantity\ of\ actula\ successful\ positive\ prediction}{true\ positive + false\ negative} \quad (12)$$

F1 score: In binary classification F1 refers to the test of accuracy and can be express by the following equation

$$F1\ score = 2 \times \frac{PPV \times TPR}{PPV + TPR} \quad (13)$$

Negative Predicted Value: NPV is the provability to describe the negative screening examination which truly does not contain positive objects and can be expressed by the following equation.

$$NPV = \frac{TN}{TN + FN} \quad (14)$$

As discussed earlier in our proposed method, we have used HSV color space, mean and local standard deviation of the corresponding channels or feature vector, unique WKNN classifier and the performance of the work is justified by sensitivity, specificity, precision, F1score, negative predicted value.

Table II
Performance correlation between different Characteristics features

Characteristics features	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	NPV (%)	F1 Score (%)
$\bar{V}, \sigma H\ and\ \sigma V$	92.6	86	94	74	97	20.4
$\sigma H, \sigma S\ and\ \sigma V$	93.2	86	95	76	97	36.4
$\bar{H}, \bar{S}, \bar{V}\ and\ \sigma H$	97.2	95	98	89	99	60.7
$\bar{H}, \bar{S}, \bar{V}\ and\ \sigma S$	96.8	92	98	91	98	29.3
$\bar{H}, \bar{S}, \bar{V}\ and\ \sigma V$	96.5	92	98	90	98	58.9
$\bar{S}, \bar{V}, \sigma H\ and\ \sigma S$	97.8	96	98	92	99	76.4
$\bar{S}, \bar{V}, \sigma H\ and\ \sigma V$	97.7	97	98	91	99	63.6
$\bar{V}, \sigma H, \sigma S\ and\ \sigma V$	86.5	94	96	81	99	47.2
$\bar{H}, \bar{S}, \bar{V}, \sigma H\ and\ \sigma S$	98.4	97	99	94	99	76.9
$\bar{H}, \bar{S}, \bar{V}, \sigma H\ and\ \sigma V$	98.6	99	98	93	99	70.7
$\bar{S}, \bar{V}, \sigma H, \sigma S\ and\ \sigma V$	98.5	99	98	93	99	69.2
$\boldsymbol{\bar{H}, \bar{S}, \bar{V}, \sigma H, \sigma S\ and\ \sigma V}$	**98.8**	**99**	**99**	**95**	**99**	**97**

Where, Mean Hue=$\bar{H}$, Mean Saturation= $\bar{S}$, Mean Value= $\bar{V}$, Standard deviation Hue=σH, Standard deviation Saturation=σS and Standard deviation Value= σV (We have observed many combinations of statistical features. However, we have indicated some selective combinations.)

In this case we have used 10-fold validation. The Table 2 shows the performance correlation between different combinations of specified features with performance parameters. It can be easily finalized that we got the best result of accuracy 98.8%, sensitivity 99%, specificity 99%, precision 95%, negative predicted value 99% and F1 score 97% by using mean of Hue, Saturation, value and local standard deviation of Hue, Saturation, and value combination than the other combinations. The Table II Shows the performance correlation between four color spaces by using the same classifier and provides the evidence that HSV color space is superior among all the color space mentioned in the Table II. The six performance parameters accuracy, sensitivity, specificity, precision, negative predicted value and F1score are higher in HSV color model than the others. Again, the Table III shows the performance correlation between different classifier that I have used to detect bleeding and non- bleeding images. It also gives the clarification that the WKNN classifier gives the better result among all the classifier mentioned in the Table IV like KNNN (K- nearest Neighbor) is the most commencement and nonparametric procedure in which a new surveillances settled toward the class of surveillance taking from the specified training set closest to the new surveillance for classification, QDA (Quadratic Discriminant Analysis) distributed the calculated value from the each class and without anticipating covariance of individual class is indistinguishable and Quadratic Support Vector Machine (QSVM) is a popular machine learning device that is considerd as a nonparametric scheme because of using kernel function.

Table III
Performance correlation between different color models

Color models	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	FPV (%)	F1 Score (%)
HSV	**98.8**	**99**	**99**	**95**	**99**	**97**
RGB	95.6	90	97	87	98	83
L*a*b*	98.4	97	99	94	99	73.7
YcbCr	98.2	97	98	93	99	84.35

Table IV
Performance correlation between different classifiers

Different classifiers	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	FPV (%)	F1 Score (%)
Fine KNN	98.5	95	99	95	99	80.6
QDA	96.5	91	98	94	98	75.8
QSVM	98.5	99	98	93	99	85.3
WKNN	**98.8**	**99**	**99**	**95**	**99**	**97**

In our research work, we have used some distinct classifier with validation process and modified the result by varying different perimeters. Just mentioned earlier, we have accuracy 98.8%, sensitivity 99%, specificity 99%, precision 95%, negative predicted value 99% and F1 score 97%by using 10-fold validation process with WKNN classifier. Though tremendous works have been done on bleeding detection, our work is unique and also improves in the arena of particular field. Comparison between the proposed method and the existing methods are shown in the following Table V and Table VI.

Table V
Performance correlation between different existing methods

Methods	Accuracy	Sensitivity	Specificity
CCV Feature [17]	92.86	93.64	-
Uniform LBP [9]	91.5	79.25	94.56
Binary feature vector	91	93.12	88.39
Red ratio in RGB space	94.8	97.02	93.56
Two fold System [20]	95.75	92	96.5
Different stricture of CNN [32]	97.18	96.94	97.18
Proposed Method	**98.8**	**99**	**99**

Table VI
Comparison between unique Performance parameters of existing methods

Methods	Precision	F1 Sore
Fu et al. [29]	86.17	83.51
Yuan et al. [30]	88.17	84.97
Jia et al.[31]	92.55	89.69
Jia et al.[32]	94.79	92.85
Proposed Method	95	97

From the above Table V and Table VI it is clearly shown that our proposed method is a developed version in the arena of bleeding detection in all terms of performance parameters among the existing work. In [33], they have used 2300 WCE images for analyzing the bleeding detection. Moreover, we have used 2393 images from the CE video to analysis which are available in [22] and give the strong evidence about the efficiency of our proposed method.

IV. CONCLUSION

This paper provides us with a unique proposed method for bleeding recognition from WCE images. This experiment exhibits that top most classification accomplishment is obtained with WKNN classifier with 10-fold validation technique, some essentials preprocessing and unique feature vector in HSV color space. The proposed method is analyzed for 2393 WCE images. The proposed features obtained the result accuracy 98.8%, sensitivity 99%, specificity 99%, precision 95%, negative predicted value 99% and F1 score 97%. It is examined that this statistical output of the proposed method is more efficient in term of accuracy, sensitivity and specificity than the indicated three comparable methods and capable to reduce the burden of the clinicians.

Acknowledgement
We would like to thank Topu Biswas for his special guidance and support.

REFERENCES

[1] "Death in Different Diseases." [Online]. Available: http://www.healthdata.org/bangladesh. [Accessed: 12-Jun-2018].

[2] G. G. Kaplan and S. C. Ng, "Understanding and preventing the global increase of inflammatory bowel disease," *Gastroenterology*, vol. 152, no. 2, p. 313–321. e2, 2017.

[3] G. G. Kaplan and T. Jess, "The changing landscape of inflammatory bowel disease: East meets West," *Gastroenterology*, vol. 150, no. 1, pp. 24–26, 2016.

[4] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless capsule endoscopy," *Nature*, vol. 405, no. 6785, p. 417, 2000.

[5] A. Wang *et al.*, "Wireless capsule endoscopy," *Gastrointest. Endosc.*, vol. 78, no. 6, pp. 805–815, 2013.

[6] D. G. Adler and C. J. Gostout, "Wireless capsule endoscopy," *Hosp. Physician*, vol. 39, no. 5, pp. 14–22, 2003.

[7] "The National digestive diseases information clearinghouse." [Online]. Available: www.niddk.nih.gov/health-information/digestive-diseases/gastrointestinal-bleeding/diagnosis. [Accessed: 14-Jun-2018].

[8] T. Ghosh, A. Das, and R. Sayed, "Automatic small intestinal ulcer detection in capsule endoscopy images," *Int. J. Sci. Eng. Res.*, vol. 7, no. 10, pp. 737–741, 2016.

[9] B. Li and M. Q.-H. Meng, "Computer-aided detection of bleeding regions for capsule endoscopy images," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1032–1039, 2009.

[10] G. Pan, G. Yan, X. Qiu, and J. Cui, "Bleeding detection in wireless capsule endoscopy based on probabilistic neural network," *J. Med. Syst.*, vol. 35, no. 6, pp. 1477–1484, 2011.

[11] S. Liangpunsakul, L. Mays, and D. K. Rex, "Performance of Given suspected blood indicator," *Am. J. Gastroenterol.*, vol. 98, no. 12, p. 2676, 2003.

[12] J. M. Buscaglia *et al.*, "Performance characteristics of the suspected blood indicator feature in capsule endoscopy according to indication for study," *Clin. Gastroenterol. Hepatol.*, vol. 6, no. 3, pp. 298–301, 2008.

[13] R. D. Francis, "Sensitivity and specificity of the red blood identification (RBIS) in video capsule endoscopy," in *3Rd, Int. Conf. on Capsule Endoscopy*, 2004.

[14] M. Abouelenien, X. Yuan, B. Giritharan, J. Liu, and S. Tang, "Cluster-based Sampling and Ensemble for Bleeding Detection in Capsule Endoscopy Videos," *Am. J. Sci. Eng.*, vol. 2, no. 1, pp. 24–32, 2013.

[15] Y. Fu, W. Zhang, M. Mandal, and M. Q. H. Meng, "Computer-aided bleeding detection in WCE video," *IEEE J. Biomed. Heal. Informatics*, 2014.

[16] J.-Y. Yeh, T.-H. Wu, and W.-J. Tsai, "Bleeding and Ulcer Detection Using Wireless Capsule Endoscopy Images," *J. Softw. Eng. Appl.*, 2014.

[17] S. Sainju, F. M. Bui, and K. A. Wahid, "Automated bleeding detection in capsule endoscopy videos using statistical features and region growing," *J. Med. Syst.*, 2014.

[18] A. R. Hassan and M. A. Haque, "Computer-aided gastrointestinal hemorrhage detection in wireless capsule endoscopy videos," *Comput. Methods Programs Biomed.*, 2015.

[19] Y. Yuan, B. Li, and M. Q.-H. Meng, "Bleeding Frame and Region Detection in the Wireless Capsule Endoscopy Video," *IEEE J. Biomed. Heal. Informatics*, vol. 20, no. 2, pp. 624–630, Mar. 2016.

[20] T. Ghosh, S. A. Fattah, and K. A. Wahid, "Automatic Computer Aided Bleeding Detection Scheme for Wireless Capsule Endoscopy (WCE) Video Based on Higher and Lower Order Statistical Features in a Composite Color," *J. Med. Biol. Eng.*, 2018.

[21] "Capusle Endoscopy." [Online]. Available: https://www.slideshare.net/patrick89/capsule-endoscopy. [Accessed: 03-Jun-2018].

[22] "The capsule endoscopy website." [Online]. Available: http://www.capsuleendoscopy.org.

[23] "What Is the HSV Color Model?" [Online]. Available: https://www.lifewire.com/what-is-hsv-in-design-1078068.

[24] "Introduction to color theory." [Online]. Available: http://infohost.nmt.edu/tcc/help/pubs/colortheory/web/hsv.html.

[25] P. Ganesan and V. Rajini, "Assessment of satellite image segmentation in RGB and HSV color space using image quality measures," in *2014 International Conference on Advances in Electrical Engineering, ICAEE 2014*, 2014.

[26] K. Hechenbichler and K. Schliep, "Weighted k-nearest-neighbor techniques and ordinal classification," 2004.

[27] B. Li and M. Q.-H. Meng, "Texture analysis for ulcer detection in capsule endoscopy images," *Image Vis. Comput.*, vol. 27, no. 9, pp. 1336–1342, 2009.

[28] "Simple guide to confusion matrix terminology." [Online]. Available: http://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/. [Accessed: 10-Jun-2018].

[29] Y. Fu, W. Zhang, M. Mandal, and M. Q.-H. Meng, "Computer-Aided Bleeding Detection in WCE Video," *IEEE J. Biomed. Heal. Informatics*, vol. 18, no. 2, pp. 636–642, Mar. 2014.

[30] Y. Yuan, B. Li, and Q. Meng, "Bleeding Frame and Region Detection in the Wireless Capsule Endoscopy Video," *IEEE J. Biomed. Heal. Informatics*, 2015.

[31] Xiao Jia and M. Q.-H. Meng, "A deep convolutional neural network for bleeding detection in Wireless Capsule Endoscopy images.," *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, 2016.

[32] X. Jia and M. Q. H. Meng, "Gastrointestinal bleeding detection in wireless capsule endoscopy images using handcrafted and CNN features," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2017.

[33] A. K. Kundu, S. A. Fattah, and M. N. Rizve, "An Automatic Bleeding Frame and Region Detection Scheme for Wireless Capsule Endoscopy Videos Based on Interplane Intensity Variation Profile in Normalized RGB Color Space," *J. Healthc. Eng.*, vol. 2018, 2018.

Easy Scheme for Ulcer Detection in Wireless Capsule Endoscopy Images

Md. Sohag Hossain
Electrical and Electronic Engineering
Pabna University of Science and Technology
msh22eee@gmail.com

Abdullah Al Mamun
Electrical and Electronic Engineering
Pabna University of Science and Technology
mamun130203@gmail.com

Md. Motaher Hossain
Electrical and Electronic Engineering
Pabna University of Science & Technology
motaher_eee@pust.ac.bd

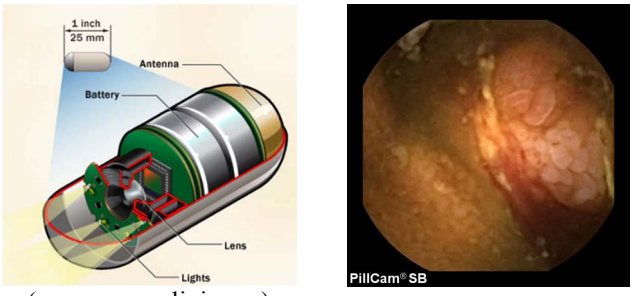
Md. Galib Hasan
Electrical and Electronic Engineering
Pabna University of Science & Technology
eeemgh@gmail.com

***Abstract*—** Among the syndrome of many dangerous diseases in the human digestive tract ulcer is one of the most common. Diagnosis of small intestine ulcers cannot be detected by using other procedures of endoscopy. For visualizing these ulcers, wireless capsule endoscopy (WCE) is excessively being used. In WCE system a huge amount of images are achieved which is very much difficult and time erosive for physicians to check manually. So, for helping the physicians it is very much needed to provide an automated computer-aided system. In this paper, an automated computer-aided diagnosis system is proposed for detecting ulcer from WCE images achieved from the capsule endoscopy video footage. In the first stage, we convert the source image into LAB color spaces for taking features. Then quadratic support vector machine (Q-SVM) classifier has been applied for separating the ulcer and non-ulcer images. Our achieved experiment results 90.39% accuracy and 99.00% sensitivity, validating the effectiveness of the proposed method. Therefore, results of the advanced methodology are given and statistical comparisons are also given related to the other works.

Keywords— Digestive tract, LAB, QSVM, Video footage, Wireless capsule endoscopy

I. INTRODUCTION

The Ulcer is a usual wound of the gastrointestinal (GI) tract. Throughout the world, about 10% of people are affected by this [1]. It is a lesion of the chronic inflammatory type which took place at internal mucous membranes. Ulcer especially digestive tract ulcer is caused by Helicobacter pylori bacteria and non-steroidal anti-inflammatory drugs (NSAIDs). The Ulcer is not fatal but it is the sign of some serious diseases especially Crohn's disease, ulcerative colitis, etc. [2]. In the early days, for ulcer detection in the GI tract, the conventional endoscopic system was used which was very much pain and discomfort because of its flexible tube, light, and camera, etc. After the invention of the transistors, it is possible to design swallowable electronic capsules for imaging the gastrointestinal tract. In 2001 Iddan first initiated video capsule endoscope called a mouth to anus (M2A) for the inspection of the small bowel. Due to its noninvasive property has numerous advantages over conventional endoscopy. Also, it is a painless technique and ability to do his daily activities [3]. A capsule endoscopy device consists of an optic dome, lens, LEDs, a CMOS camera, ASIC circuit, an antenna and a battery for power supply. After swallowing the capsule, it passes through the GI tract and captures images and transmit to a small portable recorder for storage via antenna. The images frame rate is 2 per second, approximately 50000 images were obtained within 8 hours of recording time of WCE system [4]. Signal strength is used to calculate the position of the capsule in the body. So, using this technique physician are able to identify the accurate position of ulcer in the GI tract [3].



(www.mayoclinic.org)

Fig. 1. Capsule endoscopy and sample WCE image

WCE images contain special properties of different diseases like Crohn's disease, ulcerative colitis, etc. Normally ulcer portion contains high-intensity value compared with its nearby area which seems as pinkish or white pinkish color that can be differentiated easily from non-ulcer portion [5].

A. Related Work

Many researchers devoted their efforts to make ulcer detection system fully automated. For this regard Francis *et al.* provide software for identification bleeding region in capsule endoscopy images which sensitivity and specificity is 21.50% and 41.80% respectively [6]. Then MPEG-7 visual descriptor was designed by Coimbra *et al.* for detecting the diseases such as bleeding, ulcer, and polyp [7]. Li *et al.* [8] provide a technique converting discrete curvelet transform (DCT) and local binary pattern (LBP) for feature extraction, which separates the ulcer and non-ulcer images. Karagyris *et al.* [9] proposed a technique in the HSV color space with Log Gabor filter to segment the Hue component of the WCE images to obtain the ulcer portion. Then for classifying the ulcer images from the non-ulcer, the RGB values and texture information are extracted. Yu *et al.* [10] proposed a method for ulcer detection using bag of word with spatial pyramid kernel and feature fusion techniques. Eid *et al.* [11]

introduced a method called DCT-LAC where curvelet-based lacunarity textures are used for feature extraction by calculating the lacunarity index of the WCE images. Ghosh *et al.* [5] proposed a technique for ulcer detection using histogram bin frequency of RGB color space. Another scheme proposed by Kundu *et al.* [12] using Y plane histogram of WCE images for ulcer diagnosis. In our proposed scheme features has been taken from L*a*b* color domain and classification is done by quadratic SVM classifier for separating ulcer and non-ulcer images.

II. OUR PROPOSED METHOD

The proposed scheme for ulcer identification is easy and proves low computational cost. Where some task is applied as pre-processing which is the preparation for implementing the scheme. Then the CE images are converted to L*a*b* color spaces from where features are taken for classification. In this technique, quadratic support vector machine (Q-SVM) classifier is applied for classification.

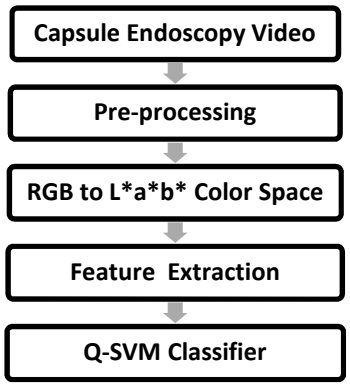


Fig. 2. Flow diagram of the proposed scheme

A. Pre-processing

The first task in our proposed scheme is the conversion of video into images. Basically, on these images, this proposed scheme is applied. The acquiring images converted from the video of capsule endoscopy will be in according to the frame rate of the capsule specification. We have used video of PillCam®SB whose frame rate is two per second with eight-hour battery life. Hence, approximately 57,600 images are achieved from the whole process. Among these huge amounts of images, we have used 110 ulcer and 135 non-ulcer images for this proposed scheme which are available in [13]. On these acquiring images, some algorithm is applied for removing the peripheral black portion of each image frame.

B. Color Space Transform

The images that we take as our database is images of RGB color image. An RGB (Red, Green, Blue) image is an image containing three-pixel values of red, green, and blue color.In RGB color space an image can be represented by the following function,

$$I_{rgb} = \left(P_R, P_G, P_B\right) \tag{1}$$

Where, $P_R(x,y)$represents pixel values (x, y) of red channel, $P_G(x,y)$represents pixel values of green channel and $P_B(x,y)$represents pixel values of blue channel [14]. The RGB color images can be transformed into L*a*b* using the following equations [15].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [M] \begin{bmatrix} r \\ g \\ b \end{bmatrix}$$

Here, [M] The transformation matrix.

$$x_r = \frac{X}{X_r} \qquad y_r = \frac{Y}{Y_r} \qquad z_r = \frac{Z}{Z_r}$$

$$f_x = \begin{cases} \sqrt[3]{x_r}, \text{if } x_r \rangle \in \\ \dfrac{kx_r + 16}{116}, \text{ otherwise} \end{cases}$$

$$f_y = \begin{cases} \sqrt[3]{y_r}, \text{if } y_r \rangle \in \\ \dfrac{ky_r + 16}{116}, \text{ otherwise} \end{cases}$$

$$f_z = \begin{cases} \sqrt[3]{z_r}, \text{ if } z_r \rangle \in \\ \dfrac{kz_r + 16}{116}, \text{ otherwise} \end{cases}$$

Where,

$$\in = \begin{cases} 0.008856, \text{ Actual CIE standard} \\ \dfrac{216}{24389}, \text{ Intent of the CIE standard} \end{cases}$$

$$k = \begin{cases} 903.3, \text{ Actual CIE standard} \\ \dfrac{24389}{27}, \text{ Intent of the CIE standard} \end{cases}$$

$$L^* = 116 f_y - 16 \tag{2}$$

$$a^* = 500\left(f_x - f_y\right) \tag{3}$$

$$c^* = 200\left(f_y - f_z\right) \tag{4}$$

In RGB images there were three channels R, G, and B. But after converting into L*a*b* color spaces we get the values of three channels L*, a*, and b*.

C. Feature Extraction

RGB To extract features from the images of L*a*b* color space, mean and standard deviation has been calculated for each color channel. After calculating, we get mean of L* (M_{L*}), mean of a* (M_{a*}), mean of b* (M_{b*}), standard deviation of L* (Sd_{L*}), standard deviation of a* (Sd_{a*}), and standard deviation of b* (Sd_{b*}). Therefore, for classification using Q-SVM classifier, these mean and standard deviation values are used as features.

D. Q-SVM Classifier

Support Vector Machine (SVM) is a popular machine learning device that is considered as a nonparametric scheme because of using kernel function. SVM classification technique was developed by Vapnik [16]. For classification in our proposed scheme, the quadratic SVM technique has been used. In our observation we have n training data a_j, $a_j \in R^N$ for j=1,….n each either an ulcer or non-ulcer. Then as separating rule [17], [18],

$$\omega^T x \geq \gamma$$

Where, $x \in \mathbb{R}^N$ is a variable with $\omega \in \mathbb{R}^N$, $\gamma \in \mathbb{R}$ are the classifier parameters.

In our proposed scheme we have used kernel function denoted by $K(x_{ul}, x_{nul})$ where x_{ul} means ulcer and x_{nul} means non-ulcer data.

$$K(x_{ul}, x_{nul}) = (x_{ul}.x_{nul})^d$$

Where, d is the degree of the polynomial.

Now, to get $\omega and\ \gamma$ from the following equation,

$$\min_{\omega\in\mathbb{R}^N,\gamma\in\mathbb{R},y\in\mathbb{R}^n_+} \frac{1}{2} y^T D\left[K(x_{ul}, x_{nul})K(x_{ul}, x_{nul})\right]Dy - Ce^T y$$

Where,

$$\omega = y^T D[K(x_{ul}, x_{nul})K(x_{ul}, x_{nul})]Dy$$

$$\gamma = -e^T Dy$$

The classification has been implemented having five-fold cross-validation technique to split the samples into training and testing sets. After applying quadratic support vector machine classifier we are able to separate the ulcer and non-ulcer images.

III. Result and Discussion

A. Dataset Description

The WCE videos are collected from a publicly available database [13]. The initial process of this research work is converting the acquiring video footage from wireless capsule endoscopy to image frame according to the frame rate specified by capsule specification for further processing such as color space transform, feature extraction, and classification. Here used capsule endoscopy of PillCam®SB whose frame rate per second is 2 and battery life 8 hours that achieved a huge amount of data for analyzing. Among this we have used 245 color images where 110 ulcer and 135 non-ulcer images from 8 capsule endoscopy videos. Image frames in ulcer videos are manually vaccinated as ulcer or non-ulcer and if ulcer, the ulcer zones are marked by expert physicians.

B. Performance Measures of Ulcer Detection

The performance of our proposed method was determined using five parameter accuracy, sensitivity, specificity, precision, and f1-score. During the classification of ulcer and non-ulcer images, four types of occurrences may occur. The occurrences are true ulcer detection (T_{ulcer}), false ulcer detection (F_{ulcer}), true non-ulcer detection ($T_{non\text{-}ulcer}$), false non-ulcer detection ($F_{non\text{-}ulcer}$). Accuracy measures how to agree with the calculated value with true value. Hence, it measures how to correct a classification test due to a given condition. The numerical value of accuracy represents the ratio of true ulcer and non-ulcer detection with total observation. Sensitivity measures the correctness of ulcer detection. The numerical value of sensitivity represents the ratio of true ulcer detection with the true positive condition. The higher sensitivity shows more accurate ulcer detection. Specificity measures the correctness of non-ulcer detection. The higher specificity shows more accurate non-ulcer detection. Precision measures how to agree the calculate value among other calculated values. F1 score measures the correctness of the classification test. It is the harmonic mean of precision and sensitivity [19], [20].

$$\text{Accuracy} = \frac{\sum T_{ulcer} + \sum T_{non-ulcer}}{\sum T_{ulcer} + \sum T_{non-ulcer} + \sum F_{ulcer} + \sum F_{non-ulcer}} \quad (5)$$

$$\text{Sensitivity} = \frac{\sum T_{ulcer}}{\sum T_{ulcer} + \sum F_{non-ulcer}} \quad (6)$$

$$\text{Specificity} = \frac{\sum T_{non-ulcer}}{\sum T_{non-ulcer} + \sum F_{ulcer}} \quad (7)$$

$$\text{Precision} = \frac{\sum T_{ulcer}}{\sum T_{ulcer} + \sum F_{ulcer}} \quad (8)$$

$$\text{F1 Score} = \frac{2\times Sensitivity\times Precision}{Sensitivity + Precision} \quad (9)$$

Using equation (5), (6), (7), (8), and (9). we can calculate accuracy, sensitivity, specificity, precision, and f1-score respectively. First, after taking features from the RGB images we get accuracy 89.51%, sensitivity 87.44%, specificity 90.08%, precision 88.32%, and f1-score 87.49%. Then we get features from the images after converting into HSV, YCbCr, and L*a*b* color spaces for choosing better color space. The comparison has been shown in TABLE I. Though specificity and precision are higher in RGB color space. We affix on L*a*b* color spaces due to its higher accuracy, sensitivity and f1-score especially - accuracy 90.39%, sensitivity 99.00%, specificity 86.63%, precision 80.16%, and f1-score 88.50%.

TABLE I
Performance comparison with color spaces

Color Space	Acc. (%)	Sen. (%)	Spe. (%)	Pre. (%)	F1-Score
RGB	89.51	87.44	90.08	88.32	87.49
HSV	88.02	88.00	90.00	88.00	88.00
YCbCr	87.70	96.00	86.00	80.12	87.27
L*a*b*	**90.39**	**99.00**	**86.63**	**80.16**	**88.50**

Analysis has done for selecting the efficient classifier among Quadratic Discriminant Analysis (QDA), Quadratic Support Vector Machine (Q-SVM), Cubic Support Vector Machine (C-SVM), and Fine KNN classifier in RGB, HSV, L*a*b*, and YCbCr color space. Among these Quadratic Support Vector Machine (Q-SVM) in L*a*b* color space provide higher accuracy, sensitivity, and f1-score. For ulcer detection sensitivity is an important parameter for testing performance and also f1 score measures the correctness of classification test. So, for these reasons, we have to choose Q-SVM in L*a*b* color space.

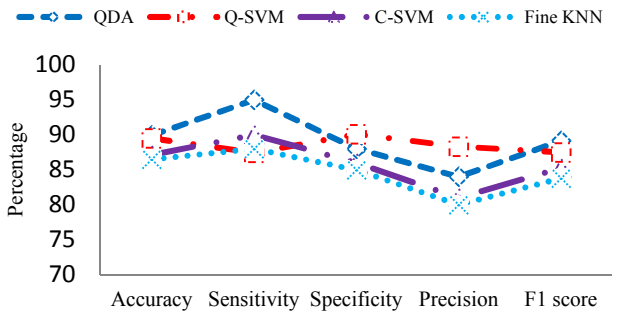


Fig. 3. Performance comparison varying classifiers in RGB color space

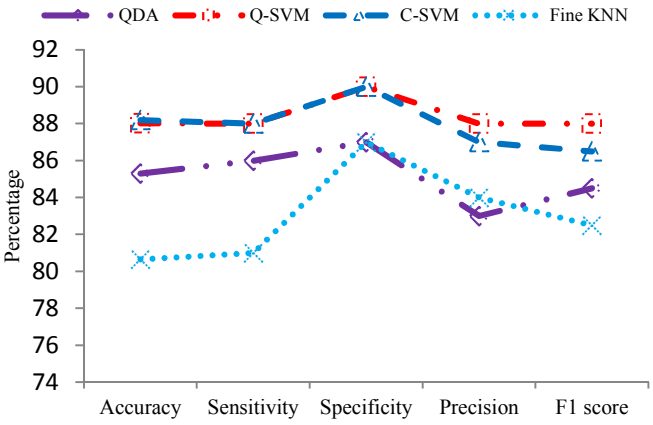


Fig. 4. Performance comparison varying classifiers in HSV color space

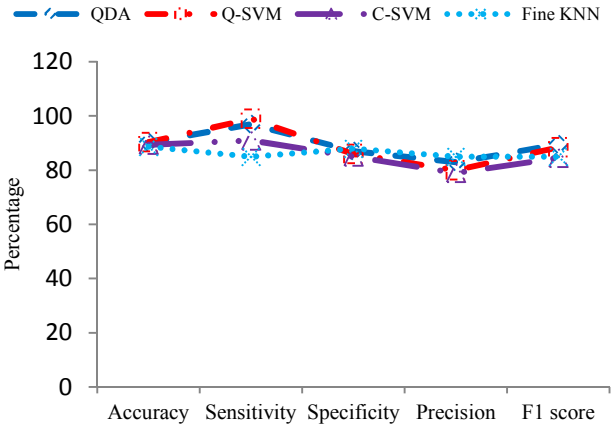


Fig. 5. Performance comparison varying classifiers in L*a*b* color space

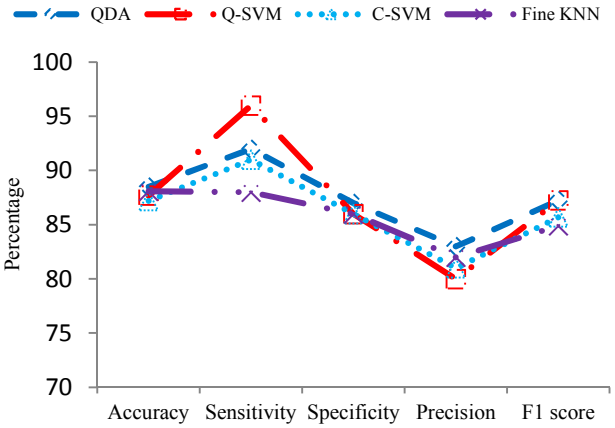


Fig. 5. Performance comparison varying classifiers in YCbCr color space

For feature selection we have chosen mean, median, min, max, moment, standard deviation, skewness, kurtosis etc. But mean and standard deviation of the values of L*, a*, and b* provide the best performance. Hence, we have used six types of features - mean of L*, a*, and b* with the standard deviation of L*, a*, and b*. By varying these all, we chose the combination of all the features which provides the best performance - accuracy 90.39%, sensitivity 99.00%, specificity 86.63%, precision 80.16%, and f1-score 88.50%.

Table II
Performance comparison varying features

FEATURES	ACC. (%)	SEN. (%)	SPE. (%)	PRE. (%)	F1 SCORE (%)
M_{a*}, M_{b*}, Sd_{L*} and Sd_{b*}	89.30	96.00	86.00	80.00	87.27
M_{b*}, Sd_{L*}, Sd_{a*} and Sd_{b*}	87.30	91.00	85.00	79.00	84.57
$M_{L*}, M_{a*}, M_{b*}, Sd_{a*}$ and Sd_{b*}	84.40	85.00	84.00	79.00	81.89
$M_{L*}, M_{b*}, Sd_{L*}, Sd_{a*}$ and Sd_{b*}	88.10	93.00	85.00	79.00	85.43
$M_{L*}, M_{a*}, Sd_{L*}, Sd_{a*}$ and Sd_{b*}	89.80	96.00	86.00	81.00	87.86
$M_{L*}, M_{a*}, M_{b*}, Sd_{L*}$ and Sd_{a*}	86.10	90.00	83.00	77.00	82.99
M_{L*}, M_{a*} and M_{b*}	77.00	75.00	78.00	72.00	73.45
$M_{L*}, M_{a*}, M_{b*}, Sd_{L*}$ and Sd_{b*}	89.80	95.00	87.00	82.00	88.02
$M_{a*}, M_{b*}, Sd_{L*}, Sd_{a*}$ and Sd_{b*}	88.10	95.00	84.00	77.00	85.06
$M_{L*}, M_{a*}, M_{b*}, Sd_{L*}, Sd_{a*}$ and Sd_{b*}	**90.39**	**99.00**	**86.63**	**80.16**	**88.50**

Mean of L* = M_{L*}; Mean of a* = M_{a*}; Mean of b* = M_{b*}; Standard deviation of L* = Sd_{L*}; Standard deviation of a* = Sd_{a*}, and Standard deviation of b* = Sd_{b*}

At last, the performance result of our proposed scheme has been compared with the performance of other methods of [5], [12], and [8]. Ghosh *et al.* method [5] features have been extracted from histogram bin frequency which was classified by KNN classifier. Kundu *et al.* method [12] were implemented via Y plane histogram with SVM classifier. Another method [8] has been proposed by Li *et al.* where CLBP technique has been used. Finally, our proposed scheme provides height accuracy and sensitivity over these all. TABLE III shows the comparison.

TABLE III
Performance comparison with other's method

	Accuracy (%)	Sensitivity (%)	Specificity (%)
Ghosh *et al.* [5]	87.27	88.64	85.75
Kundu *et al.* [12]	87.49	83.68	91.08
Li *et al.* [8]	89.20	87.53	91.46
Proposed Method	**90.39**	**99.00**	86.63

IV. Conclusion

In this paper, we proposed an easy technique with lower computational cost. This technique is a fully automated computer-aided system to diagnosis ulcer from WCE images. First, the RGB input image transformed into L*a*b* color space. Images at L*a*b* color space contains three components- L*, a* and b*. Then, the mean and standard deviation of each component has taken as features. The classification has done for separating ulcer and non-ulcer images applying Q-SVM classifier. Finally, the obtained performance – accuracy 90.39%, sensitivity 99.00%, specificity 86.63%, precision 80.16%, and f1-score 88.50%. Moreover, the comparison experiments showed that our method outperforms on others on the WCE ulcer classification.

References

[1] V. Charisis, L. Hadjileontiadis, and G. Sergiadis, "Enhanced ulcer recognition from capsule endoscopic images using texture analysis," in *New Advances in the Basic and Clinical gastroenterology*, ed: InTech, 2012.

[2] V. S. Charisis, L. J. Hadjileontiadis, J. Barroso, and G. D. Sergiadis, "Intrinsic higher-order correlation and lacunarity analysis for WCE-based ulcer classification," in *Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on*, 2012, pp. 1-6.

[3] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless capsule endoscopy," *Nature,* vol. 405, p. 417, 2000.

[4] C. Vere, S. Cazacu, C. Streba, F. Sima, D. Parvu, and A. Ionescu, "Capsule endoscopy: Diagnostic Role in Obscure Gastrointestinal Bleeding," *Curr Health Sci J,* vol. 39, pp. 154-158, 2009.

[5] T. Ghosh, A. Das, and R. Sayed, "Automatic small intestinal ulcer detection in capsule endoscopy images," *International Journal of Scientific and Engineering Research,* vol. 7, pp. 737-741, 2016.

[6] R. Francis, "Sensitivity and specificity of the red blood identification (RBIS) in video capsule endoscopy," in *3Rd, Int. Conf. on Capsule Endoscopy*, 2004.

[7] M. T. Coimbra and J. S. Cunha, "MPEG-7 visual descriptors—contributions for automated feature extraction in capsule endoscopy," *IEEE transactions on circuits and systems for video technology,* vol. 16, pp. 628-637, 2006.

[8] B. Li and M. Q.-H. Meng, "Texture analysis for ulcer detection in capsule endoscopy images," *Image and Vision computing,* vol. 27, pp. 1336-1342, 2009.

[9] A. Karargyris and N. Bourbakis, "Identification of ulcers in wireless capsule endoscopy videos," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, 2009, pp. 554-557.

[10] L. Yu, P. C. Yuen, and J. Lai, "Ulcer detection in wireless capsule endoscopy images," in *Pattern Recognition (ICPR), 2012 21st International Conference on*, 2012, pp. 45-48.

[11] A. Eid, V. S. Charisis, L. J. Hadjileontiadis, and G. D. Sergiadis, "A curvelet-based lacunarity approach for ulcer detection from wireless capsule endoscopy images," in *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*, 2013, pp. 273-278.

[12] A. Kundu, A. Bhattacharjee, S. Fattah, and C. Shahnaz, "An automatic ulcer detection scheme using histogram in YIQ domain from wireless capsule endoscopy images," in *Region 10 Conference, TENCON 2017-2017 IEEE*, 2017, pp. 1300-1303.

[13] (2016). *The capsule endoscopy website*. Available: https://www.capsuleendoscopy.org

[14] T. Kumar and K. Verma, "A Theory Based on Conversion of RGB image to Gray image," *International Journal of Computer Applications,* vol. 7, pp. 7-10, 2010.

[15] K. Leon, D. Mery, F. Pedreschi, and J. Leon, "Color measurement in L∗ a∗ b∗ units from RGB digital images," *Food research international,* vol. 39, pp. 1084-1091, 2006.

[16] V. Vapnik, "The nature of statistical learning theory Springer New York Google Scholar," 1995.

[17] P. Białoń, "Solving Support Vector Machine with Many Examples," *Journal of Telecommunications and Information Technology,* pp. 65-70, 2010.

[18] D. R. Musicant, "Data mining via mathematical programming and machine learning," University of Wisconsin--Madison, 2000.

[19] W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," *NESUG proceedings: health care and life sciences, Baltimore, Maryland,* vol. 19, p. 67, 2010.

[20] Y. Sasaki, "The truth of the F-measure," *Teach Tutor mater,* vol. 1, 2007.

[21] D. G. Altman and J. M. Bland, "Diagnostic tests. 1: Sensitivity and specificity," *BMJ: British Medical Journal,* vol. 308, p. 1552, 1994.

Investigation of EEG Signals for Brain Computer Interface

Etu Podder
Electronics and Communication Engineering Discipline
Khulna University
Khulna-9208, Bangladesh
e2.120926@gmail.com

Md. Maniruzzaman
Electronics and Communication Engineering Discipline
Khulna University
Khulna-9208, Bangladesh
m_m_zaman@hotmail.com

Alok Sarkar
Electronics and Communication Engineering Discipline
Khulna University
Khulna-9208, Bangladesh
alokece99@gmail.com

***Abstract*—Brain computer interface (BCI) is an interconnecting scheme between human brain and computer which entitles humans to interact with their environs by their brain waves instead of muscle activities. In the age of new argumentative technology, EEG signals are extensively employed in the involvement of BCI. The key motive of the work is to investigate EEG activities during imagination of hand movements. BCI competition IV dataset 2a were studied as well as analyzed in this research. We have detected the event related desynchronization and synchronization of mu and beta rythms during hand imagery movements by Power Spectral Analysis (PSA), Event Related Potential (ERP) plots, ERP comparison, Time-Frequency Analysis. Mainly, the EEG rhythms recorded at channels C3 and C4 are investigated. Results of this paper show that imagery movement of hands is usually associated with decrement in mu and beta rhythms, which is typically adversary to the moving activities. Our results are compatible with previous motor imagery researches and are of practical interest for the evolvement of BCI for motor reclamation.**

Keywords—brain computer interface, EEG signals, event related desynchronization, event related synchronization, motor imagery

I. INTRODUCTION

Electroencephalogram (EEG) is the representation of electrical activities of the brain from the scalp. The EEG signal's patterns and parameters imply the health condition of the brain. Therefore, proper investigation of EEG plays a significant role in brain studies. EEG signals contain a wide range of frequency bands. There are mainly five types of brain waves (EEG waves): δ waves (0.4-4 Hz), θ waves (4-8 Hz), α waves (8-13 Hz), β waves (13-30 Hz), and γ waves (30-100 Hz). Besides this, the motor cortical areas of brain mostly reveals 8-12 Hz EEG activities, called Rolandic μ rhythms and are commonly accompanied with 18-26 Hz β rhythms. μ rhythms (8-12 Hz) are influenced by movements or imagination of movements. Particularly, the assessment of EEG waves has been become intensively familiar to identify the difficulties of extracting the knowledge of the brain signals. One of the major applications of EEG waves is in BCI.

BCI is a direct connecting way between brain and exterior machines. BCI associates the brain activity to an external computer, as shown in Fig. 1, which enables an individual to control devices straightly with his brain waves (EEG waves) and without using his muscles. Now-a-days, controlling device with the aid of BCI has become an ultimate substitute for those who are paralyzed or suffering from crucial motor disabilities. Besides this, the EEG based BCI can be utilized in different home appliances, environment control etc. A BCI system translates reactivities of the EEG signals into a command to conduct an exterior machine fulfilling the task of an interaction [1]. This translation involves proper and efficient recognition of the user intension's patterns of the brain by proper investigation and careful analysis of EEG waves. Therefore, investigation and analysis of the recorded EEG signals by appropriate techniques are of great importance to get the powerful BCI.

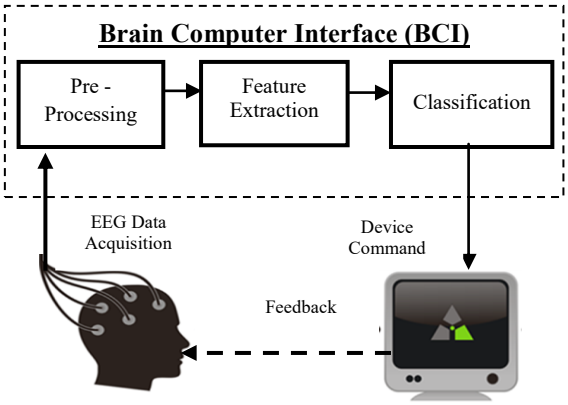


Fig. 1. Simplified block diagram of EEG based BCI.

A variety of studies have been done in order to extract information of the brain waves for robust BCI research. Over the past two decades, imagination of movements has become most popular to design EEG-based BCI systems. This entitles motor disable person to operate several BCI devices like wheelchairs [2], prosthetic devices [3], and computers [4]. Several studies have revealed that motor disable person or normal person can able to govern the amplitude of μ or β rhythms in EEG waves stored from the scalp over sensorimotor area [5]. Varity of factors suggest that the event-related desynchronization (ERD) are the decrease that appear when thinking about the movement or performing the movements. ERD begins while the subject initiates to think of a movement and reveals it as a potential desynchronization in μ and/or β rhythms. In case of motor imagery, ERD occurs in the contralateral hemisphere and event-related synchronization (ERS) occurs in the ipsilateral hemisphere. Analysis of EEG signals for BCI includes preprocessing, feature extraction techniques and classification [6, 7]. Supervised classification methods are employed to realize the mapping in between EEG waves and classes resembled to mental events [8, 9]. It was proposed that the ERD and ERS can be significant classification features for the motor imagery EEG waves [10].

In this research work, we investigate the ERD/ERS of EEG signals (μ and β rhythms) by Power Spectral Analysis (PSA), Event Related Potential (ERP) plots, ERP comparison, Time-Frequency analysis, etc. which will

indicate that EEG activity during imagery movements may be employed in a motor conducted BCI.

II. Materials and Methods

This research work has been carried out in several steps as shown in following Fig. 2.

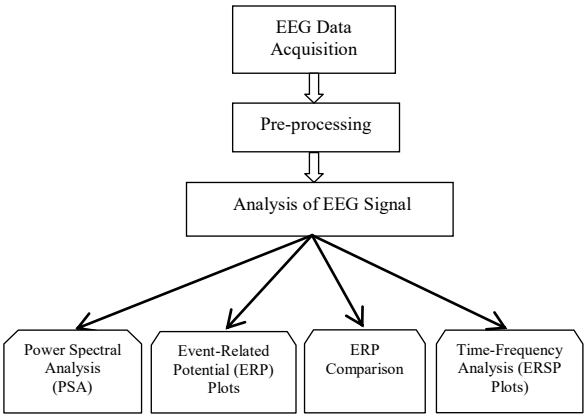


Fig. 2. Working steps.

A. Data Acquisition and Description:

In our research, the motor imagery EEG data have been taken from BCI competition IV dataset 2a provided by Graz for investigating EEG activity during imagination of movements. In this BCI dataset, EEG data are provided for 9 subjects. The BCI illustration was cue based and the data are taken for four different motor imagery tasks, such as left hand imagery movement (class 1), right hand imagery movements (class 2), both feet movements imagination (class 3), and imagery tongue movements (class 4) [11]. 22 Ag/AgCl electrodes were used to record the EEG in international 10-20 system, as shown in Fig. 3.

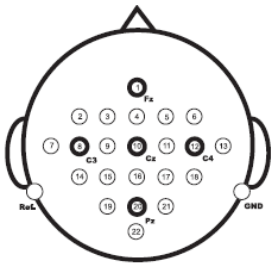


Fig. 3. Channel locations of the recorded EEG signal [11].

The signals were sampled with 250 Hz. In this database, EEG data are provided in the General Data Format (GDF) for each subject as well as session. Here, class labels are stored for only one session for all trials, but another session is imported to evaluate performances. We have used only the training datasets for 9 subjects. The events for four classes of data are shown in TABLE I.

TABLE I. DESCRIPTION OF DATA

Event	Description
769	Left hand imagery (class 1)
770	Right hand imagery (class 2)
771	Foot imagery (class 3)
772	Tongue imagery (class 4)

We have investigated the data for events of 769 and 770 corresponding to right and left hand imagery movements.

B. Pre-processing:

We have used EEGLAB (an open source tool box) provided by SCCN lab, running under the cross platform MATLAB for pre-processing of EEG data. The pre-processing steps are shown in Fig. 4.

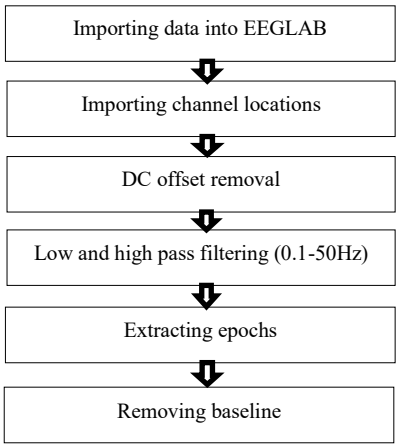


Fig. 4. Steps of pre-processing.

Firstly, the GDF data files have been imported into EEGLAB with the BioSig toolbox using EEGLAB function and plugin and then the channel locations are edited according to the electrode positions which was used to record these data as described before. After editing the channel locations, the channel data have been plotted for visualization and the DC offsets have been removed. DC offsets initiates filter artifacts at signal borders. So, it is better to remove them prior filtering. To dispel linear trends, the data have been then filtered. The data is then high-pass filtered with lower cut-off frequency of 0.5 Hz and after that low-pass filtered with cut-off frequency of 50 Hz by using Linear FIR filter [12]. This filtering of EEG data eliminated the noise and artifacts. Then we have extracted epoch corresponding to the event of right hand imagery (770) and left hand imagery (769) followed by baseline removal. We have extracted epochs from 1s before the event and 2s after the onset of event, and then EEG signals are analyzed. Thus, we have obtained the epochs of two events-

- Imagination of right hand movement (event 770)
- Imagination of left hand movement (event 769)

C. Analysis of EEG Signal:

After pre-processing, we have investigated the extracted epochs by different analysis:

1) Power Spectrum Analysis (PSA): Power spectral analysis is a well-known scheme for the analysis of EEG signals. Power Spectral plot is a colored trace which indicates signal power spectra of the activity of the channels along the range of the frequencies with colors. In our research, we have used EEGLAB to investigate the power spectrum of the selected epoch at chosen frequency. Thus, we have investigated the active part of the brain during right arm and left arm imagery. Here, the scalp distribution of power was observed by plotting the power spectral of epoched data at 4, 8, 10, 12, 18, 20, 22, 40 Hz as we have

investigated the ERD/ERs of μ (8-12 Hz) and β (18-26 Hz) rhythms, we have taken the value of frequencies of these types. Here, the channel spectra and maps have been plotted for electrode positions C3 (left side motor area) and C4 (right side motor area).

2) Event-Related Potential (ERP) Plots: An event-related potential (ERP) is the observed brain stimulus which is the direct outcome of a particular motor work. The electrophysiological data analysis mostly uses dimensional event-related potential (ERP) averages. EEGLAB provides various ERP image for realizing the motive of ERP effects. ERP-image plots are 2-D view (trial-by-trial) of the event-related data epochs in which data epochs are first sorted along some relevant dimension then smoothed and finally color-coded and imaged. In this research work, we have investigated the epoched data for two different events by plotting ERP images. ERP averages have been also utilized for the comparison of the event related potentials.

a) Channel ERP Image Plots: ERP image plot facilitates trial by trail views of a set of data epochs. An ERP image is a colored rectangular image where every single horizontal line is used to represent a potential time series for every experimental test. In our thesis, we have studied ERP image plot for channel C3 and C4 during right hand and left hand imagery where we have used the smoothing 10, down-sampling 1 and time limits 1000ms to 1996ms.

b) ERP Comparison: To make our analysis more significant, we have compared the ERPs in two different events. ERP averages have been used to compare the ERPs. We have compared the ERP of channel C3 and C4 to differentiate ERD and ERS during imagination of right hand and left hand movements.

3) Time-Frequency Analysis (ERSP Plots): When the information about changes in frequency is unknown, time-frequency analysis plays a crucial role for analyzing EEG data [13]. Event-related spectral perturbation (ERSP) and inter-trial coherence (ITC) are two methods for time or frequency decomposition, used in the basic EEGLAB. In our work, we have used ERSP plot method. ERSP quantifies the mean event variation in power spectrum at a data electrode. The synchronization as well as desynchronization of events which are narrowband was visualized by ERSP plots [14]. An ERSP can be calculated by following equation.

$$ERSP(f,t)=\frac{1}{n}\sum_{k=1}^{n}\left|F_k(f,t)\right|^2 \quad (1)$$

Here, n stands for number of trials, $F_k(f,t)$ stands for spectral estimation for k^{th} trial, f stands for frequency and t stands for time. In this research, we have used ERSP method to quantify the mean event variation in power spectrum at data electrode positions C3 and C4 for right hand and left hand imagination movements.

III. Results and Discussion

A. Power Spectral Plots:

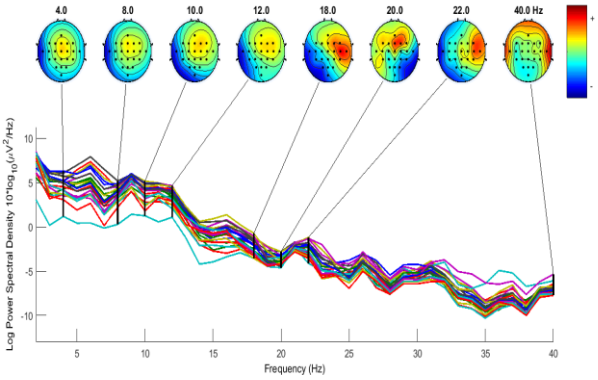


Fig. 5. Power spectral plots during right hand imagery movements at 4, 8, 10, 12, 18, 20, 22, 40 Hz.

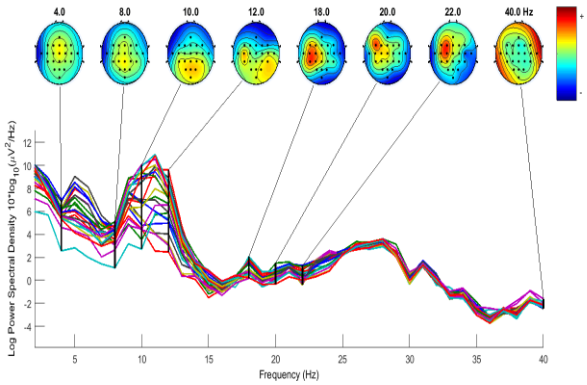


Fig. 6. Power spectral plots during left hand imagery movements at 4, 8, 10, 12, 18, 20, 22, 40 Hz.

Figs. 5 and 6 show the results of power spectral plots of EEG data during right hand imagery and left hand imagery at 4, 8, 10, 12, 18, 20, 22, 40 Hz. In power spectral plot of right hand imagery (Fig. 5), a decrease in the spectral can be shown near C3 electrode positions position (left side sensory motor area) at 8-12 Hz, 18 Hz and 22 Hz. It means a reduction in power spectrum can be seen in the left side of the brain which is represented by the blue color. Therefore, power spectrum in right side brain is relatively high and it is indicated by red color. In power spectrum during left hand imagery movements (Fig. 6), a decrease in the spectral (blue color) can be shown near C4 at around 12 Hz and 18-22 Hz which means lower power spectrum in the right hemisphere of the brain. Eventually, high power spectrum (red color) can be seen in left side motor area of the brain.

In frequency domain, ERD is identified by a decrement in the spectral power. Eventually, ERS means a high power spectrum. Therefore, the results show ERD in left side brain and ERS in right side brain during right hand imagery. Oppositely, during left hand imagery, an ERD is shown in the right side motor area and an ERS in left side motor area of the brain.

B. *Channel ERP Image Plots:*

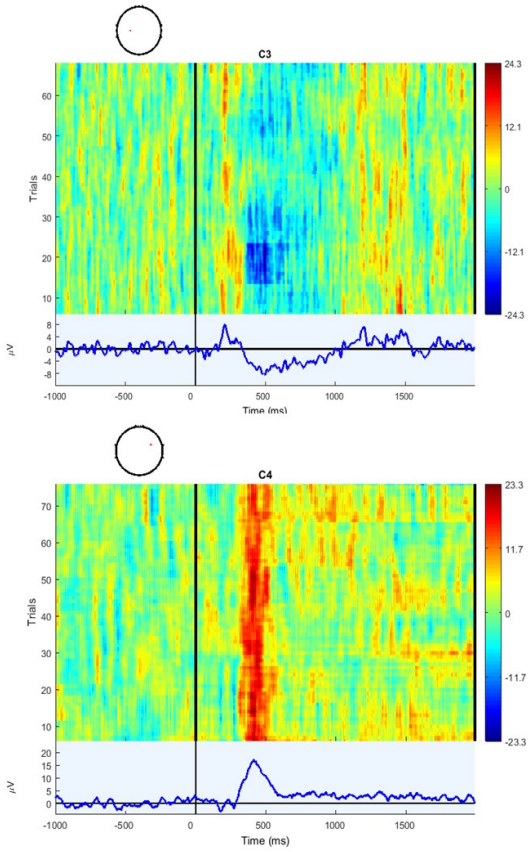


Fig. 7. Channel ERP image at electrodes C3, C4 during imagination of right hand movements.

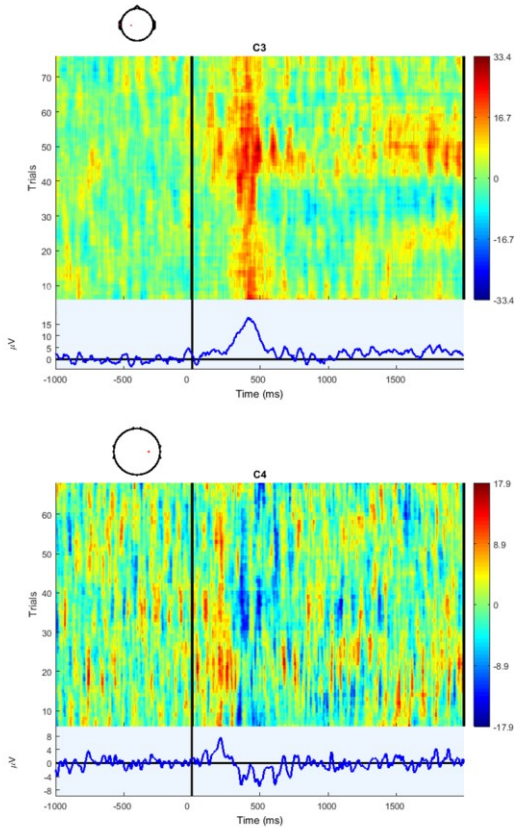


Fig. 8. Channel ERP image at electrodes C3, C4 during imagination of left hand movements.

Figs. 7 and 8 show the result of channel ERP image plots. In our research, the channel ERP images are plotted for right hand and left hand imagination at electrode positions C3 and C4. In case of right hand imagery movements (Fig. 7), the channel ERP image at electrode positions C3 (left side sensory motor area of the brain) shows a decrement in ERP at around 400ms after starting of event which indicates the ERD of μ and β rhythm. In color plot, this desynchronization is represented by blue color. Eventually, an increase in ERP at near 400ms is observed in right side brain (channel C4). Similarly, when we consider about left hand imagery movements (Fig. 8), channel ERP image of electrode positions C3 shows an ERS (red color) around 400ms after starting the event. On the other side, the channel ERP of C4 shows a decrease in ERP at 400ms. The result is summarized in the TABLE II.

TABLE II. ERD/ERS DURING IMAGERY MOVEMENTS

S	Motor Imagery	ERD	ERS
1	Right hand movement	Left side motor area of the brain	Right side motor area of the brain
2	Left hand movement	Right side motor area of the brain	Left side motor area of the brain

C. *Comparison of channel ERPs:*

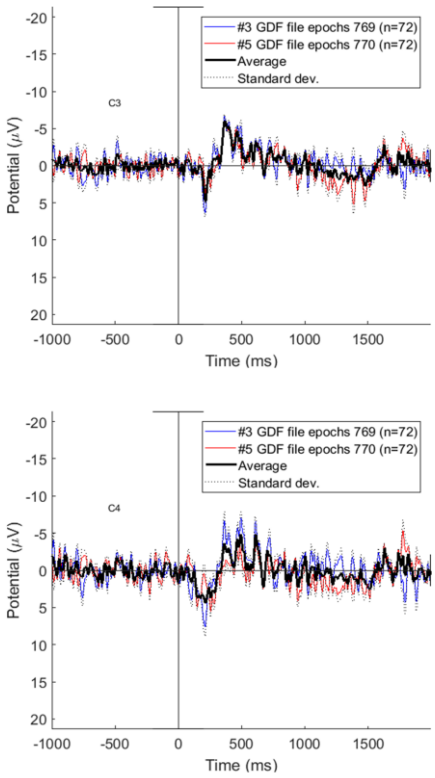


Fig. 9. Comparison of channel ERPs of right and left imagery at electrodes C3, C4.

Comparisons of channel ERPs are also plotted for the electrode positions C3 and C4 respectively which is shown in Fig. 9. The channel ERPs are also compared so that we can strengthen the result obtained from channel ERP images for right hand and left hand imagery at channel C3 and C4.

D. Time-Frequency Analysis (ERSP Plots):

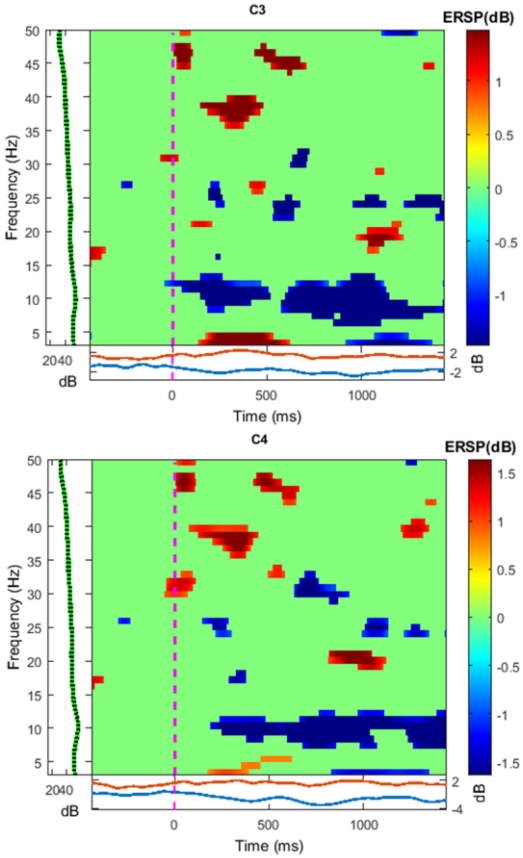


Fig. 10. ERSP plot during right hand imagery movements at channel C3 and C4.

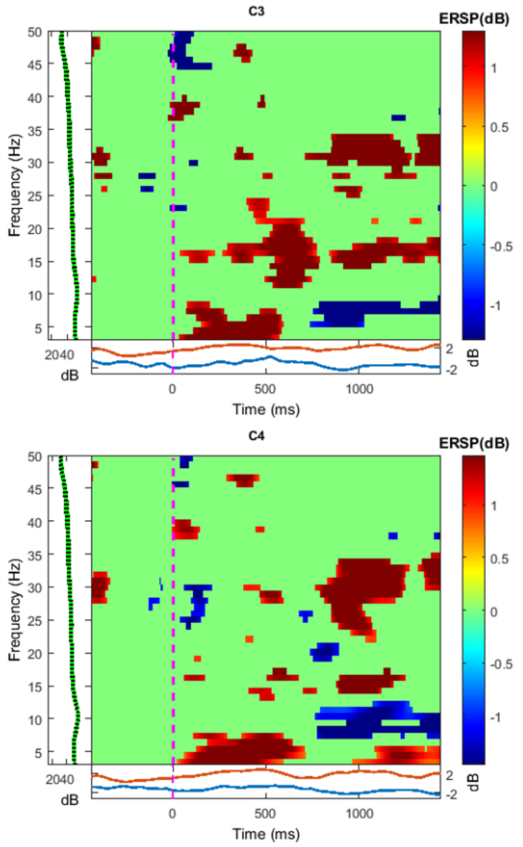


Fig. 11. ERSP plot during left hand imagery movements at channel C3 and C4.

Figs. 10 and 11 represent the ERSP plots at channel C3 and C4 during right and left hand imagery movements respectively. In Fig. 10, the ERSP plot of C3 electrode (left side brain electrode) during right hand imagery shows a decrease in the spectral power (represented by blue color) around 400-500ms after the onset of event related potential around μ rhythm (10-12 Hz). Again, the ERSP plot of channel C4 shows an enhancement around 400-500ms at 10-12 Hz. Fig. 11 shows the ERSP plot of C3 and C4 electrodes during left hand imagery. At channel C4, a decrease in the spectral power can be observed around 400-500ms after the onset of event related potential around 10-12 Hz. And again, at C3, an increment can be seen around 500ms at 10-12 Hz. The result of ERSP plots demonstrates the desynchronization of mu rhythms in the respective hand areas for hand motor imagery.

All of our analyses have given the consistent results for both hand imagery movements which are very much significant for our research. Also, our results are compatible with previous motor imagery researches and are good enough compared to some other previous motor imagery results. In 2008, the ERD and ERS of μ and β rhythms were analyzed by Yeom and Sim during imagery movements of arms [16]. But in that research, only PSA and ERSP plots are used which limits practicability. In our research, more analyzing methods have been used. In 2009, Varghese analyzed motor imagery data during left hand and right hand movements and got significant results [17]. Our hand imagination results are coincided with the research and our classification results are more indicative. In 2016, hand movement motor imagery EEG was studied by Shah and Nayak, where 'BCI competition II' dataset were utilized [18]. In our research, 'BCI competition IV dataset 2a' are used which have shown better results. Moreover, our motor imagery hand movements results are coincident with those reported by Farland et al. [19], Pfurtscheller et al. [3], and Sivakami and Devi [20].

IV. Conclusion

In this work, we analyzed the motor imagery data of 'BCI competition IV dataset 2a' [11] provided by Graz during visualization of right arm and left arm movements using EEGLAB, an open source toolbox running under MATLAB. We have investigated that how EEG activity changes with imagery movements by differentiating ERD and ERS of μ and β rythms. The result of this research substantiated that imagination of hand movements is generally associated with ERD in μ and β rhythms, typically adversary to the movements. The major significance of our research is that by our results, we are able to differentiate the features of μ and β rhythms occurring in sensory motor cortex during hand imagery by ERD/ERS quantifications. Also, our analyses give compatible results which are very much important for BCI implementations.

In future, we will try to convert these significant findings into BCI commands and thus conduct a device, such as, wheelchair, only by motor imagery EEG waves for disabled people.

REFERENCES

[1] J. R. Wolpaw, D. J. McFarland, and T. M. Vaughan, "Brain-computer interface research at the Wadsworth Center," IEEE Transactions Rehabilitation Engineering, vol. 8, pp. 222–225, 2000.

[2] G. R. Muller-Putz and G. Pfurtscheller, "Control of an electrical prosthesis with an SSVEP-based BCI," IEEE Transactions Biomedical Engineering, vol. 55, pp. 361-364, 2008.

[3] G. Pfurtscheller, C. Neuper, and C. Guger, "Current trends in Graz brain-computer interface research," IEEE Transactions Rehabilitation Engineering, vol. 8, pp. 216-219, 2000.

[4] J. R. Wolpaw, N. Birbaumer, G. Pfurtscheller, and D. J. McFarland, "Brain-computer interfaces for communication and control," Journal of Clinical Neurophysiology, vol. 113, pp. 529–538, 2002.

[5] L. Qin, L. Ding, and B. He, "Motor imagery classification by means of source analysis for brain computer interface applications," Journal of Neural Engineering, vol. 1, no. 3, pp. 135–141, 2004.

[6] D. P. Subha, P. K. Joseph, R. Acharya, and C. M. Lim, "EEG signal analysis: A survey," Springer Science, vol. 34, pp. 195–212, 2007.

[7] M. R. Lakshmi, T. V. Prasad, and V. C. Prakash, "Survey on EEG signal processing methods," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 4, no. 1, 2014.

[8] F. Lotte, M. Congedo, and A. Lecuyer, "A review of classification algorithms for EEG based brain-computer interfaces," Journal of Neural Engineering, vol. 4, pp. R1-R13, 2007.

[9] B. Kamousi, Z. Liu, and He B, "Classification of motor imagery tasks for braincomputer interface applications by means of two equivalent dipoles analysis," IEEE Transactions Neural Systems and Rehabilitation Engineering, vol. 13, no. 2, pp. 166-171, 2005.

[10] Y. Fang, L. Hao, and J. Xiaojie, "Improved classification methods for brain computer interface system," Computer Network and Information Security, vol. 2, pp. 15-21, 2012.

[11] C. Brunne , R. Leeb, G. R. Muller-Putz, A. Schlogl, and G. Pfurtscheller, "BCI competition 2008 – Graz data set A," 2008.

[12] A. Delorme, T. Fernsler, H. Serby, S. Makeig, "EEGLAB Tutorial," University of San Diego California, 2006.

[13] A. Madhu, V. K. Jayasree, and V. Thomas, "A survey paper on time frequency analysis of EEG signal," Progress in Science and Engineering Research Journal, vol. 02, pp. 46-51, 2013.

[14] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," Journal of Neuroscience Methods, pp. 9-21, 2004.

[15] E. Niedermeyer and FL. da Silva, Electroencephalography: basic principles, clinical applications and related fields, 5th ed. Lippincott, New York: Williams & Wilkins, 2005.

[16] H. G. Yeom, K. B. Sim, "ERS and ERD analysis during the imaginary movement of arms," International Conference on Control, Automation and Systems, Oct. 14-17, 2008.

[17] J. Parokaran, "Analysis of EEG signals for EEG-based brain-computer interface," 2009.

[18] R. G. Shah, and R. Nayak, "Hand movement classification using motor imagery EEG," International Journal of Innovative Research in Science, Engineering and Technology, vol. 5, pp. 6, 2016.

[19] Mc. Farland, J. Dennis, L. A. Miner, T. M. Vaughan, and J. R. Wolpaw, "Mu and beta rhythm topographies during motor imagery and actual movements," Brain topography, vol. 12, pp. 177-186, 2000.

[20] A. Sivakami, and S. S. Devi, "Analysis of EEG for motor imagery based classification of hand activities," International Journal of Biomedical Engineering and Science, vol. 2, pp. 11-22, 2015.

Temporal Variables Disorder of The Gait Cycle in Parkinson's Disease

Adharaa Neelim Dewanjee[1], Quazi Delwar Hossain[2], Mohammed Shahriar Arefin[3]
Department of Electrical and Electronic Engineering
Chittagong University of Engineering and Technology, Chittagong -4349, Bangladesh
Email: [1]adharaa.neelim.dewanjee@gmail.com, [2]qdhossain@yahoo.com, [3]shahriar.cuet2013@gmail.com

***Abstract*— The biometric method which is involved with defining pattern of movement of human limbs is called Gait. Though gait has a fixed rhythmic pattern, in case of people who are affected with neurological disease such as- Parkinson's disease (PD) the gait pattern gets distorted from normal gait pattern. In this paper, temporal variables of gait cycles are analyzed to determine the deficit in the gait cycle in PD patients. Gait cycle of 10 PD patients and 10 Controlled subjects are examined in this paper. A comparative analysis was done in respect of temporal variables such as- Single Limb Support Time and Double Limb Support Time. From the examination and investigation, it is apparent that the temporal factors of the gait cycle of PD patients are sufficiently changed in regard to the controlled subjects, all the more explicitly PD patients invest more energy in the two limbs than a single limb of a gait cycle which is absolutely in switch of the controlled subject.**

Keywords*— *Parkinson's Disease, Gait Analysis, Temporal Variables, Signal Processing, Limb Support Time

I. Introduction

Parkinson's disease which was first discovered by Dr. Parkinson is a chronic neurological disorder which predominately occurs in elderly people and associated with different neurological difficulties [1], [2]. From an in-depth survey, Parkinson's foundation stated that about 10 million people are affected by Parkinson's disease around the globe [3].

Neurons found in Substantia Nigra, a region in the brain gets infected as a result of Parkinson's. These neurons are responsible for creating dopamine which in turn controls limb movement of a human. As dopamine secretion gets distorted due to damaged neurons, the people affected by Parkinson's lose control over their movement [3]. A person affected with Parkinson's may suffer from different difficulties such as- Tremors, Muscle rigidity etc. [4]. A common phenomenon occurs among Parkinsonians is the loss of control over the gait cycle.

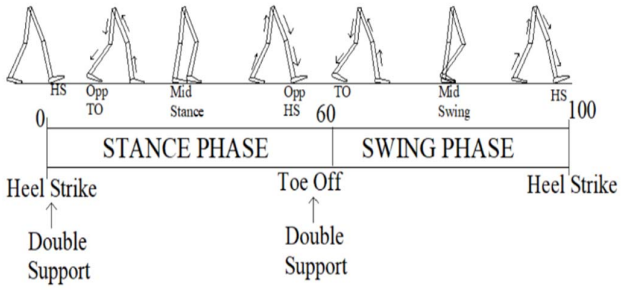


Fig 01- Illustration of Gati Cycle [HS=Heel Strike, TO=Toe Off, Opp=Opposite] [5]

Gait cycle is commonly defined as the walking cycle. In Fig. 01 a typical gait cycle is being portrayed, where stance phase (60% of gait cycle) of gait cycle is started when heel strikes and ends up when the toe is off to the ground and then swing phase (40% of gait cycle) starts [5]. Gait cycle is an important parameter to study human movement especially in the case of Parkinson's affected person. The extent of distortion can be determined by analysis of gait cycle.

The Gait Cycle can be described through two parameters: Spatial Variables and Temporal Variables. Spatial variable defines the gait cycle in terms of step length and stride length. Whereas, the temporal variable consists of several parameters: Single Limb Support Time, Double Limb Support Time and Speed. Single limb support time is defined as the time in which only one foot acts as body support. On the other hand, when body mass is carried by both feet the situation is considered as double limb support. The time spent in this posture is considered as double limb support time. In Fig. 02, temporal variables of gait cycle have been depicted.

Parkinson's disease is associated with different movement deficiencies. As Parkinson's is a chronic and long term disease, along with the time the disease gets more prominent and thus affects Parkinsonians lifestyle [7]. The Parkison's disease is responsible for truncated stride length, low walking speed and buffed double support phase [8], [11]. The situation can be improved by employing external influences such as cues [9]-[11].

There have been quite a number of studies on Parkinson's gait analysis but few of those studies focused on kinetic parameters. In one of the studies based on kinetic features, the authors examined patients under influence of external cues to rectify the spatiotemporal and kinematic parameters [12].

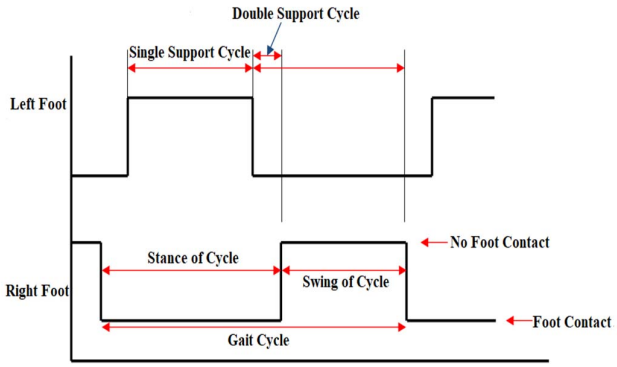


Fig. 02- Illustration of Limb Support Time [6]

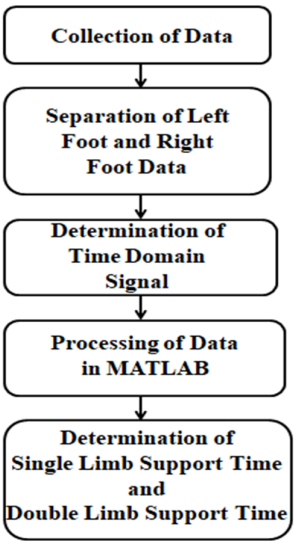


Fig. 03- Overview of Proposed Technique

Lewis showed that there is a large amount of variability present in kinematics and kinetics [13]. A quantitative analysis of gait was performed to show the difference between dopa-sensitive and dopa-resistance kinematic gait parameters. The result showed that the temporal variables are related to dopa-resistance [7]. The study lacked clarification in case of specific gait cycle parameter.
Another researcher observed an altered gait pattern in Parkinsonians. They showed that there is asymmetric timing and increased variability in gait which can be regarded as principal evidence of motor programming deficit in gait due to Parkinson's [14]. After that, it was observed that the movement time of PD patients is greater than the controlled healthy subject. External cues can improve temporal variability in case of Parkinson's disease affected person but still, there was evidence of distortion from normal subject [15]. Another researcher Pasluosta proposed a time series approach to focus on the human gait characteristics that shows a few unnecessary occasions, which are basically gait signals with pseudo-tedious nature [16]. On the other hand, researcher Chen depicted different methods for recognizing critical gait occasions which are recurring [17]. Besides, Mirelman investigated a unique signal for identification of specific parts of body with gait stability and normality [18].

In this paper, a method is proposed to quantify the gait cycle fluctuations of Parkinsonians in terms of temporal variables like- single limb support time, double limb support time and speed. The analysis also consists of a comparative study of temporal variables between controlled subjects and Parkinsonians to portray the difference between healthy subject and parkinsonians.

II. Methodology

The proposed procedure can be separated into three sections: Data Collection, Analysis of Data and Comparative Study between Healthy Subjects and Parkinson influenced people. The overall components of this investigation are portrayed in Fig. 03.

The data acquisition part comprises the gathering of information from Physionet Database named Gait in Parkinson's disease [19]. The obtained signals are vertical ground reaction force signals. The signals are fetched from 16 sensors divided into two equal number portions placed beneath both the right and left foot. The acquired data are separated into left foot and right foot data. The information is investigated independently in MATLAB to decide Time Domain Signal. By analyzing the stance phase and swing phase, it is discovered that there are just two positions in a Gait cycle when a subject needs double limb support: Heel strike position and Pre-swing Position. In light of these perceptions, time for single limb support and double limb support are determined for concerned subjects. For double limb support time, the vertical ground reaction force of left foot and the vertical ground reaction force of right foot must be greater than zero because both feet must exert the force on the ground on this case. Utilizing this logic, a square wave is created which represents double limb support. Then, difference equation topology is applied to generate a spike which represents changes of square waves of two foot. Basically, for a gait cycle, four spikes are found because in one gait cycle there are two positions when a person is in double limb support and for one double limb support there are two spikes. By using this logic, the index of the spikes is found and then the index of the first spike is subtracted from the second spike and again the same procedure is followed for fourth spike and third spike. After that, both index differences are added. As the sampling rate (100Hz) is known beforehand from the data website, the index is divided with sampling frequency to get the double limb support time. Single limb support time for a gate cycle is the difference between the stride time of a foot and double limb support time. Stride time of a certain foot is calculated by using the difference between the indexes of consecutive steps.

III. Result and Discussion

A. Data Delineation

Total data of 20 subjects were taken from the database. Among 20 subjects, 10 of the subjects were affected by Parkinson's and the rest 10 people were controlled subjects. The gender of considered subjects is male. The subjects were not under any external cues during the data acquisition phase. The age range of Parkinson's disease affected subjects is between 55 years to 70 years, whereas, the controlled subjects age spans from 60 to 75 years.

Table I: Temporal Data of Controlled Subjects

Controlled Subject Number	Mean(Double Limb Support Time) (second)± Variance	Mean(Single Limb Support Time) (second) ± Variance	Speed (m/s)
Subject 1	0.2949±0.0152	0.9518±0.0081	1.051
Subject 2	0.2338±0.0038	0.9434±0.0015	1.121
Subject 3	0.2093±0.0011	0.8595±0.0011	1.164
Subject 4	0.3112±6.6859e-04	0.9232±5.0630e-04	1.073
Subject 5	0.2441±9.9912e-05	0.8825±2.0202e-04	1.16
Subject 6	0.2901±2.7856e-04	0.9299±5.0205e-04	1.00
Subject 7	0.3073±5.2308e-04	0.8776±9.4772e-04	0.99
Subject 8	0.2014±4.7058e-04	0.8512±5.0585e-04	1.29
Subject 9	0.1918±0.0022	0.8254±0.0013	1.515
Subject 10	0.1720±0.0053	0.9262±0.0046	1.415

Table II: Temporal Data of Parkinson's Patient

Parkinson's Patient's Number	Mean(Double Limb Support Time) (second)± Variance	Mean(Single Limb Support Time) (second)± Variance	Speed (m/s)
Patient 1	0.8388±0.0012	0.3177±0.0023	0.848
Patient 2	0.9830±0.0013	0.2705±9.7951e-04	0.802
Patient 3	0.8044±0.0042	0.3111± 0.0043	0.987
Patient 4	0.8690±0.0025	0.3667±0.0034	0.825
Patient 5	0.9113±5.8292e-04	0.1941±9.9563e-04	1.013
Patient 6	0.7786±0.0032	0.2657±0.0033	0.832
Patient 7	0.7179±0.0086	0.2500±0.0044	0.785
Patient 8	0.8202±0.0080	0.2047±0.0057	1.112
Patient 9	1.0127±0.0229	0.5070±0.0291	0.413
Patient 10	0.7657±0.0053	0.2186±0.0063	0.894

The collected data were then converted into temporal variables such as- Single limb support time, Double limb support time and Speed. The temporal data are presented in Table-I and Table-II. Table-I represents controlled subjects and Table-II narrates data of Parkinson's disease affected persons and Table-III demonstrates the comparative study of temporal data. The single limb support time and double limb support time data for all subjects are depicted in Fig. 04 and Fig. 05. Fig. 04 illustrates data of controlled subjects, whereas, Fig. 05 delineates data of Parkinson's affected patients.

B. Comparative Analysis

After calculating the temporal variable data, the time required for single limb support and double limb support were compared between healthy controlled subjects and disease affected subjects.

The limb support times of controlled subject 01 and 02 are elucidated in Fig. 06 and Fig. 07. It is visible that single limb support time is greater than double limb support time in case of the controlled subject which validates the observation where it was found that only two positions in a gait cycle are based on double limb support. In Fig. 06, it is observed that for controlled subject 01 single limb support time is in between 0.3s to 1.1s with mean value 0.9518s and double limb support time is in between 0.2s to 1.2s with mean value 0.2949s. Similarly, in Fig. 07 for controlled subject 02 single limb support time is in between 0.75s to 1.5s with mean value 0.9434s and double limb support time is in between 0.19s to 0.55s with mean value 0.2338s, whereas, the speed of controlled subject 01 is 1.051m/s and controlled subject 02 is 1.121m/s.

Fig. 08 illustrates the temporal variables for Patient-03 which shows that the double limb support time is higher than single limb support time. The mean values of times are 0.8044s and 0.3111s respectively for double limb and single limb. The average speed clocks at 0.987m/s which are evidently lower than controlled subjects.

Table III- Comparative Analysis of Temporal Data

Subject	The range of Double Limb Support Time (second)	Range of Single Limb Support Time (second)	The range of Speed (m/s)
Controlled Subjects	0.1720-0.3112	0.8254-0.9518	0.99-1.515
Patients	0.7179-1.0127	0.2047-0.3667	0.413-1.112

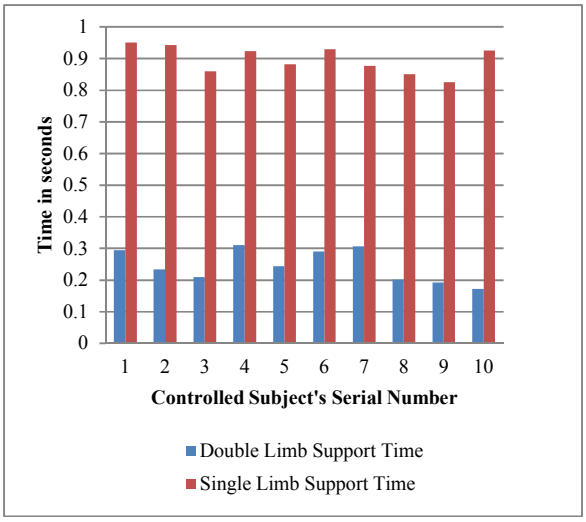


Fig. 04- Temporal Variable Data of Controlled Subjects

Fig. 09 renders the single limb support time and double limb support time for Patient 10. The figure shows that the mean value of time taken for single and double limb support is 0.2186s and 0.7657s respectively, whereas the calculated average speed is 0.894m/s.

From Fig. 6 – Fig. 9, mean value of the single limb support time and double limb support time are shown and it is noticeable that for the healthy subjects (Fig. 6 and Fig. 7) the single limb support time and double limb support time vary closer to their mean but in case of the Parkinson's patients (Fig. 8 and Fig. 9) they vary a lot from the mean.

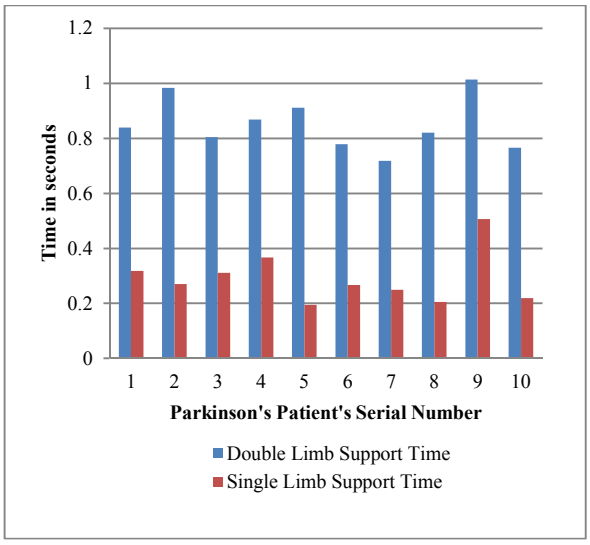


Fig. 5- Temporal Variable Data of Parkinson's Patients

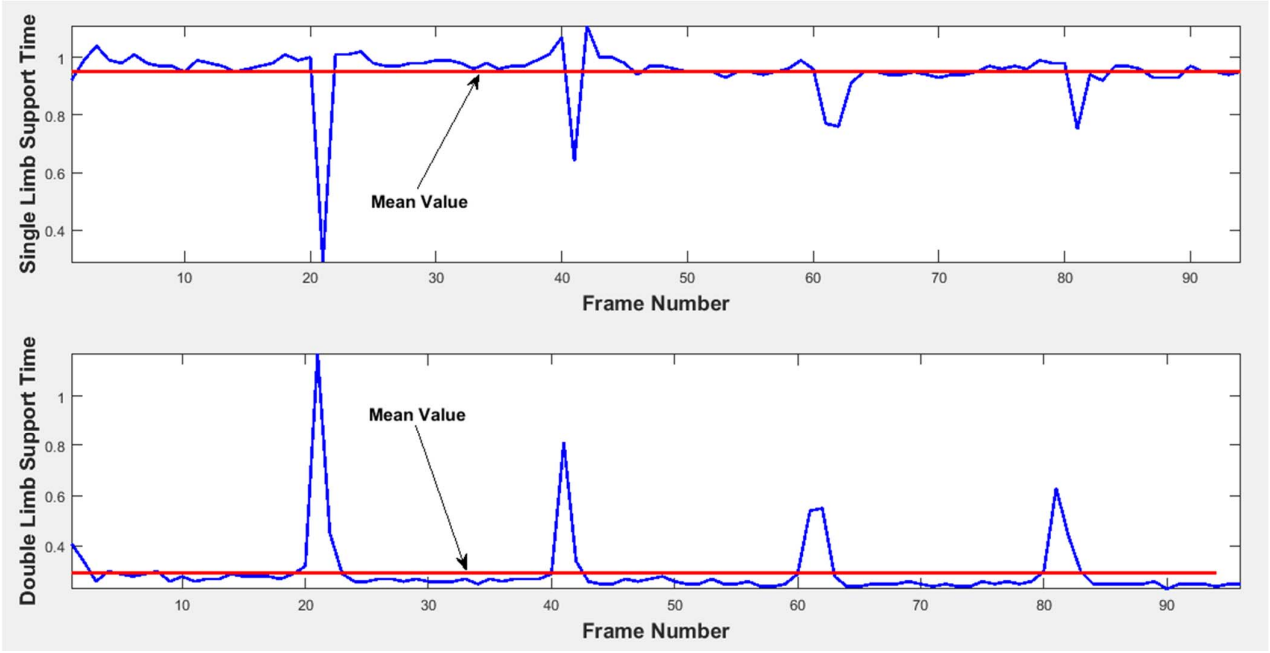


Fig. 06- Limb Support Time of Controlled Subject 01

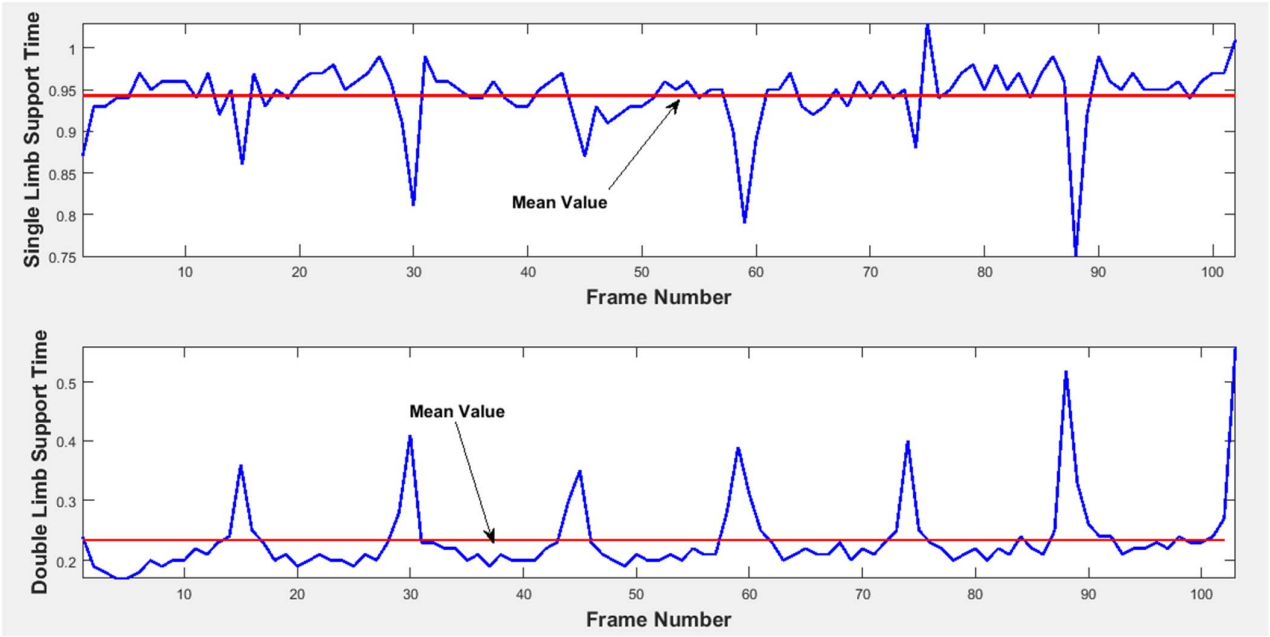


Fig. 07- Limb Support Time of Controlled Subject 02

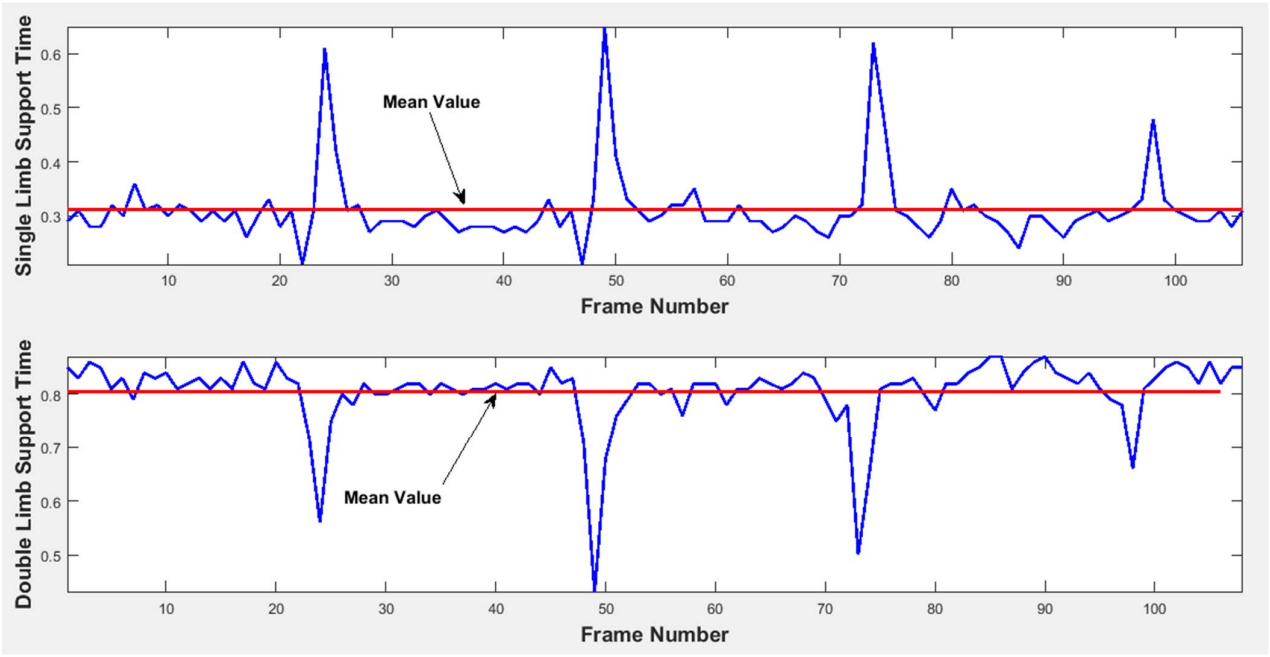


Fig. 08- Limb Support Time of Patient 03

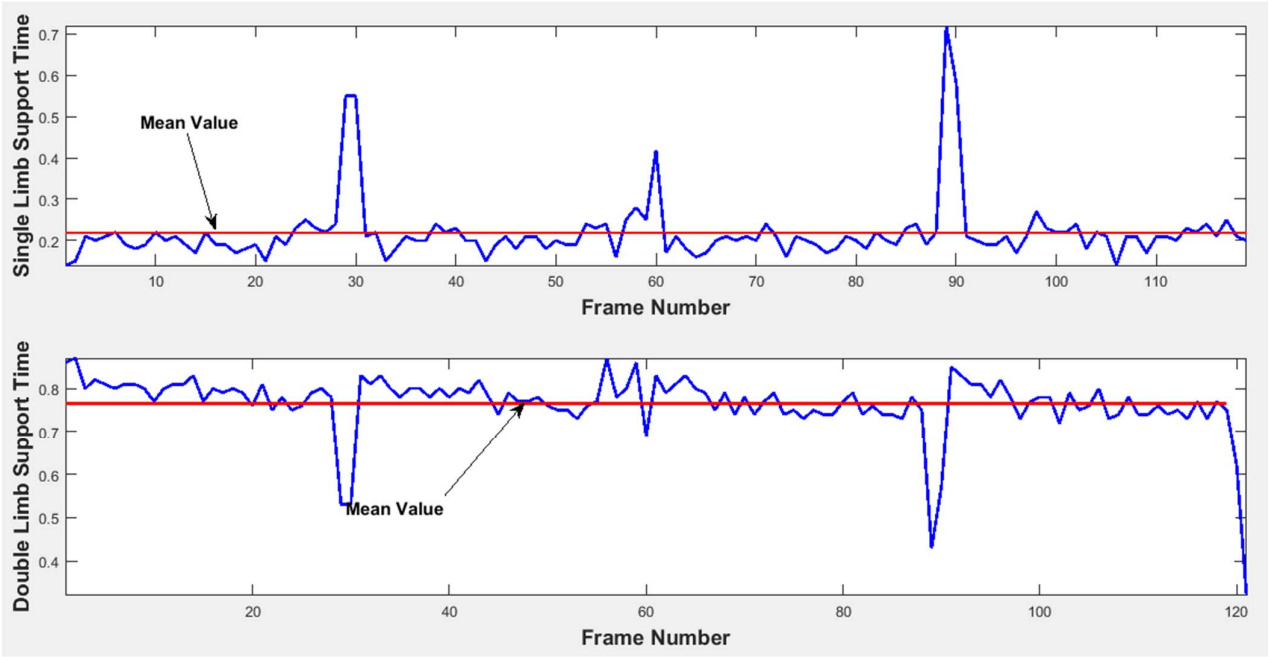


Fig. 09- Limb Support Time of Patient 10

From the precedential analysis, it is evident that double limb support time is higher than single limb support time in case of Parkinson's disease affected persons. The opposite is seen in the case of controlled subjects. The speed of Parkinson's affected subjects is lower than controlled subjects. Table-III shows discrimination among the range of temporal variables. All these data prove that the patients suffering from Parkinson's suffer greatly due to gait cycle distortion.

The results found in this comprehensive study based on temporal variable might be proved to be vital in case of assessment of gait cycle variation in Parkinsonians. In future, these data can be used with Machine Learning algorithms to predict the trend of a patient affected with Parkinson's and provide external cues according to the needs of the patient.

IV. Conclusion

In this study, a comparative analysis of temporal variable data of Parkinson's affected patients and controlled subjects are depicted. By juxtaposing all calculated data, it is discernible that the double limb support time in case of Parkinsonians is greater than controlled subjects. The reason behind this distortion is related to loss of control over the gait cycle in case of Parkinson's disease affected persons. The loss of walking speed is also visible for Parkinson's affected person. The results of this study can be used to develop assistive devices with feedback for Parkinsonians. Also, the result in conjecture with spatial variables could provide more accurate situation assessment for patients affected by Parkinson's.

Acknowledgments

The authors would like to express their gratitude to Department of Electrical and Electronic Engineering, Chittagong University of Engineering and Technology, Bangladesh for their continuous support and assistance throughout the completion of this study.

References

[1] DeMaagd G, Philip A., "Parkinson's Disease and Its Management: Part 1: Disease Entity, Risk Factors, Pathophysiology, Clinical Presentation, and Diagnosis.," *P T.* 2015; 40(8):504-32.

[2] Mei-Fen Hsieh, Zhao-Xiong Yan, "Symptoms and diagnosis of Parkinson's Disease", Family Medicine& Primary Medical Care, vol. 25, no.1, pp. 10 14, 2010. (DOI:10.6965/FMPMC.201001.0010)

[3] (2018) Parkinson's disease Foundation. [Online]. Available: http://www.pdf.org/about_pd.

[4] (2018) Parkinson's disease Information Page. [Online]. Available: www.ninds.nih.gov/Disorders/All-Disorders/Parkinsons-Disease Information- Page.

[5] (2018) Gait Cycle Visual Representation and Definition. [Online]. Available: https://fadavispt.mhmedical.com/Content.aspx?bookId=1862&sectionId=136086727.

[6] M. Kwon, Y. Kwon, Y. Park, and J. Kim, "Comparison of gait patterns in elderly fallers and non-fallers," vol. 26, pp. 427–436, 2018.

[7] M. E. Morris, "Movement Disorders in People With Parkinson Disease: A Model for Physical Therapy," *Phys. Ther.*, vol. 80, no. 6, pp. 578–597, 2000.

[8] Blin O, Ferrandez AM, Pailhous J, Serratrice G., "Dopa-sensitive and dopa-resistant gait parameters in Parkinson's disease," J Neurol Sci 1991;103:1-54.

[9] G. C. Mcintosh, S. H. Brown, R. R. Rice, and M. H. Thaut, "Rhythmic auditory-motor facilitation of gait patterns in patients with Parkinson ' s disease," J. Neurol. Neurosurg. Psychiatry, pp. 22–26, 1997.

[10] M. H. Thaut, G. C. McIntosh, R. R. Rice, R. A. Miller, J. Rathbun, and J. M. Brault, "Rhythmic auditory stimulation in gait training for Parkinson's disease patients," Mov. Disord., vol. 11, no. 2, pp. 193–200, 1996.

[11] L. E. Dibble, D. E. Nicholson, B. Shultz, B. A. MacWilliams, R. L. Marcus, and C. Moncur, "Sensory cueing effects on maximal speed gait initiation in persons with Parkinson ' s disease and healthy elders," vol. 19, pp. 215–225, 2004.

[12] M. E. Morris, J. McGinley, F. Huxham, J. Collier, and R. Iansek, "Constraints on the kinetic, kinematic and spatiotemporal parameters of gait in Parkinson's disease," Hum. Mov. Sci., vol. 18, no. 2–3, pp. 461–483, 1999.

[13] G. N. Lewis, "Stride length regulation in Parkinson's disease: the use of extrinsic, visual cues," Brain, vol. 123, no. 10, pp. 2077–2090, 2000.

[14] R. Baltadjieva, N. Giladi, L. Gruendlinger, C. Peretz, and J. M. Hausdorff, "Marked alterations in the gait timing and rhythmicity of patients with de novo Parkinson ' s disease," vol. 24, no. July, pp. 1815–1820, 2006.

[15] N. Teasdale, J. Phillips, and G. E. Stelmach, "Temporal movement control in patients with Parkinson ' s disease," pp. 862–868, 1990.

[16] C. F. Pasluost, H. Gassner, J. Winkler, J. Klucken, B. M. Eskofier, "An emerging era in the management of parkinson's disease: wearable technologies and the internet of things," Biomed Health Inform, 19(6), 1873-1881, 2015.

[17] S. Chen, J. Lach, S. Member, B. Lo, and S. Member, "Toward Pervasive Gait Analysis With Wearable Sensors : A Systematic Review," IEEE J. Biomed. Heal. Informatics, vol. 20, no. 6, pp. 1521–1537, 2016.

[18] Mirelman A., Bernad-Elazari H., Thaler A., Giladi-Yacobi E., Gurevich T., Gana-Weisz M., Saunders-Pullman R., Raymond D., Doan N., Bressman S.B., Marder K.S., Alcalay R.N., Rao A.K., Berg D., Brockmann K., Aasly J., Waro B.J., Tolosa E., Vilas D., Pont-Sunyer C., Orr-Urtreger A., Hausdorff J.M., Giladi N., "Arm Swing as a Potential New Prodromal Marker of Parkinson's Disease," Movement Disorders, 31(10), 1527-1534, 2016.

[19] (2018) Physionet Database of Gait in Parkinson's Disease. [Online]. Available: https://physionet.org/pn3/gaitpdb/.

Design and Development of a Portable Recording System for Simultaneous Acquisition of SCG and ECG Signals

Tauhid Anwar[1], Mohammad Zubair[1], Mohammad Dehan Rahman[2], Nazmus Sakib[1], Md. Ahsan-Ul Kabir[1], Tasnuva Faruk[3], and Md. Kafiul Islam[1,*]

[1] *Dept. of Electrical and Electronic Engineering, Independent University, Bangladesh (IUB), Dhaka, Bangladesh*

[2] *Dept. of Electrical and Electronic Engineering, Islamic University of Technology (IUT), Gazipur, Bangladesh*

[3] *School of Public Health, Independent University, Bangladesh (IUB), Dhaka, Bangladesh*

[*]Corresponding Author's Email: kafiul_islam@iub.edu.bd

Abstract— **In-vivo observing of patient's heart mechanics by means of seismocardiography (SCG) was recently rediscovered and presented (Demonstrated). Most of the time, seismocardiography is recorded utilizing 2D or 3D accelerometers set on patient's sternum. This research represents the ability and performance development of a PVDF piezoelectric film transducer by placing them in multiple positions on the subject to record real time mechanical movement of the heart (seismocardiogram and heart sounds) and furthermore of the breath related thorax movement. Patient's ECG signal was also recorded simultaneously to give a time reference of the cardiovascular movement and better assessment of SCG signal acquisition. The piezoelectric sensor has been able to record respiratory signal, seismocardiography and heart sounds all along. These signals later have been processed on MATLAB by constructing simple digital filter. The seismocardiogram waveform gives off an impression of being reliant on the different periods of the respiratory cycle. The recorded signals of piezoelectric sensor show important data such as pulse rate, respiratory cycle (Inhale and exhale), timing of mechanical occasion inside the heart cycle, and so forth.**

Keywords— Seismocardiography (SCG), ECG, Piezoelectric sensor, Arduino UNO, MATLAB.

I. INTRODUCTION

The Seismocardiogram (SCG) was first observed in 1961 by Bozhenko, and later applied in clinical studies for the first time in 1991 by Salerno and Zanetti [1], [2]. Seismocardiography (SCG) is a non-obtrusive strategy in view of account the vibration produced by the mechanical constriction of the heart. The pressure waves created by the siphoning activity of the heart spread through the thorax and can be distinguish noninvasively at patient's external chest divider. Customarily, vibration is recorded by putting an accelerometer on the sternum of the patient while lying prostrate, in any case, vibrations can be really estimated everywhere throughout the thorax (particularly on the precordial territory) and for any position on body. Regardless of whether SCG is generally newer, the recording of body developments related with cardiovascular movement was proposed a lot before, with respect to the ballistocardiogram (BCG), Phonocardiogram, and many other methods [3], [4], [5], [6]. A reestablished enthusiasm for these procedures is because of the vast accessibility of ease sensor (for example MEMS accelerometers PVDF piezoelectric film) and the low price for manufacturing them. SCG signals contains highlights that compares to explicit occasions inside the heart cycle as mitral conclusion, aortic opening, and so on., specifically synchronous recording of SCG and echocardiogram permitted to perceive explicit occasions timing and a SCG wave classification was proposed . Besides, SCG timing can be identified with the essential heart sounds (S1 and S2) estimated on phonocardiogram (PCG) [7]. The SCG signal range cover the infrasonic range (for example frequencies underneath 25 Hz) in light of the fact that SCG is basically produced by the movement or mechanical activity of heart through the chest wall, while the capable of being heard PCG flag is for the most part created by the blood stream. As an option in contrast to the utilization of accelerometers, the capacity to record the developments of the chest through diverse sensors has been investigated but more importantly their position on the subject thorax has been deeply researched and examined. This investigation focused around the capacity to quantify precordial acceleration on patient's chest by using different piezoelectric sensor but mostly focused on thin polymer coated piezoelectric sensors also known as PVDF piezo sensor. Specifically, the examination aims to evaluate the capacity of exceptional wide-band sensors and detect appropriate position on human body for them to measure moderate respiratory developments, seismocardiogram and the heart sounds [8].

II. MATERIALS AND METHODS

A. Sensors and their Placement

Polyvinylidene fluoride (PVDF) polymeric film not only shows clear and strong piezo-electricity but also pyro- and ferroelectric properties and therefore it is used as a sensor or transducer material whose application is found in enormous sectors like biomedical devices, mechanical, metrological, etc. Here, we have used two PVDF sensor films (CAT-PFS0006 Vibration sensor) manufactured by TE Connectivity which is based on sensitive but robust PVDF piezo film. This particular sensor provides high sensitivity to vibration while minimizes external acoustic noise. It has a stress constant of $216\ N/m$.

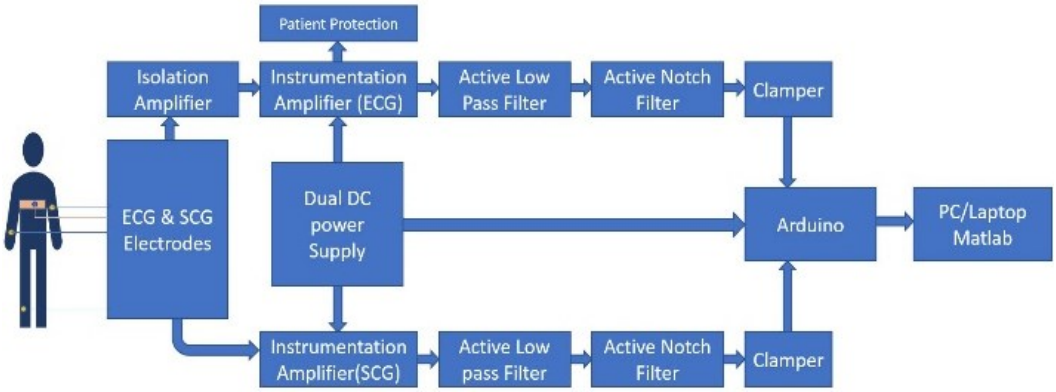


Fig. 1. Block Diagram of the proposed portable system for simultaneous acquisition of SCG and ECG signals.

This sensor can produce potential difference of $10\,mV - 100\,V$ depending on force and circuit impedance.

The sensor is mounted in a cantilever arrangement, allowing the sensing area to vibrate up and down. We have also added a small weight to the end of the sensor for getting greater sensitivity to lower frequencies. Cantilever arrangement will allow the piezo element to be deflected and this can be used to detect a vibrating object which is our subject chest when the element is flexed. Figure 1 shows the visual arrangement of cantilever system on a PVDF piezo sensor.

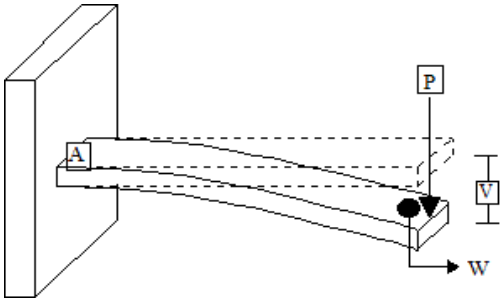


Fig. 2. Cantilever arrangement.

We have then attached the sensor on the subject's body using a belt for fixing the position. In usual cases the sensor is placed on middle of the sternum. We have moved the sensors couple of time in different places on sternum and found that putting the sensor on middle of the sternum may not give proper signals that contain all the valuable artifacts. Rather placing the sensors on the sternum it's better to set it on the left side of sternum which is known as *costal cartilage* has better chance of finding good signals. The weight on the sensor was attached in a way that it stayed touched with the subject's chest but no so tightly but tight enough to give vibration boost to the sensor. In this way the subject didn't need to stay laid on a surface but sit comfortably on a chair. The output voltage signal from the sensor was amplified by an instrumentation amplifier. At the same time ECG was also recorded for time and cardiac cycle reference. The amplified signal then passed through a low pass filter. Then another filter has utilized after the low pass filter known as notch filter. This process has some benefit over having just raw output signal. Pre-filtered signals have better visual on screen that has been recording and also has potential information whether the signal is really available or just random RF noise. Both ECG and SCG signal then being recorded by Arduino mega. The recorded signals then filtered out for extraction of potential information for the respiratory signal, SCG signal, and the heart sounds with the frequency bands of $0.05 - 1\,Hz$ for, $1 - 50\,Hz$, and $50 - 150\,Hz$ respectively [9], [10].

B. Hardware Implementation

Figure 2 shows the potable system design that has been used. The system is has the capability of running in battery power or using an AC-DC wall adapter. Both ECG and SCG circuitry contain Instrumentation amplifier, low pass filter(suppressing higher frequency noise) and notch filter (to minimize power line interference), voltage adder for shifting signal's negative portion to the positive side to make sure ADC can read analog input. Lastly Arduino mega is being used as a data acquisition device to simultaneously record both ECG and SCG and store it in Computer for further processing.

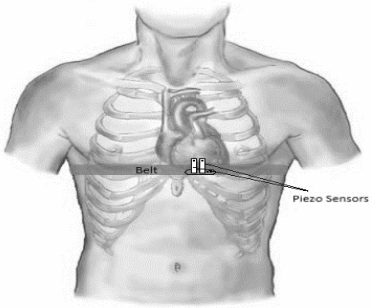


Fig. 3. Piezo sensor placement.

C. Instrumentation Amplifier

Two instrumentation amplifiers are used in this project to extract the ECG and SCG signal from subject. The instrumentation amplifiers used in this research are AD620 and INA121 for pre-amplifying ECG and SCG signals respectively [11] [12], [13]. Both amplifiers have high accuracy, but in comparison with AD620, INA 121 has bit more accuracy and advantages. INA121 has 4pA bias current while AD620 has 1nA. This gives the INA121 an impressive common mode noise cancelation about 106dB at G = 100. Both of this amplifier require an external resistor for setting the gain between 1 and 10,000. The gain of the amplifier (for AD620) can be found using the following formula:

$$A_v = 1 + \frac{49.4k\Omega}{R_{gain}} \qquad (1)$$

For INA121, the gain of the amplifier can be computed by using the formula given below:

$$A_v = 1 + \frac{50k\Omega}{R_{gain}} \qquad (2)$$

The following diagrams shows the schematic diagram of ECG and SCG circuit.

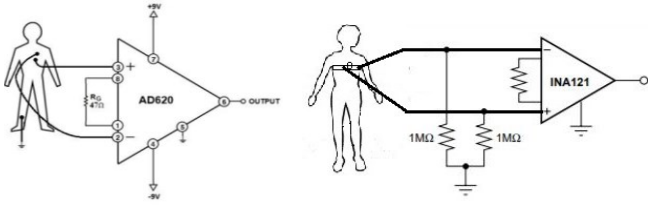


Fig. 4. Circuit Schematic used to extract the ECG (left) and SCG (right) signal from the chest.

D. *Isolation Amplifier*

Isolation amplifier, an op-amp based circuit which is also known as unity-gain amplifier, provides isolation of one side from the other side by optically transmitting data. This is done in such a way that the power is not wasted in any side of the circuit. The isolation amplifier has unity gain. For protection of the subject, an isolation amplifier is used in this research and isolated the subject form the circuit for safety purpose. Here Broadcom (HCPL-7800) general purpose isolation amplifier is used. The high CMR ability of this isolation amplifier gives the required precision and also stability. This IC design is based on fully differential circuit topology Functional diagram of the IC is given below.

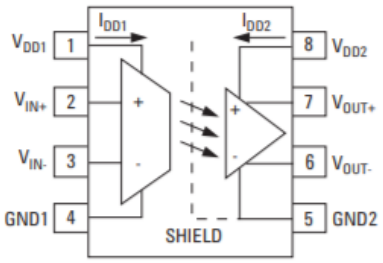


Fig. 5. Pin diagram of the isolation amplifier.

E. *Low-Pass Filter*

In this research two low pass filter was used to eliminate noise from the signals. One low pass filter is used for ECG and other one for SCG. A second order low pass filter having cutoff frequency of 80Hz is used for ECG to obtain decent frequency response. The cutoff frequency can be set much lower than 80Hz but there is a huge chance of losing high frequency components of the actual ECG signals. Cut off frequency of second order can be calculated by using the formula given below.

$$f_c = \frac{1}{2\sqrt{R_1 R_2 C_1 C_2}} \qquad (3)$$

A 10th order low pass filter used for SCG for getting clearer and better frequency response as SCG sensors tends to detect a lot of vibration from outer environment. The low pass filter used for SCG has 150Hz cut off frequency. Typical SCG signal contain three type of valuable information like Respiratory signal, Seismocardiography signal and Heart sound. To preserve all this information, we need to set the cutoff frequency to 150Hz. Higher filter order was used for making cutoff region as close as ideal response. Again, the cutoff frequency can be calculated by using the formula given below.

$$f_c = \frac{1}{2\sqrt{R_1 R_2 \ldots R_{10} \times C_1 C_2 \ldots C_{10}}} \qquad (4)$$

F. *Notch Filter*

To remove the 50Hz power line interference some sort of band rejection filter was needed. This research has used an active notch filter to reject the power line noise as the system has dual dc power system that draws power form main line. This band rejection filter is second order. Bandwidth was fixed at 2Hz and the quality factor Q was 25. Bainter topology was followed all along while designing such filter. The circuit schematic has given below.

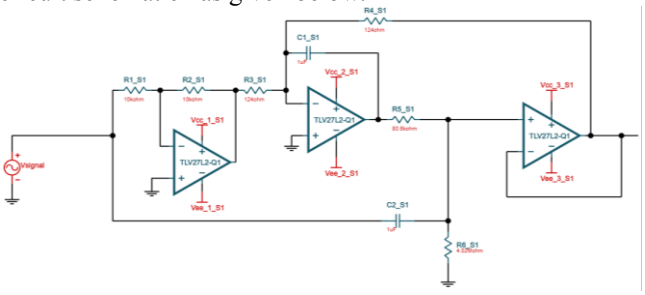


Fig. 6. Circuit schematic of the 50-Hz notch filter used in our system.

G. *Clamper*

ECG and SCG signals have both positive and negative cycle. However, Arduino can only sense the positive cycle of an analog input. In order to feed both positive and negative cycle of the signal, the level must be shifted to the positive side. To do that the system needs a clamper circuit which can be two types: active and passive clamper. In this research, an active clamper circuit was used. Three general purpose Op-Amps are used for construction of the clamper circuit. One for voltage buffer circuit and other two for adding the dc offset voltage to shift the level. The clamper circuit schematic is shown in figure 7.

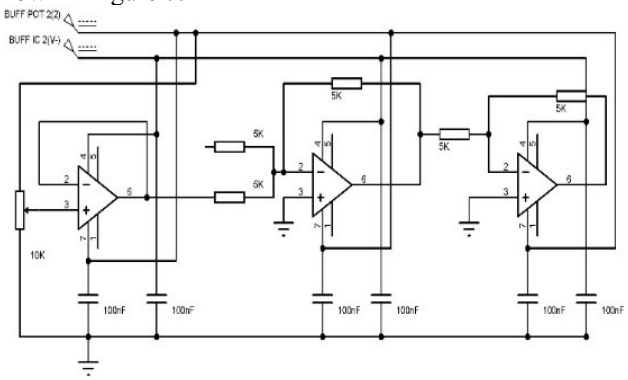


Fig. 7. Clamper circuit scematic used in the proposed system.

III. INTERFACING WITH PC

The ECG and SCG circuit that has been used for this research was interfaced with a Laptop/PC through an Arduino Mega. Signals from both ECG and SCG was fed to different analog pin of Arduino mega. Arduino was programed in such way that it can read both analog input data, perform ADC process and store the data in the PC/Laptop and repeat the whole processes as long as possible. Basically, this process runs in a continuous loop. This Arduino recording process was done by using telemetry viewer. 800 Hz sampling frequency was used in this process. Once the data was store in the PC/Laptop, further processing and analysis can be done. MATLAB is a resourceful and suitable software when it comes to process the raw signals and analyze them. Digital signal processing can be easily done to remove unwanted noise and interference as well as to analyze the signal characteristics to extract important features in terms of statistical parameters for diagnosis purpose.

The data from the Arduino was collected on pc/laptop using a special software name Telemetry Viewer. This software has lot of potential to work with such kind of miniature microcontroller device. The data was sent to PC/Laptop using serial communication at a baud rate of 115200. The resolution of the internal ADC of Arduino is 10-bit which means it can detect 2^{10} =1,024 different analog levels. Telemetry viewer software has capability to store large amount of data in a single shot. For experimental purpose every data that have been collected from different subject was length of 5 minutes and having 1040574 samples on average.

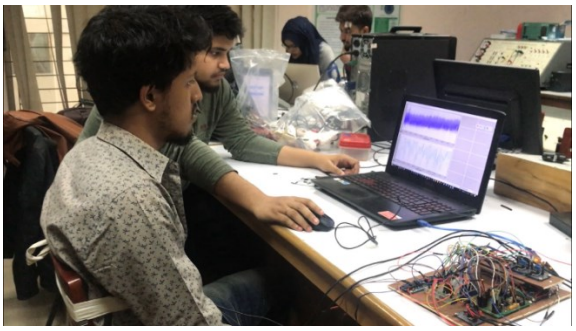
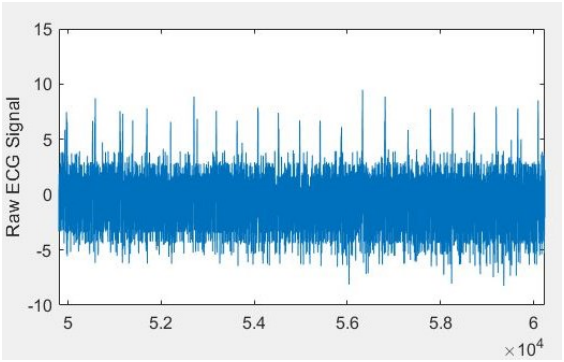


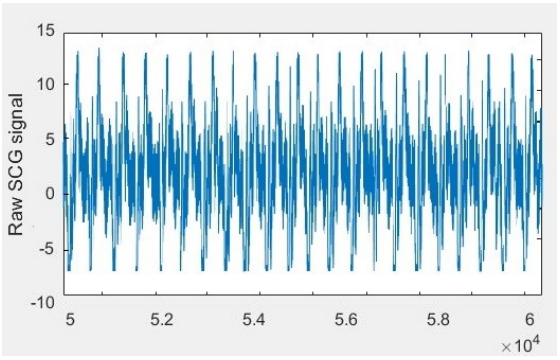
Fig. 8. Experimental set-up to record the simultaneous SCG and ECG.

IV. RESULTS

Figure 9 shows some raw signal recordings of ECG and SCG below.



(a)



(b)

Fig. 9. Recorded raw signals: ECG (a); PVDF piezo output (b).

Figure 10 represents the raw signals both ECG and SCG decomposed into the four frequency component which is being plotted subsequently. As for the respiration signal on the top trace has frequency band of $0.05 - 1Hz$. Second trace is SCG waveforms having a frequency band of $1 - 50Hz$. Third Trace is the ECG signal that has been taken as a time and cardiac cycle referance(R peak has been considred) [14]. Finally heart sound is plotted at the bottom of the figure and the frequency bad is $50 - 150Hz$. Some portion of the cleaned ECG and SCG signals are given below.

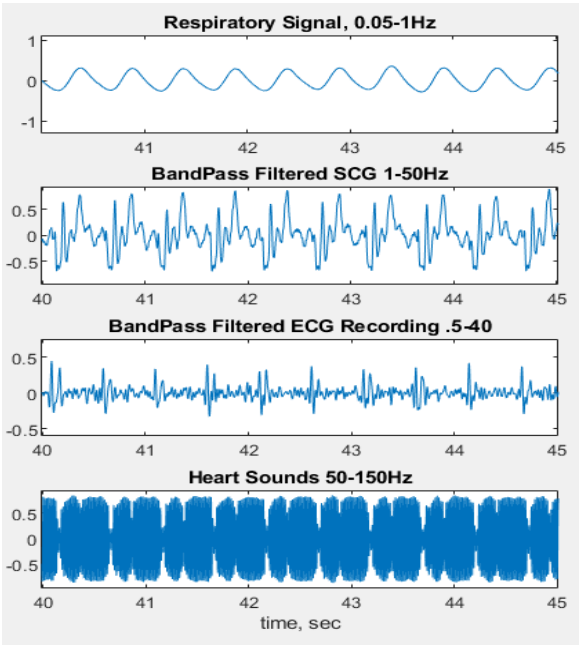


Fig. 10. The output signal of Piezo is decomposed in different frequency bands: from top to bottom: ECG; Respiration; SCG; and Heart Sounds respectively.

A portion of cleaned ECG and SCG signal are given in figure 11. The clean shows primary visual of SCG signal. To verify that the signal pattern has seismocardiographic wave the result is being compared with an accepted seismocardiographic wave form that has been represented by Amirtaha Taebi (2018) proposed that during each cardiac cycle, typical SCG contains actually two main events that can be called the first SCG (*SCG1*) and the second SCG (*SCG2*). Analyzing signals form this figure we found that it also contains these two events that have been identified. In the figure, the cardiac cycles are shown.

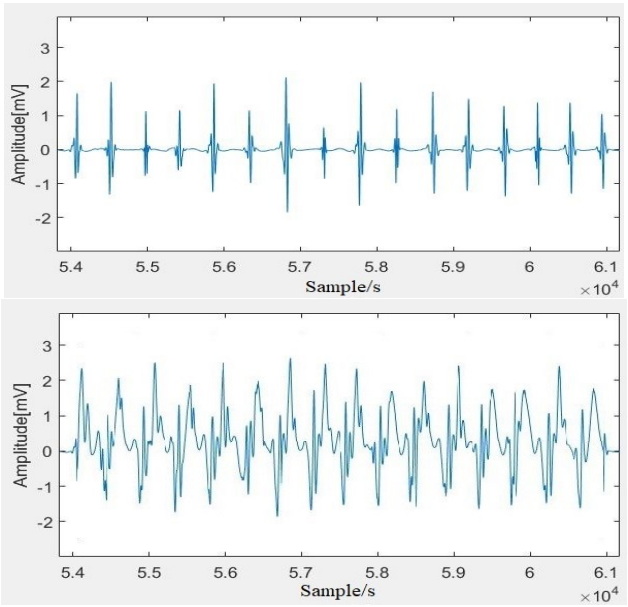


Fig. 11. Clean, recorded signals: ECG (top trace); SCG (lower trace).

After analyzing the cardiac events a single waveform is taken from filtered signal for pointing out all the annotations that have been previously proposed and took as a reference signal to compare with filtered signal [15].

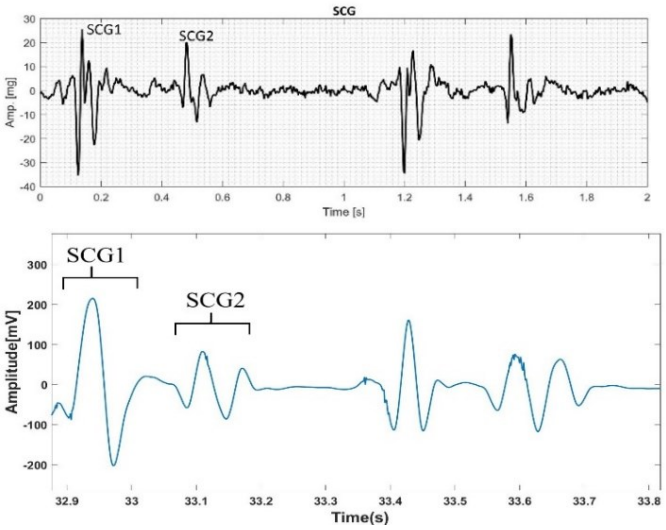
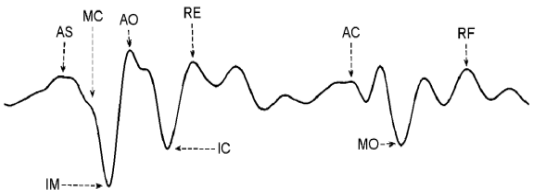
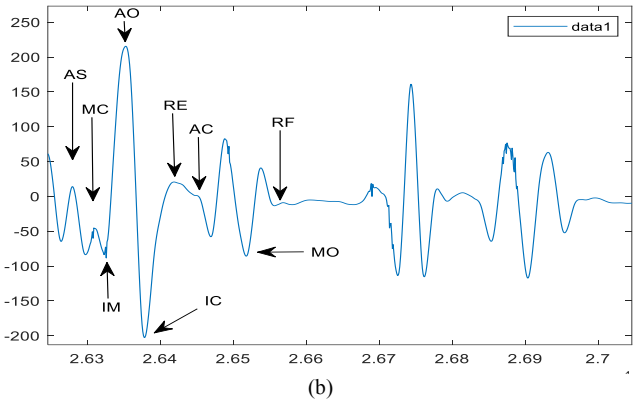


Fig. 11. SCG signal waveform shape comparison using standard annotations

Below the comparison result shows that all the annotations are present in the cleaned signal. Visual comparison is given in figure 12.



(a)



(b)

Fig. 12. Annotation of seismocardiographic signals using the annotation scheme proposed by Crow et al. (1994). (a) along with our recorded clean signal (b).

A synchronized averaging of the SCG waveforms identifying with individual heart cycles by considering the R-peak of ECG as time reference has been performed to compare the transient connection of ECG with SCG signal.

Figure 13 shows the results obtained from the operation of synchronized average made of 10 consecutive cardiac cycles. The average signals of ECG and SCG show the continuity of the annotations of the seismocardiographic signals. Comparing this average signal with another standard annotation scheme proposed by Anh Dinh (2011) also verify this research result [16]. Comparison of these two signals are given in figure 14.

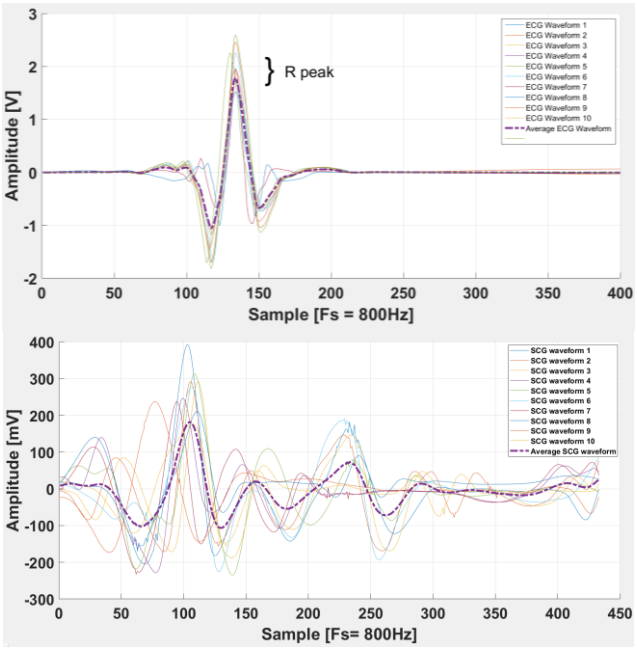


Fig. 13. Synchronized average of ECG (top) and SCG (bottom) signals over 10 cardiac cycles. The R-peak of ECG was considered as time reference. Average traces are highlighted in bold dashed lines.

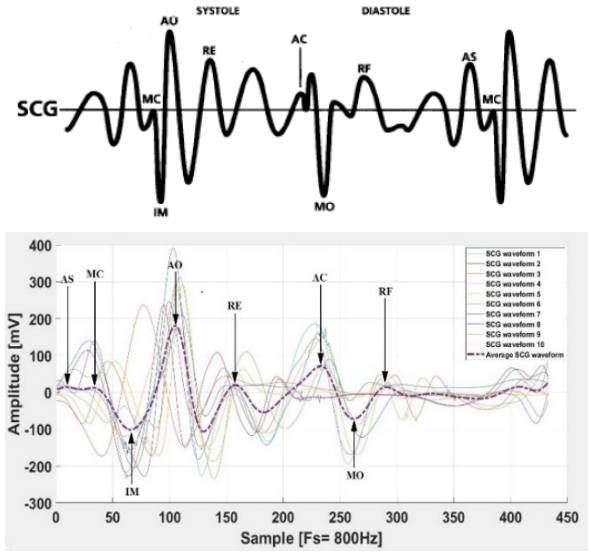


Fig. 14. Annotation of seismocardiographic signals using the annotation scheme proposed by Dinh (2011) et al. (Top trace) along with our recorded clean signal (Lower trace)

V. Conclusion

Two cardiovascular assessment systems were implemented in this research. The system was relatively simple and easy to handle. Current system implementation demonstrated the feasibility of SCG by using piezo sensors. SCG can be a great complementary analysis method along with ECG as it promises a lot handier information on heart's mechanical activity. Using two sensors instead of one promises more reliable and less noise full signal. Despite numerous studies in the field, there are still several limitations that must be addressed in the future work. These limitations include, but are not limited to:

- SCG variability (Age, Gender, thickness of chest wall Tissue)
- More realistic simulated SCG signals
- A comprehensive library and database of SCG signals
- Investigating correlation between SCG and other signals that are related to the mechanical aspects of the cardiac activity such as PCG and BCG:

References

[1] O. Kinoshita, G. . Fontaine, F. . Rosas, J. . Elias, T. . Iwa, J. . Tonet, G. . Lascault and R. . Frank, "Time- and Frequency-Domain Analyses of the Signal-Averaged ECG in Patients With Arrhythmogenic Right Ventricular Dysplasia," Circulation, vol. 91, no. 3, pp. 715-721, 1995.

[2] M. Blanco-Velasco, B. . Weng and K. E. Barner, "ECG signal denoising and baseline wander correction based on the empirical mode decomposition," Computers in Biology and Medicine, vol. 38, no. 1, pp. 1-13, 2008.

[3] T. Jeon, B. . Kim, M. . Jeon and B.-G. . Lee, "Implementation of a portable device for real-time ECG signal analysis," Biomedical Engineering Online, vol. 13, no. 1, pp. 160-160, 2014.

[4] J. Ma, J. . Xu, H. B. Xu, Y. . Wang and S. X. Yin, "Design of ECG Signal Acquisition and Processing Circuit," Applied Mechanics and Materials, vol. , no. , pp. 856-861, 2012.

[5] S. Lee and J. . Kruse, "Biopotential Electrode Sensors in ECG/EEG/EMG Systems," , vol. , no. , p. , .

[6] V. E. Krahl, "The electric strain gauge ballistocardiograph.," American Heart Journal, vol. 39, no. 2, pp. 161-173, 1950.

[7] J. Zanetti and K. . Tavakolian, "Seismocardiography: Past, present and future," , 2013.

[8] K. Tavakolian, G. . Portacio, N. R. Tamddondoust, G. . Jahns, B. . Ngai, G. A. Dumont and A. P. Blaber, "Myocardial contractility: A seismocardiography approach," , 2012.

[9] C. A. Wick, J.-J. . Su, O. . Brand, J. H. McClellan, P. T. Bhatti and S. . Tridandapani, "A Trimodal System for the Acquisition of Synchronous Echocardiography, Electrocardiography, and Seismocardiography Data," IEEE Engineering in Medicine and Biology Magazine, vol. 2011, no. , pp. 6911-6914, 2011.

[10] S. Schmidt, A. S. Jensen, J. . Melgaard, C. . Graff, J. . Hansen, T. A. Bhuiyan and J. J. Struijk, "Three-dimensional apex-seismocardiography," Computing in Cardiology, vol. 41, no. , pp. 1105-1108, 2014.

[11] R. A. Wilson, V. S. Bamrah, J. . Lindsay, M. . Schwaiger and J. . Morganroth, "Diagnostic Accuracy of Seismocardiography Compared with Electrocardiography for the Anatomic and Physiologic Diagnosis of Coronary Artery Disease During Exercise Testing," American Journal of Cardiology, vol. 71, no. 7, pp. 536-545, 1993.

[12] C. A. Wick, O. T. Inan, J. H. McClellan and S. Tridandapani, "Seismocardiography-Based Detection of Cardiac Quiescence," IEEE Transactions on Biomedical Engineering, vol. 62, no. 8, pp. 2025-2032, 2015.

[13] I. Korzeniowska-Kubacka and R. . Piotrowicz, "Seismocardiography--a non-invasive technique for estimating left ventricular function: preliminary results.," Acta Cardiologica, vol. 57, no. 1, pp. 51-52, 2002.

[14] M. Becker, A. . Roehl, U. . Siekmann, A. . Koch, M. d. l. Fuente, R. . Roissant, K. . Radermacher, N. . Marx and M. . Hein, "Simplified detection of myocardial ischemia by seismocardiography," Herz, vol. 39, no. 5, pp. 586-592, 2013.

[15] Z. J. Salerno DM, "Seismocardiography for monitoring changes in left ventricular function during ischemia," Chest, vol. 100, no. 4, p. 991–93, 1991.

[16] J. Wahlström, I. . Skog, P. . Händel, F. . Khosrow-Khavar, K. . Tavakolian, P. K. Stein and A. . Nehorai, "A Hidden Markov Model for Seismocardiography," IEEE Transactions on Biomedical Engineering, vol. 64, no. 10, pp. 2361-2372, 2017.

[17] P. K. Jain, A. K. Tiwari and V. S. Chourasia, "Performance analysis of seismocardiography for heart sound signal recording in noisy scenarios," Journal of Medical Engineering & Technology, vol. 40, no. 3, pp. 106-118, 2016.

[18] H. Lamport, "THE ORIGIN OF THE BALLISTOCARDIOGRAPH.," Science, vol. 93, no. 2413, pp. 305-305, 1941.

[19] S. Zoneraich, "Seismocardiography is not a new term.," American Journal of Cardiology, vol. 69, no. 5, pp. 573-574, 1992.

[20] D. M. Salerno, J. M. Zanetti, L. C. Poliac, R. S. Crow, P. J. Hannan, K. . Wang, I. F. Goldenberg and R. A. V. Tassel, "Exercise seismocardiography for detection of coronary artery disease," American journal of noninvasive cardiology, vol. 6, no. 5, pp. 321-330, 1992.

[21] B. Bs, "[Seismocardiography--a new method in the study of functional conditions of the heart].," Terapevticheskii Arkhiv, vol. 33, no. , p. 55, 1961.

[22] M. Paukkunen, "Seismocardiography: Practical implementation and feasibility," , 2014.

[23] P. F. Migeotte, S. D. Ridder, X. . Neyt, N. . Pattyn, M. D. Rienzo, L. . Beck, P. . Gauger, U. . Limper, P. . Gk, V. . Rusanov, I. I. Funtova, R. M. Baevsky and J. . Tank, "Three-dimensional ballistocardiography and seismocardiography in parabolic flight: Preliminary results from the esa b3d project," , 2013.

[24] J. D. Lawson, "The Hemienvelope Phonocardiogram.," Annals of Internal Medicine, vol. 58, no. 4, pp. 710-711, 1963.

[25] C. Aravanis and R. . Harris, "The Normal Phonocardiogram of the Aged," Chest, vol. 33, no. 2, pp. 214-219, 1958.

[26] Taebi, Amirtaha, "Characterization, Classification, and Genesis of Seismocardiographic Signals" (2018). Electronic Theses and Dissertations. 5832.

A lossless speech signal compression technique

Md. Atiqur Rahman[1], Most. Jannatul Ferdous[2], Md. Mamun Hossain[3], Md Rashedul Islam[4], Mohamed Hamada [5,*]
[2,3]Department of Computer Science and Engineering
Bangladesh University of Business and Technology, Dhaka, Bangladesh
[1,4,5]School of Computer Science and Engineering
The University of Aizu, Aizu-Wakamatsu City, Fukushima, Japan
[1]atick.rasel@gmail.com, [2]jannatul.083035@gmail.com, [3]csecu.mamun@gmail.com, [4]rashed.cse@gmail.com,
[5]mhamada2000@gmail.com

***Abstract*—Day by day, the use of the internet is increasing; as a result, the way of communication through the internet is growing dramatically. So, speech signal compression is being a vital issue in the area of signal processing. There are many techniques used to compress a speech; such as Huffman coding, LZW coding, Arithmetic coding etc. These techniques convert a speech signal in such a way so that it takes very less memory and can be sent quickly over the internet than original data. From the point of view, this article introduces a novel speech coding method which use a set of binary combinations as needed to encode speech data. In this technique, the binary combinations are set to the probabilities of a speech signal where 0 and 1 will be assigned to the first two highest probabilities. The third, fourth, fifth, sixth, seventh and eighth probabilities are assigned by 00,01,10,11,000,001 and the process will continue until all probabilities are finished. Here, Average code length (ACL), compression ratio (CR), compressed file size and PSNR are considered to evaluate the performance of the proposed technique with the state of the art techniques. From the experimental outcomes, we have seen that the proposed method provides a better result than the state of the art procedures.**

Keywords: Speech compression, Huffman, LZW, Arithmetic, compression ratio, average bits length, speech signal to noise ratio.

I. Introduction

Speech can be expressed as the response of the vocal tract to one or more excitation signals. A statement is used to carry a message of human all over the world to each other [1,2]. A speech signal is stored in digital device or sent over the internet as a digital data. A large memory is required to store a massive amount of data in computer and a huge amount of time is needed to send the data over the internet. So, storing a speech signal in computer and transferring over the internet are two big challenges.

From the above points of view, Mahmut et al. proposed Huffman coding based speech signal compression procedure in [3]. In this algorithm, first of all, bandwidth diversity is analyzed for a quantized speech and then applied Huffman coding to compress the speech data. In the article [4,5], a data compression procedure using run-length-Huffman coding scheme is introduced. In the technique, a vector quantization is applied by a first stage lossy compression and then the lossy-compressed speech data are treated by a lossless compression at the second state. Finally, a run-length Huffman coding together with unary code is used to compress the data. Though Huffman coding based compression is good but it does not provide better results when the number of probabilities is large. Mustafa et al. proposed a new style speech data compression technique using Discrete Wavelet Transform (DWT) in [6,7] and shown that it provides better compression than the state of the art procedure. In [8,9], Sinha proposed a wavelet base speech compression procedure using median function thresholding. In the technique, a model is created using a generic compression scheme to represent a data in a compressed bitstream. It shows that 39.54% more compression is done than the state of the art technique and also demonstrates that it keeps 99.8184% signal energy. The article [10] discusses Discrete Cosine Transform (DCT), Wavelet Transform (WT) and Wavelet Packet Transform (WPT) based data compression and shows the comparison among them. Finally, it states that Wavelet Transform and Wavelet Packet Transform provides a better performance than Discrete Cosine Transform. Kinsner et al. propose a wavelet-based image compression procedure in [11,12]. Instead of the usual short-window ways, a wavelet-based time-frequency multi-resolution investigation is considered for compression. It shows that the wavelet-based technique can produce till 30 dB peak signal to noise ratio(PSNR) when 90% wavelet coefficients are truncated.

Hosny et al. proposes a speech data compression procedure using Wavelet Packet Decompositions (WPD) methods in this paper[13]. They take a speech data of Malayalam language for an experiment and used MSE, PSNR, NRMSE and MOS to calculate the performance of the proposed procedure. The author shows that the proposed procedure provides better quality than the state of the art techniques. In this technique, first of all, under a defined threshold level, all the high-frequency coefficients of the wavelet decomposition are killed using zero wavelet transform (ZWT). Secondly, the average zero wavelet transform (AZWT) is used to calculate the average of the estimated coefficients to achieve the goal of the first point. Finally, Lloyd's algorithm is used to quantize all the coefficients before encoding. The procedure provides

a better result in case of compression ratio than traditional discrete wavelet transform (DWT).

Speech coding techniques are two types; one is Lossy, and another one is Lossless. Kabir et al. discussed in the paper [14] that lossless procedure keeps the same copy of original signal without holding the entire bits of the message. As a result, the file encoded by lossless achieves smaller size than the original. Normally, the result of lossless compression is larger than lossy compression. Lossy audio compression method such as MPEG or MP3 may not be admissible for some use such as a stereo CD quality digital audio signal. Since, internet resources are limited, and everybody wants to transfer data quickly through it. Music distribution applications will offer highly compressed audio clips to the user for transferring.

Several techniques are introduced for digital audio signal compression in [15,16]. A popular one is based on sub-band coding . In sub-band coding, digital signals are partitioned into multiple sub-bands by multi-rate [U+FB01]lter-banks and then quantized according to the power of each sub-band signal. Another technique named Filter bank is used for audio coding using an iterative bit allocation. It was not useful to real-time communication directly because of the iterative bit allocation manner [17,18]. Because, the frequency energy distribution of audio signals is Non-uniform, so it is difficult to determine how to allocate the whole number of bits to each sub-band while reducing of quantization noise arising. Moving Picture Experts Group 1(MPEG-1) is a sub-band coding scheme, broadly used over the Internet or with digital devices [19-21]. There are some limitations of MPEG-1 given below though it provides an e[U+FB03]cient compression way [22].

- Bits are not allotted often for high-frequency sub-bands at a low bit-rate transmission.
- The quantizer of MPEG-1 Audio uses Fast Fourier transform of original signals to define how much bits be allotted. This process is not only computationally expensive but also not very suitable for real-time transmission.

Abdul Mawla et al. explained that in case of speech coding, Waveform is the most straightforward technique which reproduces a signal analysing the code sample-wise. Here the actual pattern of the speech signal is reconstructed without considering the characteristics of speech using the coding. A uniform 16-bits PCM and ADPCm are the well-known waveform can be investigated in both time and frequency domain[23]. In this paper, it has also been shown that if arithmetic coding is used then in case of decoding, the full code-word must be taken and an entire file could be corrupted for a single bit loss. Stearns et al. proposes an arithmetic coding based waveform compression in this paper[24]. In this technique, a loss linear prediction is applied to a digitized waveform to drop the compatible components. Finally, Arithmetic coding is applied to compress the data. A lossless coding has been proposed in the paper [25]. In this technique, a set of binary consecutive patters has been used to encode image data.

An innovative coding style based on binary sequence has been introduced for a lossless speech compression in this paper. In this technique, a set of binary combination are taken which are assigned to the probabilities of an original speech signal where the binary combination of the shortest length is assigned to the highest probability.The rest of this article has been made as follows. An existing algorithms have been discussed in section II. In section III, the proposed system has been introduced. Finally, The experimental results analysis and conclusion are shown in in section IV and V respectively.

II. Background study

There are many speech signal compression algorithm; Such as, Run-length coding, Huffman coding,LZW coding, Arithmetic coding etc. Huffman is a variable-length coding procedure used to compress data. A set of distinct symbols and their corresponding probabilities are used to encode a file which uses the tiniest code for the symbol containing the highest probabilities that provides a better reduction of the original signal.

In Huffman encoding scheme, all probabilities are sorted in descending order, and a tree is constructed using the possibilities. The two symbols with lowest probabilities from the list of probabilities are added, and the computed value is used as the root of the two's at each step. This process is continued until finished. When a tree is made, then it is traversed to encode the file. Huffman algorithm builds a binary coding tree in a greedy style, beginning with the leaves and again and again merges the two nodes with the least probabilities. A priority queue is used as the principal data structure to reserve the nodes. The pseudo code of Huffman Binary Coding looks below.

Input: Array P [1...n] of statistical probabilities. Output: Huffman coding tree with n leaves that has the shortest code length for P.

HuffmanCoding (P [1...n])

a) HT = Empty binary tree
b) Q = Priority queue of pairs (i, P[i]), i = 1...n, with P as the identification sign
c) foreach m = 1...n − 1
   i) a = extractMinP(Q)
   ii) b = extractMinP(Q)
   iii) P[n + m] = P[a] + P[b]
   iv) insertNode(HT, n + m) with children a, b
   v) insertRear(Q,(n + m, P[n + m]))
d) return HT

As an example, Huffman coding is demonstrated for eight symbols (S1, S2, S3, S4, S5, S6, S7 AND S8) and their probabilities (0.3, 0.2, 0.15, 0.15, 0.1, 0.07, 0.02 and 0.01) in Fig. 1. Average code length(ACL), Compression ratio(CR) for the same probabilities are given in the following TABLE I.

From the Table I, it is looking that Huffman coding uses 2.63 bits as an average code length which means it reduces 67.13%.

Shannon–Fano coding is slightly related to the Huffman coding. When the probabilities of data are negative powers of 2, then Huffman coding and Shannon-Fano produces the

Figure 1. Huffman encoding fashion

Table I
HUFFMAN CODING

S	P(i)	B2S	CLB2S(i)	P(i)*LBS(i)	CR
S1	0.3	00	2	0.6	6.0837
S2	0.2	01	2	0.4	
S3	0.15	010	3	0.45	
S4	0.15	110	3	0.45	
S5	0.1	011	3	0.3	
S6	0.07	0111	4	0.28	
S7	0.02	01111	5	0.1	
S8	0.01	11111	5	0.05	
				ACL=2.63	

best variable length codes. The main distinction between two procedures is that Huffman coding constructs a code tree from the bottom up while Shannon-Fano constructs its codes from top to bottom. Shannon determined that the concept of entropy from statistical mechanics play an essential role. Entropy estimates the degree of disorder in a mechanical system, which is related to the number of possible states of the system. There are some limitations of Shannon-Fano coding, one of them is, it does not guarantee optimal codes and another one is , two different codes for the same symbol can be generated. From the point of view, Huffman coding is better than Shannon-Fano coding.

III. RECOMMENDED SCHEME

The main interest of this article is speech data compression. A distinct method has been introduced for speech signal compression. In this technique, 0 and 1 are assigned to the first two highest probabilities. Two bits combinations (00, 01, 10, and 11) are assigned for the third, fourth, fifth and sixth greatest possibilities respectively. In the same manner, for next eight probabilities, three bits combinations (000, 001, 010, 011, 100, 101, 110, and 111) are assigned sequentially. The similar process is applied for next probabilities to encode data. The encoding and decoding procedure of the proposed coding are presented in Fig. 2.

Figure 2. Diagram of the Proposed encoding and decoding procedure

The encoding and decoding procedure of the proposed coding style are shown below: Finally send the encoded speech

Algorithm 1 Encoding Procedure

1) Read a speech signal and set i = 1
2) Calculates the probabilities (P) of the speech signal and sort them in descending order.
3) Calculate N = number of probabilities
4) Repeat step 5 to 6 until i $\leq$ N
5) Generate a binary combination X_i
6) Assign $P_i = X_i$
7) Scan the whole speech data and assign the binary combination.

data and the list of symbols where Symbol arrangements are arranged as the arrangement of the probabilities. The proposed

Algorithm 2 Decoding Procedure

1) Read the distinct symbols as the order were sent and assign 0,1, 00,01,10,11,000,001,010 code-word etc respectively.
2) Scan the whole encoded speech data and replaces each encoded code-word by the symbol where the encoded code-word matches with the code-word assigned to the symbol in the step 1.

coding scheme is given in Table-II using the same probabilities used in TABLE I.

Table II
PROPOSED CODING

Symbols	Probabilities	Proposed coding
S1	0.3	0
S2	0.2	1
S3	0.15	00
S4	0.15	01
S5	0.1	10
S6	0.07	11
S7	0.02	000
S8	0.01	001

Average code length(ACL) and compression ratio(CR) are demonstrated in Table III Using the same probabilities where CLB2S, and B2S indicate the code length of Proposed coding and Proposed coding respectively.

Table III
PROPOSED CODING

Symbols	P(i)	B2S	CLB2S(i)	P(i)*CLB2S(i)	CR
S1	0.3	0	1	0.3	10.4575
S2	0.2	1	1	0.2	
S3	0.15	00	2	0.3	
S4	0.15	01	2	0.3	
S5	0.1	10	2	0.2	
S6	0.07	11	2	0.14	
S7	0.02	000	3	0.06	
S8	0.01	001	3	0.03	
				ACL=1.53	

From the TABLE I and II, it looks that the Huffman coding uses 2.63 bits as an average code length and provides 6.0837 compression ratio where the proposed coding uses 1.53 bits as an average code length and provides 10.4575 compression ratio. It has been calculated that the proposed coding uses 41.8251% less memory than Huffman coding and provides 71.8938% more compression.

IV. RESULT ANALYSIS

Several human speech signals used as an input sample is exhibited in the following Fig. 3 where each symbols of the speech signals used in experiment hold 8, 16 bits space in memory. All speech signals are mono, sampled at 8 kHz. In this paper, compressed signal size, compression ratio (CR) and average code length (ACL) are used to analyze the results. Low average code length indicates the use of less memory whereas higher compression ratio (CR) is used to quantify the higher reduction. If a speech data contains n distinct symbols then average code length, compression ratio and peak signal to noise ratio (PSNR) are measured using the equation 1,2 and 3 respectively where P(i) is the probability of ith symbol.

$$ACL = \sum_{i=1}^{n} P(i)L(i) \tag{1}$$

$$CR = \frac{Original\ Speech\ Data\ size}{Compressed\ Speech\ data\ size} \tag{2}$$

$$PSNR = 10xLog_{10}\frac{Max\ data\ of\ original\ signal}{\sqrt{\sum_{i=1}^{s}\frac{(OS-CS)^2}{M}}} \tag{3}$$

Where OS, CS and L(i) represent original signal and compressed signal and the code length of ith probability and s is the total number of sample of the M sized speech signal. Based on the original signal size, compressed signal size, PSNR, ACL and CR, the comparison between the proposed coding and Huffman coding are given in TABLE IV. In TABLE IV, NC, HC and PC stand for no coding, Huffman coding and the proposed coding respectively.

Figure 3. Original Speech Signals

Table IV
THE PROPOSED AND HUFFMAN CODING COMPARISON

Sam	CM	OSS (KB)	CSS (KB)	PSNR	ACL	CR
1	NC	4.1953	4.1953	inf	16	1
	HC	4.1953	2.4989	inf	9.5303	1.6789
	PC	4.1953	1.3213	inf	5.0391	3.1752
2	NC	13.2735	13.2734	inf	16	1
	HC	13.2734	10.0719	inf	12.1408	1.3179
	PC	13.2734	6.0872	inf	7.3376	2.1806

We have applied Huffman coding, LZW coding, arithmetic coding and the proposed coding to seven different speech signals already shown in Fig. 3. After applying the state of the art techniques, we have got the results demonstrated in TABLE V.

From the TABLE V, for the first three speech samples, we look that Huffman coding uses 5.671, 11.725 and 9.530 bits; LZW coding uses 10.718, 12.618 and 1.051 and Arithmetic coding (AC) use 4.1374, 9.4335 and 8.4360 bits as average code length (ACL) whereas the proposed coding use 3.5, 5.5 and 5.039 bits for the same case. From this point of view, for the first three speech samples, we see that Huffman coding uses 38.282%, 53.092% and 47.125% ; LZW coding uses 67.345%, 56.411% and 54.402% and Arithmetic coding uses 15.398%, 41.7% and 40.268% more memory than the proposed procedure.

For the same three samples Huffman coding provides 2.822 1.365 and 1.679; LZW coding provides 1.529 1.106 and 1.448; Arithmetic coding estimates 3.867 1.696 and 1.897; and the

Table V
EXPERIMENTAL RESULTS

Samples Size, Bit Depth & Samples	Coding Method	Compressed Signal Size (KB)	ACL (bits)	CR
3.20 KB, 16 & 1600	Huffman	1.108	5.671	2.822
	LZW	2.093	10.718	1.529
	Arithmetic	0.808	4.137	3.867
	Proposed	0.684	3.5	4.571
15.625 KB, 16 & 8000	Huffman	11.450	11.725	1.365
	LZW	12.322	12.618	1.106
	Arithmetic	9.212	9.434	1.696
	Proposed	5.371	5.5	2.909
4.195 KB, 16 & 2148	Huffman	2.499	9.530	1.679
	LZW	2.898	11.051	1.448
	Arithmetic	2.212	8.436	1.897
	Proposed	1.321	5.039	3.175
13.273 KB 16 & 6796	Huffman	10.072	12.141	1.318
	LZW	11.109	13.391	1.195
	AC	9.489	11.436	1.399
	Proposed	6.087	7.338	2.181
20.4082 KB 16 & 10449	Huffman	14.421	11.306	1.415
	LZW	14.520	11.384	1.406
	Arithmetic	11.603	9.096	1.759
	Proposed	10.853	8.509	1.881
25.3066 KB, 16 & 12957	Huffman	19.931	12.602	1.270
	LZW	20.967	13.257	1.207
	Arithmetic	18.774	11.870	1.348
	Proposed	15.989	10.109	1.583
5.3066 KB, 8 & 5434	Huffman	3.818	5.756	1.390
	LZW	4.121	6.213	1.288
	Arithmetic	3.072	4.631	1.727
	Proposed	2.131	3.212	2.490

proposed coding provides 4.571 2.909 and 3.175 compression ratio.

Analyzing these compression ratio, we see that the proposed coding provides better results in case of both average code length and compression ratio.

After applying the proposed coding on the seven speech samples of different sizes, the signals have been reconstructed perfectly that has been shown in Fig. 4. We have also calculated peak signal to noise ratio (PSNR) of each signal and seen that it is inf for each signal. So, from the PSNR point of view, it can be stated that there is no any difference between the original signals and compressed signals.

A graphical representation of average code length of the seven speech signals have been demonstrated in Fig. 5 In this figure, four different bars have been shown for each speech sample where the bar to the left and right side represent Huffman coding and the proposed coding, and the middle-left and middle-right bar represent LZW and arithmetic coding. It is observing from the Fig. 5 that the bar for the proposed coding is smaller than that of Huffman, LZW and arithmetic coding which indicates that the suggested procedure use a few number of bits averagely than Huffman, LZW and Arithmetic coding.

On the other hand, the compression ratio of the same signals have been demonstrated in Fig. 6 and it looks that the proposed

Figure 4. Compressed speech signals

Figure 5. Comparison of average code length

coding provides a better compression than the state of the art techniques.The proposed coding reduces more than the state of the art procedures retaining the same quality. From the points mentioned above, it can be concluded that the proposed coding is one of the best lossless speech signal compression technique.

V. CONCLUSION

Speech compression is an important issue in the area of signal processing used to store in computer or send over the internet at high speed. The speech compression is obtained by representing each sample of digitized data by lesser number

Figure 6. Comparison of compression ratio

of bits. A speech signal compression system based on binary combination, a completely new coding style has been proposed in this article which work better than the state of the art techniques. The suggested strategy is illustrated to be more useful in speech signal compression than other techniques, such as Shannon-Fano Coding, Huffman Coding, LZW coding, Arithmetic coding etc. Moreover, the suggested algorithm needs less memory than others compression algorithm. This technique can be used to compress any data.

REFERENCES

[1] Tong Lai Yu, "Speech data compression for embedded systems," *Proceedings of Data Compression Conference - DCC '96, Snowbird, UT, USA, 1996, pp. 460-. doi: 10.1109/DCC.1996.488392*,

[2] Elaydi, H., Jaber, M.I. and Tanboura, M.B., 2003. Speech compression using wavelets. *In The International Arab Conference on Information Technology (pp. 1-8).*,

[3] M. E. Demirhan and U. Arıöz, "Huffman coding of a PCM-quantized speech signal," *2015 23nd Signal Processing and Communications Applications Conference (SIU), Malatya, 2015, pp. 160-163. doi: 10.1109/SIU.2015.7130379*,

[4] M. A. Rahman, M. M. Fazle Rabbi, M. M. Rahman, M. M. Islam and M. R. Islam, "Histogram modification based lossy image compression scheme using Huffman coding," *2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT), Dhaka, Bangladesh, 2018, pp. 279-284. doi: 10.1109/CEEICT.2018.8628092*,

[5] M. A. Rahman, S. M. S. Islam, J. Shin and M. R. Islam, "Histogram Alternation Based Digital Image Compression using Base-2 Coding," *2018 Digital Image Computing: Techniques and Applications (DICTA), Canberra, Australia, 2018, pp. 1-8. doi: 10.1109/DICTA.2018.8615830*,

[6] A. K. Sinha, S. K. Dutta, B. Dev, Jaydipta, R. Ranjan and R. Kumari, "Wavelet based Speech Coding technique using median function thresholding," *2014 International Conference on Electronics and Communication Systems (ICECS), Coimbatore, 2014, pp. 1-4. doi: 10.1109/ECS.2014.6892538*,

[7] Rahman, M.A. and Rabbi, M.F., 2015. DWT-SVD based New Watermarking Idea in RGB Color Space. *International Journal of Signal Processing, Image Processing and Pattern Recognition, 8(6), pp.193-198.*

[8] Wonyong Chong and Jongsoo Kim, "Speech and image compressions by DCT, wavelet, and wavelet packet," *Proceedings of ICICS, 1997 International Conference on Information, Communications and Signal Processing. Theme: Trends in Information Systems Engineering and Wireless Multimedia Communications (Cat., Singapore, 1997, pp. 1353-1357 vol.3. doi: 10.1109/ICICS.1997.652210*,

[9] Rahman, M.A. and Rabbi, M.F., 2015. Non-Blind DWT-SVD based Watermarking Technique for RGB Image. *Global Journal of Research In Engineering.Vol 15, No 4-F (2015)*

[10] W. Kinsner and A. Langi, "Speech and image signal compression with wavelets," *IEEE WESCANEX 93 Communications, Computers and Power in the Modern Environment - Conference Proceedings, Saskatoon, Saskatchewan, Canada, 1993, pp. 368-375. doi: 10.1109/WESCAN.1993.270520*,

[11] S. M. Joseph and P. B. Anto, "Speech compression using wavelet transform," *2011 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, Tamil Nadu, 2011, pp. 754-758. doi: 10.1109/ICRTIT.2011.5972258*,

[12] Rahman, M.A., Rabbi, M.F. and Islam, M.A., Dynamic Hand Gesture recognition from a complex background and finger identification using RGB colours.*International Journal of Computer Science & Engineering Technology (IJCSET), ISSN : 2229-3345 Vol. 9 No. 03 May-Jun 2018*,

[13] N. M. Hosny, S. H. El-Ramly and M. H. El-Said, "Novel techniques for speech compression using wavelet transform," *ICM'99. Proceedings. Eleventh International Conference on Microelectronics (IEEE Cat. No.99EX388), Kuwait, 2000, pp. 225-229. doi: 10.1109/ICM.2000.884846*,

[14] Hussain Mohammed Dipu Kabir, Syed Bahauddin Alam, M. A. Matin and Imranul Kabir Chowdhury, "A loss-less compression technique for high quality speech signals and its implementation with MPEG-4 ALS for better compression," *2010 IEEE International Conference on Information Theory and Information Security, Beijing, 2010, pp. 781-785. doi: 10.1109/ICITIS.2010.5689685*,

[15] H. Saito and S. Nakamura, "Speech signal compression using subband division and sub-adaptive polynomial mapping quantization," *Proceedings of Digital Processing Applications (TENCON '96), Perth, WA, Australia, 1996, pp. 184-188 vol.1. doi: 10.1109/TENCON.1996.608781*,

[16] Sun, L., Mkwawa, I.H., Jammeh, E. and Ifeachor, E., 2013. Speech compression. *In Guide to Voice and Video over IP (pp. 17-51). Springer, London.*,

[17] Gopi, E.S., 2014. Digital speech processing using Matlab (pp. 73-93). *Springer India.*,

[18] Gibson, J.D., Berger, T., Lookabaugh, T., Baker, R. and Lindbergh, D., 1998. Digital compression for multimedia: principles and standards. *Morgan Kaufmann.*

[19] Chan, W.Y. and Falk, T.H., 2016. -Machine Assessment of Speech Communication Quality. *In Mobile Communications Handbook (pp. 603-618). CRC Press.* ,

[20] M. A. Rahman, S. P. Paul, M. Das, M. M. Hossain, R. Haque and M. A. Rahman, "Convolutional Neural Networks based multi-object recognition from a RGB image," *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox'sBazar, Bangladesh, 2019, pp. 1-6. doi: 10.1109/ECACE.2019.8679409*,

[21] Mitchell, J.L., 2004. Introduction to digital audio coding and standards. *Journal of Electronic Imaging, 13(2), p.399.*,

[22] T. Painter and A. Spanias, "A review of algorithms for perceptual coding of digital audio signals," *Proceedings of 13th International Conference on Digital Signal Processing, Santorini, Greece, 1997, pp. 179-208 vol.1. doi: 10.1109/ICDSP.1997.628010*,

[23] Mawla, A., Najih, M.A., Ramli, A.R. and Prakash, V., 2003, January. Speech Compression Using Discrete Wavelet Transform. *In 4th National Conference on Telecommunication Technology Proceedings (pp. 1-4).*

[24] S. D. Stearns, "Arithmetic coding in lossless waveform compression," *in IEEE Transactions on Signal Processing, vol. 43, no. 8, pp. 1874-1879, Aug. 1995. doi: 10.1109/78.403346*

[25] M. A. Rahman, J. Shin, A. K. Saha and M. Rashedul Islam, "A Novel Lossless Coding Technique for Image Compression," *2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Kitakyushu, Japan, 2018, pp. 82-86. doi: 10.1109/ICIEV.2018.8641065*

Effects of Different Environmental noises and Sampling Frequencies on the performance of MFCC and PLP based Bangla Isolated Word Recognition System

Oli Lowna Baroi, Azhar Niaz, Md. Jahidul Islam, Md. Shaikh Abrar Kabir, Emrul Islam, Md. Jakaria Rahimi
Dept. Of EEE
Ahsanullah University of Science & Technology
Dhaka, Bangladesh
olilowna@gmail.com, azharniazkakon@gmail.com, jahidshawon10@gmail.com, shaikhak707@gmail.com, emrul.aust@gmail.com, mjrahimi@gmail.com

***Abstract*—The paper has observed that different environmental noises and sampling frequencies severely affect the performance of the MFCC and PLP based Bangla Isolated Word Recognition System. We have observed the effect of different types of environments on MFCC and PLP for 39 and 52 coefficients for 8 KHz, 16 KHz, 32 KHz, and 44.1 kHz sampling frequencies. From the experimental results, we have observed that for different sampling frequency both in noiseless and noisy medium PLP models detects better than MFCC except in CLASSROOM environment where different types of noise are present simultaneously.**

***Keywords*— MFCC, PLP, Clean and Noisy Environment, Different Sampling frequency, HMM, Bangla ASR.**

I. INTRODUCTION

Automatic speech recognition (ASR) technique helps a computer to identify the words spoken by a person into a microphone or telephone and converts it into a written text. For ASR research, a computer can perceive words with highest accuracy level which are intelligently spoken by any person and also independent of vocabulary size, noise, speaker characteristics, language or accent. Therefore, different types of feature extraction techniques have already been introduced. Among them, MFCC and PLP are most admired feature extraction techniques. That's why we have taken an effort to comprehend which feature helps us to get better detection rate for both noisy and noiseless medium at different sampling frequency for Bangla Corpus.

One of the most popular feature extraction techniques for speech analysis is MFCCs (Mel Frequency Cepstral Coefficients). This technique facilitates auditory processing and relatively fast to compute. It works based on the critical bandwidth frequency of human ears with filter spaced linearly at low frequency and logarithmic at high frequency. That's why it approximates the human system response more closely than any other system. But this system is not robust in the presence of additive noise and therefore it normalizes their values at speech recognition to minimize the influence of noise.

For this reason, another new technique based on the short-term spectrum for the analysis of speech is developed termed PLP which is the modified version of LPC. PLP use Bark scale instead of the Mel scale. It uses an all-pole model to smooth the modified power spectrum and then estimate the output cepstral coefficients. That's why in some particular conditions PLP outperforms MFCC.

Many researchers showed a great effort to find out the effects of different parameters [1] [4] on MFCC and PLP. Comparison among different feature extraction techniques [6] [9] [10] were also being observed using different language corpus. Furthermore, some documents exhibited us how different types of feature extraction techniques work [2] [3] [5] [7].

Therefore in this paper, we desired to observe the effects of different types of environment and sampling frequency for Bangla corpus language using delta, acceleration and third differential coefficients. Our experimental corpora have been made using isolated sentences. So, we used Hidden Markov Model (HMM) based on monophone model for our experiment as it is better for isolated voice recognition. HMM models have been built up using HTK toolkit and MATLAB software helped us to do this operation.

We have divided our paper into six parts including this introduction section. Part-II delineates the difference between PLP and MFCC. Part-III & Part-IV represents the experimental setup and the changes of configuration files that should have been done before creating a model. In Part-V we have tried to analyze the Result. And lastly, conclusion and future work are discussed in Part-VI.

II. PLP AND MFCC

MFCC is a robust, dynamic and most commonly used Feature Extraction Technique for speech which uses Mel-spaced filter banks. Therefore behave more likely to a human ear. The Mel filter bank is a series of a triangular bandpass filter. This scale behaves approximately linear below 1000 Hz and logarithmic for frequencies above 1000 Hz. The following formula is used to compute the Mel for a given frequency [3]

$$Mel(f) = 2595 \log_{10}(1 + \frac{f}{700})$$

The triangular filters are arranged in such a manner that the lower value of one filter is placed in the middle of the

previous filter and the higher value is placed within the center of the next filter.

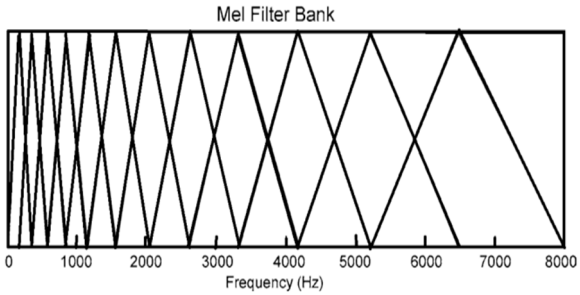


Fig. 1. Mel scale filter Bank

On the contrary, PLP is a new feature extraction technique which works on the psychophysics of human hearing. It is similar to linear predictive coding (LPC), except that the PLP technique also uses three concepts which are the critical-band spectral resolution, equal-loudness curve, and intensity- loudness power law [2].

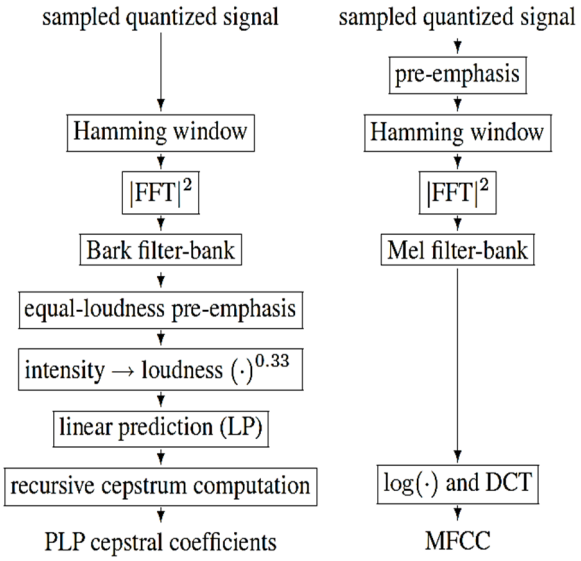


Fig. 2. The computation steps of PLP (left) and MFCC (right).

PLP uses Bark spaced filter bank instead of Mel scaled filter Bank where 1 Bark is approximately equal to 100 Mels. The Bark filter bank is a series of a trapezoidal shaped bandpass filter. This scale becomes more and more linear below 500 Hz and more or equal logarithmic above 500 Hz. The Bark for a particular frequency can be estimated using the equation given below,

$$Bark(f) = \frac{26.81f}{1660 + f} - 0.53$$

If result< 2, add 0.15*(2-result)

If result > 20.1 add 0.22*(result-20.1)

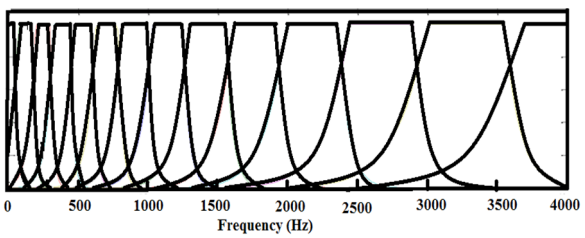


Fig. 3. Bark Filter Bank

PLP is a very low order. As it needs fewer coefficients, they are more efficient in terms of both storage and recognition complexity than most other spectral representation. They could show less sensitivity to speaker variation than most other spectral representation. Whereas MFCC is a fast and simplified model. But it's less sensitive in additional noise.

III. EXPERIMENTAL SETUP

For our experiment, we have taken isolated as well as some naturally spoken voices at 8KHz, 16KHz, 32KHz and 44.1KHz sampling frequency in both noiseless and noisy medium and then made some model using the monophone model which is actually fast and better for isolated voices. We have recorded voices in four different medium. They are (1) noiseless environment, (2) only AC noise present, (3) only FAN noise present and (4) CLASSROOM environment. We used both MFCC and PLP configurations for our experiments and tried to observe the effects of delta, acceleration, third differential and cepstral coefficients after appending them into the configuration for different situations. We have made four models for our experiments.

TABLE I. Experimental Models

Models	No. of voices for train corpus	Description
Model-1	800	In a clear medium, sentences are uttered by a number of male and female speakers for building the train corpus in different frequencies (8 KHz, 16 KHz, 32 KHz, and 44.1 KHz) and then use the base voices for the test.
Model-2	800	Sentences are uttered by a number of male and female speakers for building the train corpus in such a medium where the noise of **AC** is interrupting the sentences and then use the same base voices for the test.
Model-3	800	Sentences are uttered by a number of male and female speakers for building the train corpus in such a medium where the noise of **FAN** is interrupting the sentences and then use the same base voices for the test.
Model-4	800	Sentences are uttered by a number of male and female speakers for building the train corpus in a **CLASSROOM** where sentences are interrupted by several things and then use the same base voices for the test.

IV. CHANGE IN CONFIGURATION FILES

The main configuration files for MFCC and PLP are given in TABLE II.

TABLE II. Configuration Files

PLP	MFCC
SOURCEKIND = HAUDIO SOURCEFORMAT = HTK SOURCERATE = 226 TARGETKIND = PLP TARGETRATE = 100000.0 WINDOWSIZE = 250000.0 USEHAMMING = T PREEMCOEF = 0.97 NUMCHANS = 20 CEPLIFTER = 22 NUMCEPS = 12 USEPOWER = T LPCORDER = 12	SOURCEKIND=WAVEFORM SOURCEFORMAT = WAV SOURCERATE = 226 TARGETKIND = MFCC TARGETRATE = 100000.0 SAVECOMPRESSED = T SAVEWITHCRC = T WINDOWSIZE = 250000.0 USEHAMMING = T PREEMCOEF = 0.97 NUMCHANS = 26 CEPLIFTER = 22 NUMCEPS = 12 ENORMALISE = T

The qualifiers must append in the TARGETKIND for changing the model and number of coefficients. We have used the following TARGETKIND for our paper:

i. MFCC_D_A_0(39 coefficients) monophone HMM
ii. MFCC_D_A_0_T (52 coefficients) monophone HMM
iii. PLP_D_A_0(39 coefficients) monophone HMM
iv. PLP_D_A_0_T (52 coefficients) monophone HMM

SOURCERATE, TARGETRATE, and WINDOWSIZE require a floating point value in 100 ns units. In TABLE II. SOURCERATE = 226 indicates that the source is direct audio with sample period 22.6μs. Frequency is inversely proportional to period. So, frequency, in this case, is $\frac{1}{22.6\times10^{-6}}$ = 44.1 KHz.

TABLE III. SOURCERATE at each frequency

FREQUENCY (KHz)	SOURCECREATE
8	1250
16	625
32	312
44.1	226

V. RESULT ANALYSIS

Our database consists of 800 speech signals acquired during different sessions at the different sampling frequency. TABLE IV.toTABLE VII. shows the sentence detection rate (SDR) in percentage for Model-1. Here, 800 sentences are spoken by five male and three female in a noiseless medium which is used for making the train corpus. After that 100 voices have taken for the test purpose from the base voices.

TABLE IV. Model-1 at 8 KHz noiseless medium

SL. No	MFCC_D_A_0	MFCC_D_A_0_T	PLP_D_A_0	PLP_D_A_0_T
M1	84	89	92	90
M2	96	97	98	99
M3	95	93	94	91
M4	90	86	94	92
M5	81	86	91	91
F1	91	88	90	91
F2	90	89	94	96

TABLE V. Model-1 at 16 KHz noiseless medium

SL.No	MFCC_D_A_0	MFCC_D_A_0_T	PLP_D_A_0	PLP_D_A_0_T
M1	87	87	90	88
M2	96	96	99	96
M3	93	91	94	90
M4	88	89	88	89
M5	82	84	85	83
F1	96	92	98	97
F2	99	98	98	97

TABLE VI. Model-1 at 32 KHz noiseless medium

SL. No	MFCC_D_A_0	MFCC_D_A_0_T	PLP_D_A_0	PLP_D_A_0_T
M1	88	87	89	87
M2	97	98	99	98
M3	94	95	94	92
M4	90	86	93	85
M5	80	83	82	84
F1	92	91	95	94
F2	98	98	98	98

TABLE VII. Model-1 at 44.1 KHz noiseless medium

SL. NO	MFCC_D_A_0	MFCC_D_A_0_T	PLP_D_A_0	PLP_D_A_0_T
M1	74	72	79	75
M2	86	86	91	90
M3	87	82	90	89
M4	60	56	63	61
M5	45	52	56	56
F1	93	96	97	96
F2	95	94	97	98

From TABLE IV.toTABLE VII.we can visualize the detection rate is quite good for each of the models at every sampling frequency except 44.1 KHz. At 44.1 KHz the detection rate is slightly poor than other frequencies. Also, we have observed that PLP detects better than MFCC in different sampling frequency in the noiseless medium. Moreover, PLP_D_A_0 & PLP_D_A_0_T both models gives better detection for 8 KHz sampling frequency. Otherwise, PLP_D_A_0 gives the highest detection at other sampling frequencies. In Fig. 4 the detection of PLP_D_A_0 in the noiseless medium is given.

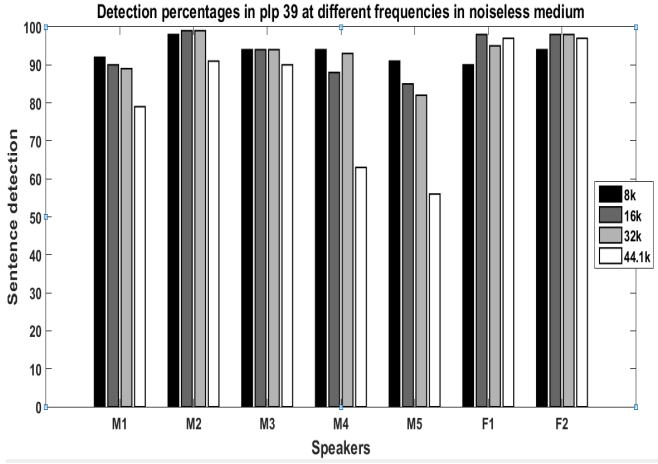


Fig. 4. PLP_D_A_0 at different frequencies in noiseless medium

Fig. 5 delineates the sentence detection rate (SDR) in percentage at 8 KHz sampling frequency in pure noiseless and AC, FAN &CLASSROOM noise based medium.

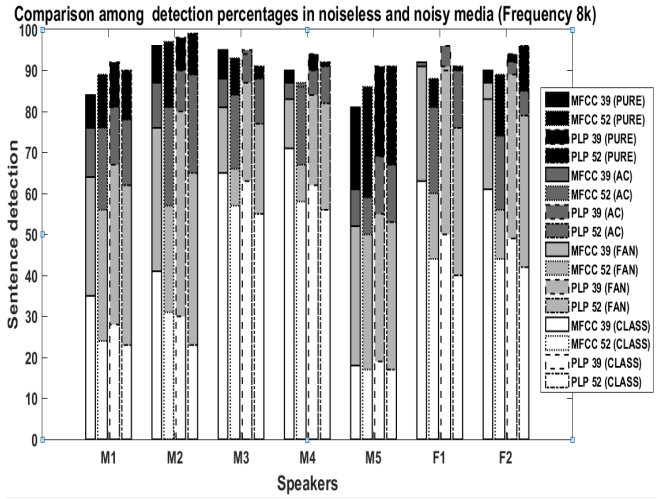


Fig. 5. Comparison between MFCC and PLP in a different medium at 8 kHz frequency

From Fig. 5 we can observe that the detection rate is highest in the noiseless medium and lowest in the CLASSROOM environment. We can also visualize that rarely AC noised medium's detection is approximately similar to the noiseless medium's detection rate.

Fig. 6 exhibits the sentence detection rate (SDR) in percentage at 16 KHz sampling frequency in pure noiseless and AC, FAN &CLASSROOM noise based medium.

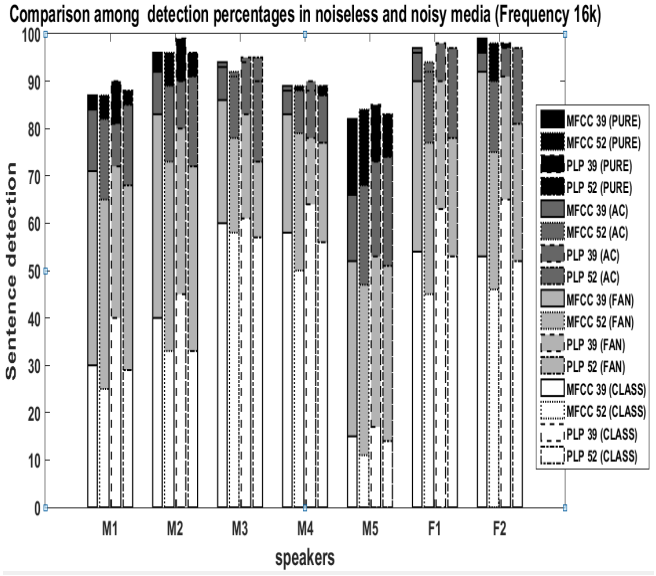


Fig. 6. Comparison between MFCC and PLP in a different medium at 16 kHz frequency

From Fig. 6 we can depict that again the detection rate is highest in the noiseless medium and lowest in the CLASSROOM environment. We can also exhibit that AC noised medium's detection is approximately similar to the noiseless medium's detection rate for a number of samples in many cases.

Fig. 7 & Fig. 8 visualize the sentence detection rate (SDR) in percentage at 32 KHz and 44.1 KHz sampling frequency in pure noiseless and AC, FAN & CLASSROOM noise based medium.

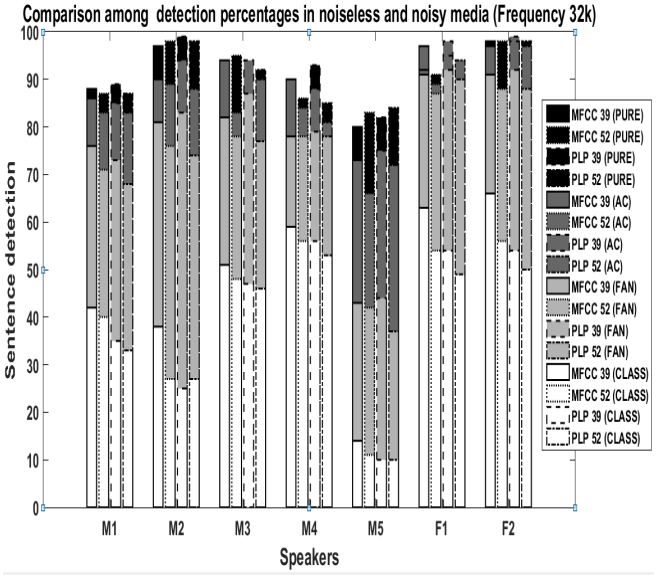


Fig. 7. Comparison between MFCC and PLP in a different medium at 32 kHz frequency

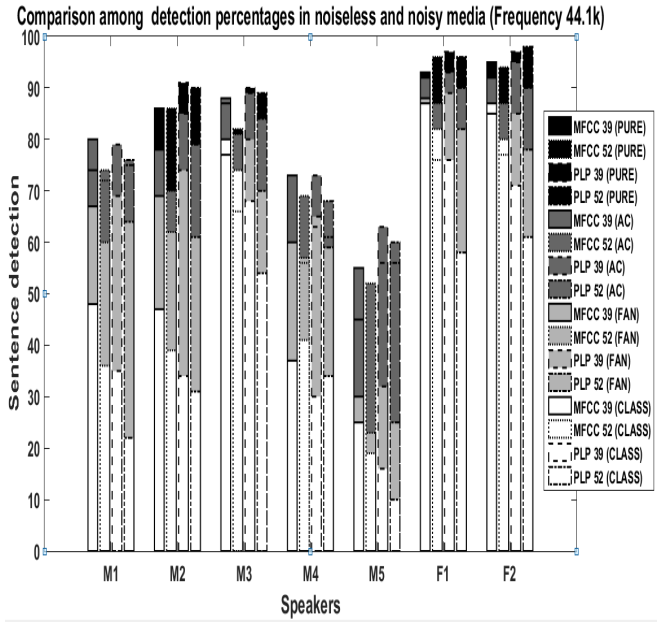


Fig. 8. Comparison between MFCC and PLP in a different medium at 44.1 KHz frequency

Fig. 7 & Fig. 8 show a similar pattern for highest and lowest detection as 8 KHz and 16 KHz. In some cases of 32 KHz, noiseless and AC noised medium detection are similar. But for 44.1 KHz frequency, the detection rate for both noiseless and AC noised medium are quite similar to each other and also less than other frequencies.

Furthermore, by observing Fig. 5 to Fig. 8 we can also identify that PLP_D_A_0 detects better than others in AC noised and FAN noised medium at different frequencies. But for CLASSROOM environment MFCC_D_A_0 outperforms all other models for each sampling frequency. In TABLE VIII.we can see the entire maximum detection rate per sample and detectors.

TABLE VIII. Maximum detection rate per sample in noisy medium at different frequencies

Medium	Frequency (KHz)	Maximum detector	M1	M2	M3	M4	M5	F1	F2
AC noise mixed	8	PLP_D_A_0	81	90	95	90	69	96	92
	16	PLP_D_A_0	81	90	95	90	73	98	97
	32	PLP_D_A_0	85	94	94	88	75	98	99
	44.1	PLP_D_A_0	79	85	89	83	70	93	95
FAN noise mixed	8	PLP_D_A_0	67	80	87	84	55	91	89
	16	PLP_D_A_0	72	80	83	78	53	90	91
	32	PLP_D_A_0	73	83	87	79	44	92	92
	44.1	PLP_D_A_0	69	74	80	67	32	89	85
CALSS ROOM	8	MFCC_D_A_0	35	41	65	71	18	63	61
	16	MFCC_D_A_0	88	86	92	88	54	97	97
	32	MFCC_D_A_0	42	38	51	59	14	63	66
	44.1	MFCC_D_A_0	48	47	80	37	25	87	87

From here we can say, though detection is poor in higher sampling frequency for noiseless or AC & Fan noised medium, it gives the highest detection in the CLASSROOM environment.

VI. CONCLUSION AND FUTURE WORKS

So, after all the experimental results we can come to this conclusion that PLP detects better than MFCC in noiseless as well as in noisy medium. But the detection rate varies depending on frequency. For higher frequency, this detection rate becomes poor however for a real-life noised medium like CLASSROOM it gives the highest detection. Though PLP outperforms MFCC in many noisy media, in a real environment like CLASSROOM, MFCC works better than PLP at any frequency. Moreover, 39 coefficient models dominate 52 coefficient models both in PLP and MFCC models. However, as we expected, in most of the cases, the detection rate of PLP is greater than MFCC. In this paper, we observed the effects of different environment and frequencies. So, in the future, we can enrich our database and can try to observe the effects of environment and frequencies using triphone model for naturally speaking a language.

REFERENCE

[1] Josef Psutka, Ludek Muller and Josef V Pustka, "Comparison of MFCC and PLP parameterizations in the speaker independent continuous speech recognition task" In: Eurospeech 2001, pp.1813-1816(2001).

[2] Florian Honig, Georg Stemmer, Christian Hacker, Fabio Brugnara, "Revising Perceptual Linear Prediction (PLP)". In: INTERSPEECH 2005.

[3] Athiramenon.G, Anjusha.V.K, "Analysis of Feature Extraction Methods for Speech Recognition", In: IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 4 Issue 4, April 2017.

[4] Debnath, S., Saha, S., Aziz, M.T., Sajol, R.H., Rahimi, M.J.: Performance comparison of MFCC based Bangla ASR system in presence and absence of third differential coefficients. In: Proceeding of the International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), pp. 1–6 (2016).

[5] H. Hermansky, "Perceptual linear predictive (PLP) analysis of speech," In: Acoustical Society of America Journal, vol. 87, no. 4, pp. 1738–1752, 1990.

[6] Sai Chen ,Hongcui Wang, JiaJia, Yeteng An and Jianwu Dang , "Comparison of Mel Frequency Ceptrum Coefficient and Perceptual Linear Predictive in Perceptual Measurement of Chinese Initials", In: Trans Tech Publications, SwitzerlandApplied Mechanics and Materials Vols. 411-414, pp 291-297 (2013).

[7] Namrata Dave, "Feature Extraction Methods LPC, PLP and MFCC In Speech Recognition, " In: International Journal for advance Research in Engineering and Technology, Volume 1, Issue VI, July 2013.

[8] FatmazohraChelali, Amar.Djeradi, Rachida.Djeradi, "Speaker Identification System based on PLP Coefficients and Artificial Neural Network", In:Proceedings of the World Congress on Engineering 2011 Vol II WCE 2011, July 6 - 8, 2011, London, U.K.

[9] H.B.Chauhan, Prof. B.A. Tanawala, "Comparative Study of MFCC and LPC Algorithms for Gujrati isolated Word Recognition," In: International Journal of Innovative Research in Computer and Communication Engineering, Vol.3, Issue 2, February 2015.

[10] Corneliu Octavian Dumitru, Inge Gavat, "A Comparative Study of Feature Extraction Methods Applied to Continuous Speech Recognition in Romanian Language," International Symphosium ELMAR, 07-09 June, 2006, Zadar, Croatia.

Design and Characterization of low loss Single Mode Fiber for Terahertz Signal Transmission

Akash Bosu
Department of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Bangladesh
akashruet13@gmail.com

Chaity Basak
Department of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Bangladesh
basak.ruet.ete@gmail.com

Shumaia Sharmin
Department of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Bangladesh
Shumaia.Sharmin.Ruet@gmail.com

Khandaker Md. Tasrif Hossain
Department of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Bangladesh
tasrif.hossain948996@gmail.com

Sharmina Zaman Urmi
Department of Electrical & Electronics Engineering
Rajshahi University of Engineering & Technology
Bangladesh
urmisharmina03669@gmail.com

Mohammad Al Mahfuz
Department of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Bangladesh
almahfuz.ete.ruet@gmail.com

***Abstract*—In this paper, Zeonex based hexagonal photonic crystal fiber (PCF) is proposed to obtain low confinement loss and moderate effective material loss (EML) at terahertz frequency. We use the finite element method (FEM) with a perfectly matched layer (PML) boundary condition to investigate the modal properties of the PCF. Simulated results demonstrate that this structure transfers signal in single mode condition. Besides, excessive low confinement loss of 6.73×10^{-11} dB/cm, and EML of 0.1491 dB/cm have obtained at an operating frequency of 0.55 THz. In this paper, other crucial guiding parameters such as air-core power fraction, effective area, bending loss, and dispersion of the fiber have also discussed. This terahertz fiber can be a good candidate for several applications in the terahertz regime.**

Keywords—Terahertz, single-mode fiber, effective material loss, confinement loss, bending loss.

I. INTRODUCTION

The wavelength of an electromagnetic wave within 3-0.003 mm (0.1-10 THz) is known as terahertz (THz) frequency band or sub-millimeter wave [1]. THz frequency band lies between the far end infrared radiation and microwave radiation. Because of its powerful applications, THz draws the attention of the researcher's eyes. THz radiation has been used in gas sensing, imaging in the medical sector, security, oral treatment, and microRNA based diagnostic [1]. THz radiation widely uses in biomedical applications like colon cancer diagnosis, ovarian cancer diagnosis, skin cancer detection, and to detect breast tumors. THz radiation has also used in the communication sector. It brings benefit in the local area networks (LAN) for the wireless communication system. Comparing with X-ray, terahertz radiation has some advantages in the case of its non-ionizing property.

Conventional THz waveguide suffers from high absorption loss and confinement loss which generally based on free-space line-of-sight (LOS) communications. Therefore, it is very challenging to design an effective waveguide that transmits the THz wave. To overcome these existing problems, the researchers have proposed different forms of waveguides. Metallic circular waveguides, slit waveguides, hollow metallic waveguides, bare metal wires are the example of metallic waveguide [1]. Metallic waveguides face trouble in the long-distance transmission due to ohmic loses, and dispersion. But in parallel-plate waveguides, there is no dispersion. However, they have diverge lose [1]. Researchers tried to demonstrate various types of polymer fibers such as Bragg fibers, dielectric plastic fibers [1], and porous core honeycomb fibers [1] to transmit THz signal effectively. In the THz regime, these Polymers fibers showed high birefringence. However, they face problems to transmit THz signal due to high absorption loss, low confinement of light, and the small possibility of fabrication. Polymers fibers have very high and positive dispersion too.

To overcome these shortcomings researchers demonstrate photonic crystal fiber (PCF). But solid core PCF has a high effective material loss (EML) in the THz frequency band. Porous-core PCFs show their uniqueness. They are capable of achieving low material loss and flexible in fabrication.

Many researchers keep their footprint to improve the transmission possibility in the THz frequency band. Reference [2] showed, polarization maintaining spiral PCF for THz transmission with the EML of 0.37 dB/cm and low confinement loss of 1.91×10^{-3} dB/cm at an operating frequency of 1 THz. This work also showed about 31% and 37% of the total power flows in the porous core for the core size of 410 μm. Atakaramians *et al.* proposed an asymmetrical sub-wavelength porous-core waveguide that exhibited an EML of >0.025 and confinement loss of 2.51 dB/cm at 0.65 THz using PMMA as the background material [3]. Islam *et al.* suggested a diamond-shaped porous-core which showed the dispersion of 2.92±0.55 ps/THz/cm and confinement loss about 0.1 dB/cm at an operating frequency of 0.70 THz [4]. They used Topas as their PCF's background material. Md. Hasan *et al.* [5] proposed a new type of THz fiber. This research work offered a cladding with square lattice and porous core which exhibited low and flattened dispersion from 0.65-1.05 THz. They showed their structure worked at single mode condition between a large frequencies ranges. Hasan *et al.* [6] demonstrated Topas based novel porous-core fiber with EML of 0.39 dB/cm and confinement loss of 1.17×10^{-3} dB/cm.

In this paper, using Zeonex as the background material, we design and characterize a hexagonal THz photonic crystal

fiber. In this paper, we use four rings of the air holes. In existing most of the work in this literature, several air rings were introduced. But using more air holes in the cladding region makes the structure large, and these decreases the fabrication possibility. This paper aim is to lower the confinement loss and increase power propagation into the core air-holes. We also analyze different significant modal parameters including dispersion, EML or material absorption loss, effective area, V-parameter, numerical aperture, FOM and bending loss.

II. Design Procedure

Fig. 1 shows the schematic structure of the proposed PCF with the expanded view of the porous core. In cladding region, four rings of the hexagonal structure are proposed which make the design compact and gives the adjustability of the fiber. Four rings are used because with the increasing of the air holes the refractive index of the cladding decreases. Many air holes in the cladding reduces the confinement loss. We have chosen hexagonal lattice in the cladding because it provides high confinement of light and easy to fabricate. In the hexagonal structure, identical air holes are used that provides fabrication flexibility. Moreover, air-filling fraction (AFF) can be better controlled which determines the loss on light confinement into the PCF.

In the cladding the hole-hole distance $\Lambda = D_{core}/0.43$. The diameter of the hexagonal cladding air hole is denoted by D and $D=0.42\times\Lambda$. For the simplicity the value of D is fixed for all the simulation. The value of D_{core} has been changed to calculate different parameters. In the core, rectangular shaped structure with eight horizontal line of air holes are chosen. In the rectangular line, air holes are missing after one another. We remove air holes because this gives better confinement of light through the core. In the upper and bellow of the rectangular core, we use two triangular shaped to maintain polarization. The diameter of core air hole is a_c. The distance from one line to another line is d_c. The value of $d_c = 0.30\times a_c$. The pitch of the core air hole is denoted by a_c, and the value of $a_c = 0.230\times D_{core}$ .The total diameter of the proposed PCF is taken $(4.1\times\Lambda + D)$. One of the important layers of PCF is perfectly matched layer (PML). PML is used to trim the computational period. In our proposed PCF the outer Layer is PML. In our design, PML is taken as 15% of the total fiber.

Polymethyle methacrylate (PMMA), Teflon, Topas, Zeonex are several polymers which were applied as a back-ground material for a THz waveguide and among of these polymers, Cyclic-Olefin copolymer (COC) commonly known as Topas used as the background material in most of the previous structures.

We consider Zeonex as the background material to design an effective THz waveguide. This material is selected due to its some advantages. It has a stable refractive index of 1.53 in THz region, low absorption of 0.14-0.30 cm^{-1} for 0.2-1.2 THz [7], insufficient water absorption of <0.1% [7], negligible moisture absorption [7]. Zeonex shows high transparency, sterilization, silicon oil-free, and drug compatibility. It also gives better stability in Heat exposures and humidity.

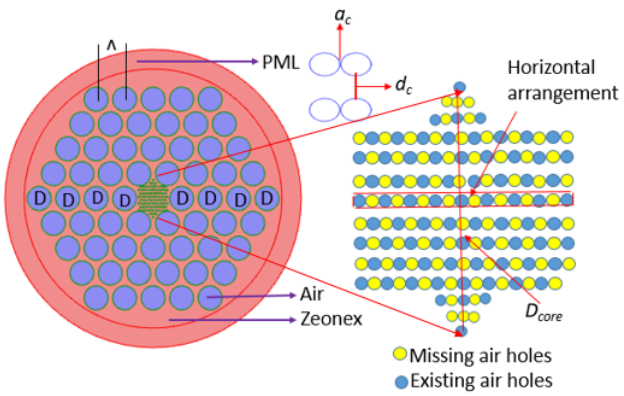


Fig. 1. Schematic sample of the proposed THz waveguide with a large view of core is shown beside.

III. Simulation Results And Discussion

To design and simulate this proposed PCF finite element method (FEM) based commercially available computer software package COMSOL v.4.2 has been used. Besides scattering boundary condition has been added to simulate and analysis the proposed structure. The complete mesh of our PCF contains 616 vertex elements, 4220 boundary elements and 0.05711 minimum element quality. To represent the complete structure 60378 elements are required.

The bulk absorption loss of Zeonex is 0.20 cm^{-1} that has been added in the simulation time. This structure has been analysed for different values of D_{core}. For D_{core}=300 µm, fundamental core mode power flow distribution of the proposed fiber is showed in Fig. 2 at 0.65 THz. This confirmed that the light is strongly confined in the centre of the proposed PCF.

Signal strength should be reduced due to modal distortion. Modal distortion is occurred by the PCF when the polarized light propagates in more than one mode. To avoid distortion, single mode fiber is desired. Using V-parameter the single mode condition of the proposed PCF can be validated. The characteristics of single mode can be measured by the following equation:

$$V = \frac{2\pi f}{c}\sqrt{n^2_{co} - n^2_{cl}} \le 2.405 \qquad (1)$$

When the fiber satisfies the above condition, a signal is transmitted by the fiber with zero interference.

In equation 1, n_{cl}= effective index of refraction for cladding, n_{co}= effective index of refraction for core, f= operating frequency. We calculate V-parameter for different operating frequency. Besides we also analysis the effect of D_{core} for a fixed frequency.

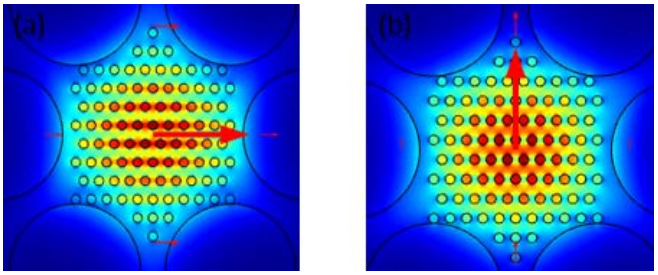


Fig. 2. Electric field distribution of the proposed fiber at 0.65 THz (a) x-polarized beam and (b) y-polarized beam.

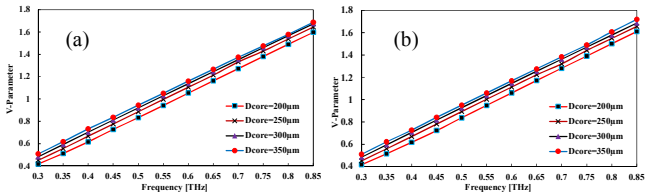


Fig. 3. V-parameter of the proposed waveguide as based on frequency for different D_{core} and n_{cl}=1 (a) x-polarized beam (b) y-polarized beam.

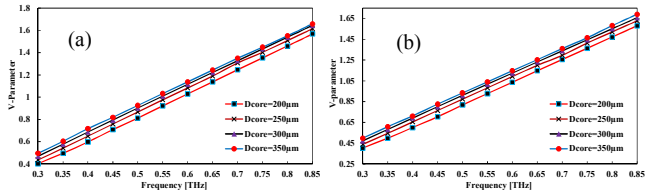


Fig. 4. V-parameter of the proposed waveguide as based on frequency for different D_{core} and n_{cl}=1.03 (a) x-polarized beam (b) y-polarized beam.

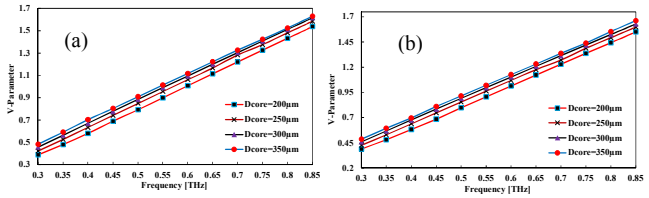


Fig. 5. V-parameter of the proposed waveguide as based on frequency for different D_{core} and n_{cl}=1.06. (a) x-polarized beam (b) y-polarized beam.

It can be reported that to design a waveguide, the cladding of the PCF primarily consists of a large sized air holes [3]. Besides the core consists of small sized air holes and bulk material respectively. Thus the effective index of refraction for the core should be found from the guided mode and the effective index of refraction for the cladding should be taken unity. From Fig. 3 it can be observed that V-parameter is an increasing function of frequency but it does not exceed the value 2.405. But the cladding region is not perfectly air in nature. Thus the effective index of refraction for cladding is unity is not true. Therefore, we calculate V-parameter for n_{cl}=1.03 and 1.06 respectively. From Fig. 3-5 it is clear that our proposed PCF propagates in single mode condition.

The background material of the fiber makes the signal weaker when light-wave travels through the PCF. Material absorption loss or EML is the main loss factor for a PCF. EML indicates that how much polarized light is absorbed by the background material. This important parameter is weighted by the following formula [2]:

$$\alpha_{eff} = \frac{(\varepsilon_0/\mu_0)^{(1/2)} \int_{A_{mat}} n\alpha_{mat} |E|^2 dA}{2\int_{ALL} S_z dA} cm^{-1} \quad (2)$$

Where, s_z=0.5(E×H*)z, s_z is the pointing vector.

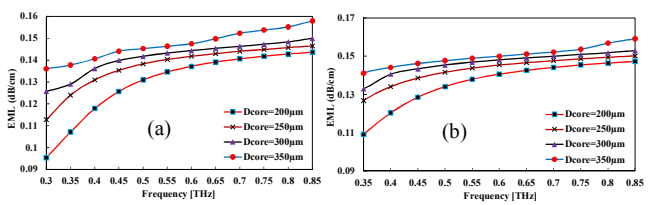


Fig. 6. EML of proposed waveguide as based on frequency for (a) x-polarized beam (b) y-polarized beam.

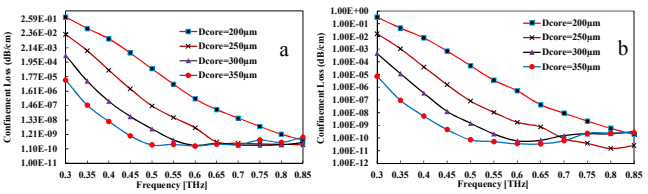


Fig. 7. Confinement loss of proposed waveguide as based on frequency for (a) x-polarized beam (b) y-polarized beam.

For a different value of D_{core}, Fig. 6 shows the EML as a function of operating frequency. From the figure, it can be seen that x-polarized mode and y- polarized mode exhibits approximate same amount of EML for a certain frequency. The calculated EML for the proposed structure are 0.1454 dB/cm and 0.1491 dB/cm for x-polarized and y-polarized modes, respectively for D_{core}=300 μm. The unit of cm^{-1} can also be expressed in terms of dB/cm by simply multiply it with 4.34. These values are found for a fixed frequency of 0.65 THz. By analyzing these value one can observe that x and y-modes absorb approximate same amount of the polarized light by the Zeonex.

Another loss factor of that proposed PCF is confinement loss. When polarized light propagates through the core some light is gripped by the nearest air holes. Confinement loss reduces the transmission possibility. The value of confinement loss should be kept low for efficient signal transmission. Due to the leaky nature of the modes confinement loss arises. It relies upon the arrangement of the air holes that used in the core and cladding. This important loss factor can be estimated by the following formula [2]:

$$\alpha_{cl} = 8.686 \times \frac{2\pi f}{c} \mathrm{Im}(n_{eff}) \times 10^{-2} \text{ dB/cm} \quad (3)$$

Where, f is the operating frequency, $Im(n_{eff})$ is the imaginary part of the effective index of refractive for fundamental core mode and c is the speed of light in vacuum space.

Fig. 7 shows the confinement loss of the proposed fiber for different D_{core} as a function of different frequency. α_{cl} is a decreasing function of D_{core} for both x-polarized and y-polarized modes. With the increase of D_{core} more light is fixedly confined in the core which gives surety that confinement loss is inversely proportional to D_{core}. From the Fig. 7 it is understood that x-polarized modes exhibits more loss than y-polarized modes. The calculated confinement loss of the proposed PCF for D_{core}=300 μm at 0.65 THz is found 2.71×10^{-10} dB/cm and 6.73×10^{-11} dB/cm for x-polarized and y-polarized modes, respectively. It can also be observed from the figure that for a fixed D_{core} confinement loss decreases with the increases of frequency. This happens because polarized light propagates through the higher refractive index region and that's why less light propagates through the cladding region at high frequency.

The power that is limited to different region of the PCF in the guided mode is called the fraction of power. Power fraction within the core in guided mode is a crucial parameter for characterizing the waveguide. But in many research paper power fraction was not calculated.

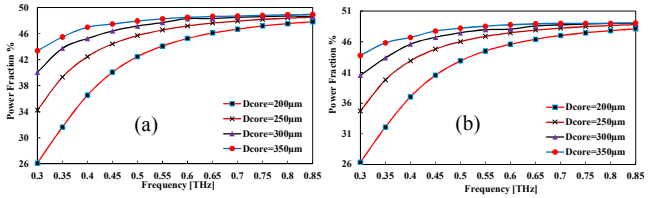


Fig. 8. Power fraction of proposed waveguide as based on frequency for (a) x-polarized beam (b) y-polarized beam.

Fraction of power in different region of the fiber can be computed by the following equation [2]:

$$Power\ fraction = \frac{\int_{X} S_Z dA}{\int_{ALL} S_Z dA} \tag{4}$$

In the numerator, the region of an interested area is denoted by X and in the denominator the total area of the fiber is denoted by ALL. For this proposed design power fraction of porous core is calculated as a function of operating frequency.

From Fig. 8 it can be observed that power fraction in the core region for the guided mood increases with the increases of the core diameter. Light is well confined in both x and y polarized beam. It is reported that the fraction of power for x-polarization and for y-polarization is 48.32% and 48.57% at D_{core}=300 μm and f=0.55 THz. It is also evident from the figure that the fraction of power through the core is same for a large bandwidth (0.60-0.85 THz).

For characterizing a PCF modal effective area is an important parameter. It indicates the actual area through which the polarized light propagates. In laser application and communication effective area plays a significant role. Higher values of effective area are required for laser application and lower values are required for communication. The area inside the core field of a fiber which is covered by the EM waves can be defined by the effective area and quantitatively measured by the following equation [6]:

$$A_{eff} = \frac{|\int I(r) r dr|^2}{|\int I^2(r) r dr|^2} \tag{5}$$

Where, I(r) =$|E_t|^2$=intensity distribution of the transverse electric field.

Modal effective area of the fundamental core mode as a function of frequency is shown in Fig. 8, for different values of D_{core}. From the Fig. 8 it can be observed that the effective area is reduced with the increase of the frequency. In this paper modal effective area is calculated for both x-polarized and y-polarized beam. From the figure it shown that the value of A_{eff} is 5.3×10^5 and 4.8×10^5 μm^2 for x and y-polarization.

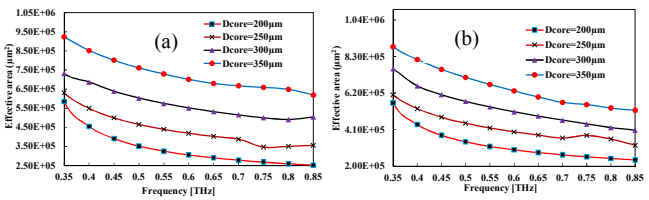


Fig. 9. Effective area of the proposed waveguide as based on frequency for (a) x-polarized beam (b) y-polarized beam.

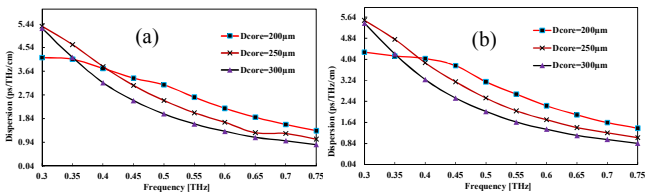


Fig. 10. Dispersion of proposed waveguide as based on frequency for (a) x-polarized beam (b) y-polarized beam.

Now we are interested to analysis another important parameter termed dispersion. It indicates that how much pulse is spreading when it propagates through the PCF. For characterizing an effective waveguide our principal target is to reduce the modal dispersion. From the waveguide structure, the modal dispersion creates. It also originates from the background material. Material dispersion and waveguide dispersion are two important dispersion factors for a PCF. But the novel material Zeonex has very low material dispersion in THz frequency band. Thus the effect of material dispersion is not considered for our calculation. Therefore, waveguide modal dispersion is observed for our proposed structure. Basically waveguide dispersion depends on the geometrical structure of the PCF and can be calculated by the following equation [2]:

$$\beta_2 = \frac{2}{c}\frac{dn_{eff}}{d\omega} + \frac{\omega}{c}\frac{d^2 n_{eff}}{d^2\omega} \text{ ps/THz/cm} \tag{9}$$

β_2 indicates modal waveguide dispersion. The 1st term of the equation is the 1st derivate of the effective refractive index in core mode with respect to the angular frequency. 2nd term of the equation is the 2nd derivative of the effective refractive index in core mode with respect to angular frequency, $\omega = 2\pi f$. Here f=operating frequency. From Fig. 9 it can be observed that the modal waveguide dispersion decreases with the increasing of the frequency. It can also be evident that the modal waveguide dispersion is low for both x and y-polarized modes at a certain frequency range.

For the overall assessment of a waveguide, the figure of merit (FOM) is an important guiding parameter. This crucial parameter can be quantified by the following equation [8]:

$$FOM = \frac{\alpha_{eff}}{A_{eff}} \tag{7}$$

From the equation it is clear that this parameter is simply found from the EML and Modal effective area. It is the ratio of normalized EML and effective area. Using this equation we calculate FOM for different values of D_{core}. The value of FOM is approximate same for both polarized beam. With the increasing of frequencies FOM is decreasing. For D_{core}=300um FOM is 3.66×10^{-6} um^{-2} for x polarized modes. Besides from the figure it is noted that the FOM is 3.26×10^{-6} um^{-2} for the same conditions.

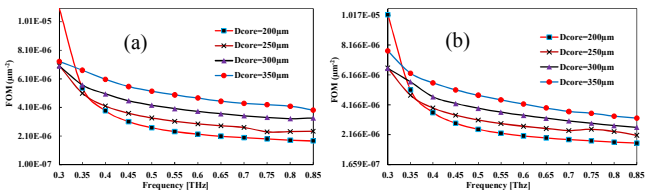


Fig. 11. FOM of proposed waveguide as based on frequency for (a) x-polarized beam (b) y-polarized beam.

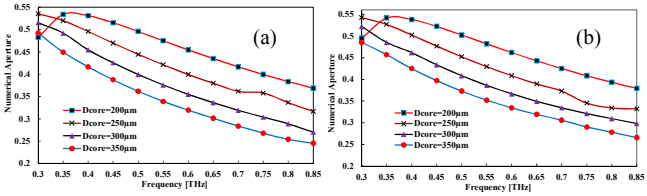


Fig. 12. NA of proposed waveguide as based on frequency for (a) x-polarized beam (b) y-polarized beam.

Light capturing capability is an important factor for a PCF. How much light is captured by a PCF can be defined by the Numerical aperture (NA). A large numerical aperture is needed for the broad sensing operation. High numerical aperture can be found when the effective refractive index difference between the core and cladding is made higher. If this happens more light will propagate through the core. This is the basic of PCF operation. We enhance NA by introducing new material Zeonex. We can calculate NA by using the following equation:

$$NA = \frac{1}{\sqrt{1+\frac{\pi A_{eff} f^2}{c^2}}} \quad (8)$$

From the Fig. 11 it can be seen that light capturing capability is high for our Proposed PCF. We analysis NA for different values of D_{core}. But For D_{core}=200 μm NA is high and for D_{core}=350 μm it is low.

For practical use, it is needed to bend the PCF. When a fiber is bent then the light propagates outer the core and it comes out in the cladding. If this happens light energy is absent at the ending point. Besides transmission length will be reduced. Thus for practical use calculation of bending loss is very important. Bending loss has a great impact on the guided mode power transfer by altering its architecture. Bending loss of a fiber can be not measured directly. Therefore bending loss has been calculated by using a simple easy-to-evaluate analytical method:

$$\alpha_{BL} = \frac{1}{8}\sqrt{\frac{2\pi}{3}}\frac{1}{A_{eff}}\frac{1}{\beta_{co}}F[\frac{2}{3}R\frac{(\beta^2_{co}-\beta^2_{cl})^{(3/2)}}{\beta^2_{co}}] \text{ cm}^{-1} \quad (10)$$

Hear, A_{eff}=modal effective area, R=bend radius, $F(x)=x^{-1/2}e^{-x}$. the proposed fiber has been observed for x polarized mode only. From Fig. 13 it is seen that bending loss decreases with the increasing of operating frequency.

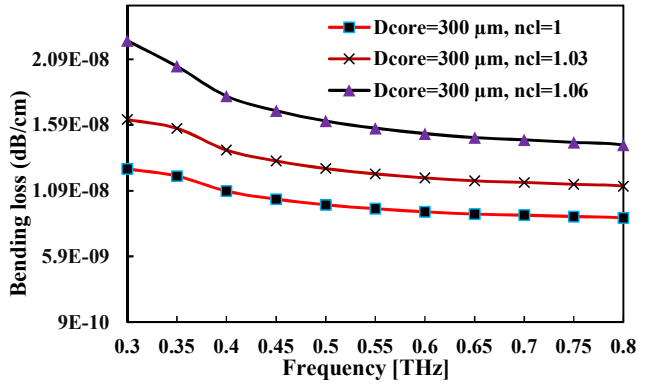


Fig. 13. Bending loss of the proposed waveguide as based on frequency.

For n_{cl}=1 the bending loss is low but for n_{cl}=1.06 the bending loss is high. So from the figure, it is clear that if the more holes are used in the cladding the bending loss will be low. If more air holes are used then the light will not propagates outer layer. Then it is possible to bend the fiber more. From the equation one can see that the unit of bending loss is cm^{-1}. But this term can also be expressed in dB/cm using conversion process. The bending loss is calculated for n_{cl}=1 is 9.13×10^{-9} dB/cm at 0.65 THz.

The performance parameters of the proposed fiber and existing works in the literature are tabulated Table. I.

TABLE I
CHARACTERISTICS COMPARISON

Ref.	Material	*f* (THz)	EML (dB/cm)	α_{cl} (dB/cm)	Power fraction %
[2]	Topas	1	0.37	1.91×10^{-3}	37
[3]	PMMA	0.65	>1.09	2.51	-
[4]	Topas	0.70	0.304	>0.1	-
[5]	Topas	0.75	0.169	0.087	34
[6]	Topas	1	0.39	1.17×10^{-3}	37
proposed	Zeonex	0.65	0.1491	6.7×10^{-11}	48.57

This table includes frequency, modal EML, confinement loss, and fraction of power through the core. From the Table. 1 it is clear that our proposed PCF shows less loss and high power propagates through the core in THz range.

It is crucial to address the manufacture possibility of our introduced waveguide. Several fabrication techniques have developed in recent years [2]. The proposed PCF consists of small-sized of air holes with two different dimensions. In cladding, we use a large air hole and in the core, we use a small air holes. Kiang *et al.* [9] demonstrated a new technique that gives fabrication freedom. This technique helps to demonstrate a complex structure practically. Capillary stacking technique mentioned in [10] used to fabricate circular air holes. The dimension of the air holes are adjusted freely in Sol-gel [2] technique.

IV. Conclusions

A Zeonex based THz waveguide is proposed and analysed that contains rectangular porous-core which offers high fraction of power through the core, low confinement loss, moderate EML and large effective area. Numerical investigation shows high fraction of power through the core over a wide frequency range of 0.60 to 0.85 THz. At 0.65 THz the power fraction of the proposed fiber is 48.57%. A flattened waveguide dispersion is obtained over a large frequency of 0.60 to 0.75 THz. Moreover, moderate EML of 0.1491 dB/cm and about 4.8×10^5 μm^2 modal effective area can be obtained for the optimum design parameters. This proposed structure acts as a single mode fiber, therefore it can be used for long distance transmission.

References

[1] S. Atakaramians, S. Afshar, T. M M. Monro, D. Abbott, "Terahertz dielectric Waveguides," *Adv. in opt. and photonics*, vol. 5, pp. 169-215, 2013.

[2] M. Hasan, M. S. Anower, M. A. Islam, S.M.A Razzak, "Polarization-maintaining low-loss porous-core spiral photonic crystal fiber for

terahertz wave guidance," *Appl. Optics*., vol. 55, pp. 4145-4152, 2016.

[3] S. Atakaramians, V. Shahraam Afshar, B. M. Fischer, D. Abbott, and T. M. Monro, "THz porous fibers: design, fabrication and experimental characterization*," Opt. Express*, vol. 17, pp. 14053–15062, 2009.

[4] R. Islam, M. S. Habib, G. K. M. Hasanuzzaman, S. Rana, M. A. Sadat, and C. Markos, "A novel low loss diamondcore porous fiber for polarization maintaining terahertz transmission," *IEEE Photon. Technol. Lett.,* vol. 28, pp. 537-1540, 2016.

[5] Md. Rabiul Hasan, S. Ali, S.A. Emi, "Ultra-low material loss microstructure highly fiber for terahertz guidance," *Photonics Lett.of poland,* vol. 9, pp. 66-68, 2017.

[6] M. Hasan, Md. Ariful, M.S. Anower, S.M.A. Razzak, "Low loss and bend insensitive terahertz fiber using a rhombic shaped core," *Appl.Opt.,* vol. 55, pp. 8441-8447, October 2016.

[7] G. Woyessa, A. Fasano, C. Markos, A. Stefani, Henrik, "Zeonex microstructured polymer optical fiber. Fabrication friendly fibers for high temperature and humidity insensitive Bragg grating sensing," *Opt. Mater. Exp.*, vol. 7, pp. 286-295, 2017.

[8] J. Liang, L. Ren, N. Chen, C. Zhou, "Broadband, low-loss, dispersion flattened porous core photonic band gap fiber for terahertz (THz) wavepropagation"*, Opt. Commun*., vol. 295, pp. 257–261, 2013.

[9] K. M. Kiang, K. Frampton, T. M. Monro, R. Moore, J. Tucknott, D. W. Hewak, D. J. Richardson, and H. N. Rutt, "Extruded singlemode non-silica glass holey optical fibres," *Electron. Lett.,* vol. 38, pp. 546–547, 2002.

[10] M.I Hasan, M.S Habib, S.M Abdur Razzak, "Design of hibrid photonic crystal fiber, polarization, dispersion propertices, photonic nano structure'', *fundam. Appl.,* vol. 12, pp. 205-211, 2014.

Classification of Intrinsic Mode Spectra of EEG Signal for Epilepsy Detection

Ankush Sarker
Department of Electrical & Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
ankushsarker@gmail.com

Md. Masudur Rahman
Department of Electrical & Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
rahman.masud1994@gmail.com

***Abstract*— In this paper, we proposed EEG based epileptic seizure detection where statistical bandwidth of intrinsic mode functions (IMFs) are used. EEG signals are decomposed into IMFs through the implementation of Empirical Mode Decomposition. These IMFs are turned into analytical functions by means of Hilbert transformation. Euclidean value of Amplitude and Frequency Modulated Bandwidths are calculated from these analytic functions. Finally, the statistical distribution bandwidth of these different IMFs is considered as a feature. The classification was done by means of Support Vector Machine (SVM) with the Gaussian kernel. For training purpose, 75% of the data is utilized. Finally, 89.6% accuracy was gained for the remaining 25% test data in the SVM classifier.**

Keywords*— Empirical Mode Decomposition (EMD), Intrinsic Mode Functions (IMF), Hilbert Transform, Support Vector Machine.

I. INTRODUCTION

Epilepsy is the 4th common neurological disorder where the neurons start firing abnormally which finally leads to the formation of seizure. It is mostly found in people at the age of 7-15 and 65-70 and about 1 % of the worldwide population is affected by this disease [1]. Eighty percent of epileptic seizure severity can be minimized if identified properly. Bangladesh, being a developing country, the portion of epileptic patients are increasing because of the inadequacy of accurate diagnosis technique.

EEG is a low-frequency signal which can give info about epileptic seizure inside the brain. Manual detection of epilepsy can be time-consuming and inaccurate. That's why computerized EEG based automatic epilepsy detection is necessary. EEG based epilepsy detection has mainly two steps: feature extraction and classification. Many renowned EEG feature extraction processes have been used worldwide. Fourier transform of EEG signal is commonly used for classification of seizure given the signal is stationary [2]. But there persists a window problem where an exchange between time and frequency resolution has to be considered. But as EEG is a non-stationary signal, methods like the smoothed pseudo-Wigner-Ville distribution is used which depends on time-frequency domain [3]. But the method uses a genetic algorithm which fails to define an optimum solution. Consequently, DWT extracted line length features for sub-band decomposition of EEG signals are used to detect epileptic seizure [4]. But its computational expense increases in case of fine analysis. Recently, many nonlinear methods are used for parameter extraction to analyze EEG signals. Among these parameters, both complexity and spectrum analysis features have been popular for the identification of epileptic EEG signals [5]. As an EEG signal is non-linear signal and the total data span is too short [10]. That's why it is quite challenging to accurately extract feature for seizure EEG signal detection. For solving these problems in a non-linear signal, EMD based feature extraction technique was first introduced by Huang et al. in 1996 [10]. Moreover, it has been revealed that the conventional EMD method has lesser computational complexity and more classification accuracy compared to multivariate EMD [7]. So nowadays it has become a state of art of field of signal processing. A non-stationary EEG signal is broken into some predictable numbers of intrinsic mode functions (IMF) and a residue by this process. The IMFs are a form of data expansion for the decomposition. Their orthogonality cannot be predicted but their scale is predetermined. So IMFs have to be transformed into analytic signals. Because the analytic functions provide time-dependent instantaneous frequency (IF) and amplitude functions which has valuable information on the local properties of the signal. Hilbert Transform is applied on these IMFs to determine analytic function. Hilbert Transform is a useful tool in situations where frequency fluctuates with time, namely non-stationary signals. It is considered much more efficient than FFT in case of non-stationary signals [6]. So after the conversion into the analytic signals, these instantaneous frequency functions and amplitude functions are calculated. Because they are needed for extracting the amplitude modulated (AM) and frequency modulated (FM) bandwidths. These AM and FM bandwidths were taken as features for seizure detection erstwhile [13]. But these two features show little promise for accurate detection. Therefore, our proposal is the calculation of their Euclidean total bandwidth. Because the variation among these bandwidths for different IMF carries important information for epileptic seizure detection. So, the mean and variance of the total Euclidean bandwidth of all IMFs of a single EEG signal are considered as corresponding features. This distribution of bandwidth feature is supposed to increase the ability to detect epileptic seizure.

Finally, for data classification, Different types of classifier like the random forest, Naïve bias, neural network, SVM, KNN are used. But SVM has global minima that avoid overfitting and require less number of data for data training. SVM has broadly used in classification, regression and feature reduction task. In most cases, it is used for classification purpose for supervised learning. It mainly maximizes the margin to create a hyperplane with the help of kernel function. Kernels are used to transform the feature vectors from input space to kernel space for linearly non-separable datasets. Different types of kernels are used like linear, polynomial, sigmoid and Gaussian in SVM classifier. The linear kernel is used for a huge number of input features. For a few feature space, the Gaussian kernel has performed better. That's why SVM classifier with Gaussian kernel is used for classification purpose.

The above mention EEG seizure detection algorithm is used EMD based feature extraction and SVM with Gaussian kernel. Autoregressive Coefficient Extraction with the SVM classifier based system has an accuracy of 54.77%. Using Power Estimation with FFT and Fisher Discriminant Analysis based system accuracy is 72.4%. Spectral Density Estimation based Bayes Quadratic Classifier has increased the rate to 81.5%. The Stock-well transform feature extraction process associated with KNN classifier produced 84% accuracy rate. For signal energy estimation using DWT and feature classification using SVM picked up the accuracy rate to 84.73% [14]. Compared to these five methods, our EMD feature based proposed method needs less computation complexity and it performs better in small noise signal and SVM classifier required less amount of data and less time required for data training. So it is expected SVM classifier with EMD feature extraction will perform better.

As from here, section II will describe our method of feature extraction and classification process. The result and discussion of this method are given in section III. Finally, we have a conclusion for the paper in section IV.

II. Methodology

A. Dataset:

EEG dataset which is used in this paper is publicly available online [8]. The dataset is divided into five categories which are symbolized by Z, O, N, F, and S. There are 100 single channel EEG signals in each set which makes a total data of 500. Each one has 23.6s duration and a sampling rate of 173.61 Hz. A bandpass filter was used where the band frequency was approximately 0.50~40 Hz (12 dB/oct). The subset N, F and S were recorded intracranially and subset Z and O were recorded extracranially from five healthy volunteers with eyes opened and closed respectively.

The EEG signals from subset F and N are taken from five patients when they were in epilepsy-free intervals. The subset S contains epileptic activity which is mainly complex partial seizure. So here signals from subset Z, O, N and F are considered as non-seizure class and signals from subset S is considered as seizure class. Fig.1 shows one signal from each category.

.

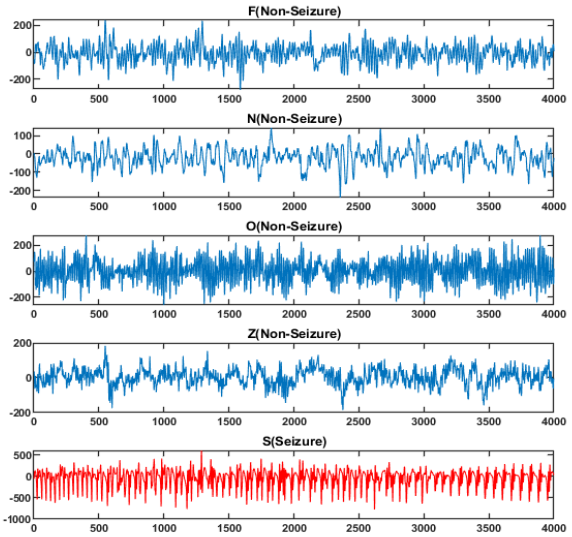


Fig. 1. EEG signal of five sets.

B. IMF Generation using EMD:

The hypothesis that a number of different simple intrinsic

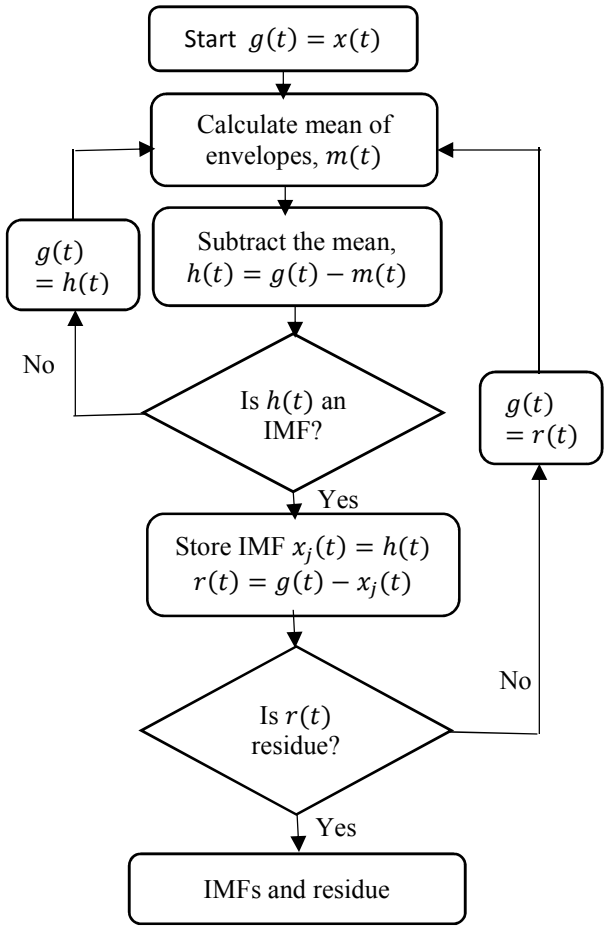


Fig. 2. Flow Diagram of the EMD Algorithm.

modes of oscillations forms a non-stationary and non-linear time series established the EMD method. It is a form of time domain analysis like Fourier transform and Wavelet decomposition. The objective of the method is to empirically recognize these intrinsic modes by their local characteristics of time scales in the data, and then decompose the data into IMF accordingly [9]. EMD filters out the time-frequency information of the IMF. It converts a signal $x(t)$ into its constituent IMFs maintaining two conditions: 1) An IMF has one single extremum between two subsequent zero crossings. 2) An IMF has a mean value of zero.

EMD algorithm: The following algorithm shows the process for extracting the IMF from the EEG signal. (Fig. 2). It decomposes a signal $x(t)$ into a set of IMF $c_n(t)$ and one residue r(t). So the signal can be represented as their summation like the equation below

$$x(t) = \sum_1^n c_n(t) + r(t) \quad (1)$$

After EMD is executed one signal is represented by several IMFs and one residue as shown in Fig. 3.

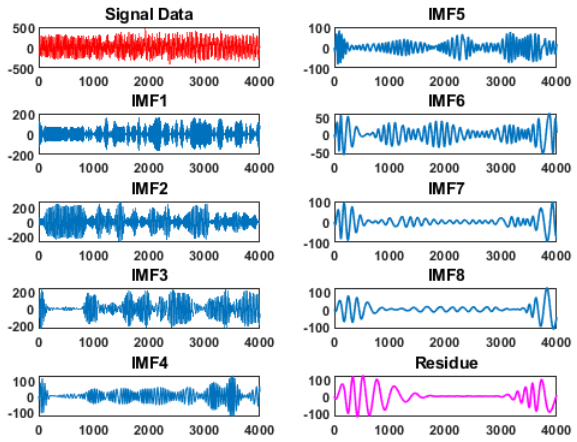


Fig. 3. Empirical Mode Decomposition of a Seizure Data Sample.

C. IMF to Analytic signal conversion:

An analytic signal contains complex functions which have no negative valued frequency component. The Hilbert Transform is done to obtain analytic functions from IMFs [10]. It instructs every Fourier function component to make a phase shift of 90°. Hilbert transform is the convolution of any real function with another function called Cauchy kernel which is, $h(t) = 1/\pi t$.
The final function is given by $c_H(t)$

$$c_H(t) = c(t) * h(t) \quad (2)$$

Every real IMF can be converted to an analytic signal using the equation below

$$z(t) = c(t) + jc_H(t) \quad (3)$$

Equation (3) can be written as

$$z(t) = A(t)e^{j\varphi(t)} \quad (4)$$

The amplitude $A(t)$ and instantaneous phase $\varphi(t)$ of the analytic signal can be determined using equations as follows

$$A(t) = \sqrt{c^2(t) + c_H^2(t)} \quad (5)$$

$$\varphi(t) = \arctan[\frac{c_H(t)}{c(t)}] \quad (6)$$

The instantaneous frequency (IF) of the analytic signals can be gained from equation (6) as given below

$$\omega(t) = \frac{d\varphi(t)}{dt} \quad (7)$$

$A(t)$ and $\omega(t)$ of a random signal from 'S' category is shown in Fig. 4.

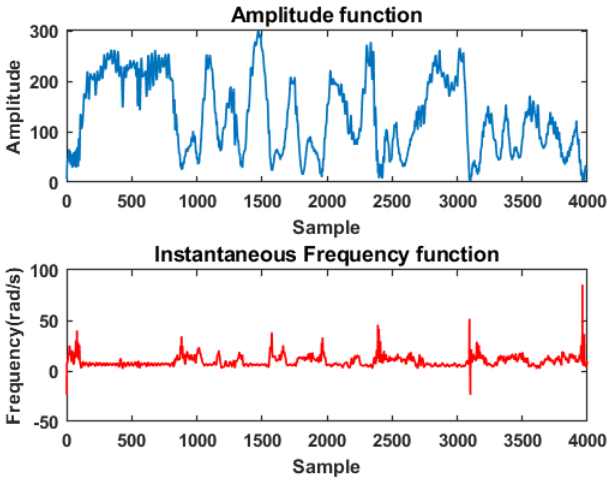


Fig. 4. Amplitude and Instantaneous Frequency (IF) function of the analytic signal.

D. Acquisition of Total Bandwidth:

The bandwidth is a measure of the spread in frequencies for the total time of the signal. Amplitude change and mean frequency of a signal can deviate the frequency range. But the bandwidth cannot precisely determine for which parameter the range changes in a signal. In order to determine, for which parameter the range changes, two different quantities are measured in this case namely amplitude modulated bandwidth (B_{AM}) and frequency modulated bandwidth (B_{FM}) [11]. Each analytic signal contains different energy levels for different working conditions. The energy of a signal z(t) is determined by the eq. (8)

$$E = \int |z(t)|^2 dt \quad (8)$$

If $Z(\omega)$ denotes Fourier transform of signal $z(t)$ then the center frequency of the analytic functions $z(t)$ is given by

$$< \omega >= \frac{1}{E} \int \omega |Z(\omega)|^2 d\omega \quad (9)$$

The amplitude modulated and frequency modulated bandwidth can be given as

$$B_{AM}^2 = \frac{1}{E}\int(\frac{dA(t)}{dt})^2 dt \quad (10)$$

$$B_{FM}^2 = \frac{1}{E}\int\left(\frac{d\varphi(t)}{dt} - < \omega >\right)^2 A^2(t)dt \quad (11)$$

The Euclidean value of Amplitude and Frequency modulated bandwidths in any analytic IMF $z(t)$ is given by

$$B = \sqrt{B_{AM}^2 + B_{FM}^2} \quad (12)$$

The Euclidean total bandwidths of different IMF for a random signal from each category is shown in Fig. 5.

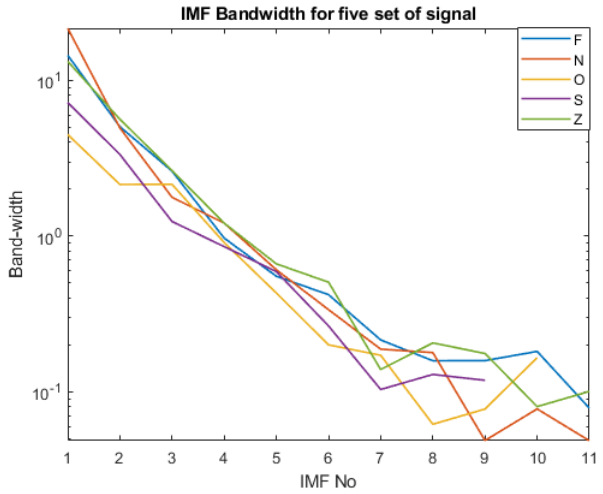


Fig. 5. IMF Bandwidth of sample signal from five datasets.

The mean and variance of all IMF Euclidean Bandwidth are taken as features. These two feature value is diverse for the Seizure and non-Seizure of EEG signal which is visualized in Fig. 6.

E. Data classifiers:

For non-linear classification, Support vector machine, K-nearest Neighbor and Neural Network are mainly used. But SVM is the most popular technique for solving the pattern recognition problem. SVM classifier mainly creates the decision boundary using N-dimensional separating hyperplane that maximizes the margin and identifying the different classes of data [12]. For non-linear classification, kernels are used. It calculates similarities between observations in the feature space of training set example that's why the selection of proper kernel function is important. Gaussian Kernel is the most popular kernel choice in SVM. The Gaussian kernel can be defined as [13]

$$K(x, x_i) = e^{\frac{-|x-x_i|^2}{2\alpha^2}} \quad (13)$$

Where α controls the width of the RBF function.

So The SVM classifier decision function is

$$f(x) = sign[\sum_{i=1}^{N} \alpha_i y_i k(x, x_i) + b\,] \quad (14)$$

Where α_i is called Lagrangian multiplier and b is is called bias.

The mean and variance of bandwidth are given as input feature in the Support Vector Machine. For training purpose, 75% data of each subset as a total of 375 data is used as training of SVM classifier. Fig. 6 shows a scattered plot of two different data classes- seizure and non-seizure.

After training the data, we have found an SVM trained network. Then 125 test data (25% of overall data) is tested in these SVM networks for predicting these unknown class. From Fig. 7 show the test data scattered plot and for proper visualization, SVM decision boundary is also included where the purple colored region is seizure and green is non-seizure. It shows maximum test data is contained in there following proper region.

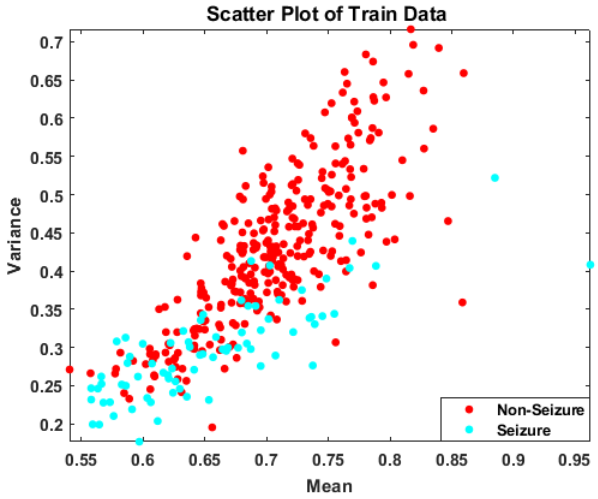


Fig. 6. Scattered plot of train data

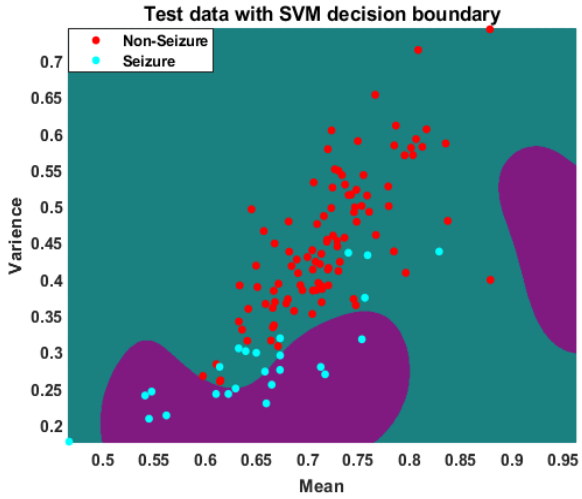


Fig.7. SVM decision boundary with test data.

III. Results and Discussion

A very important part of research work is to calculate the success rate and accuracy. Without it, the research might not have any value. So we are going to discuss the validity and accuracy of our work.

After extracting the features from IMFs of each signal we can see that the seizure signal features are less in magnitude than the non-seizure signal features. So thus we can effectively understand that our features can actually be useful in SVM.

The outcomes of our SVM classifier are represented in the confusion matrix given below. A confusion matrix is a tool for the execution of measurement of any classification algorithm which uses a specific table layout for supervised learning. The rows and columns are represented by output and target class respectively. Where diagonal element shows the number and percentage of the correctly classified and while off-diagonal term represents the misclassified observations. The cell at the last row of the last column shows the overall accuracy

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (15)$$

Where the total number of correctly detected positive events is indicated by *TP,* the number of correctly detected negative events indicated by *TN,* the total number of incorrectly detected positive events is indicated by *FP* and the number of incorrectly detected negative events indicated by *FN*.

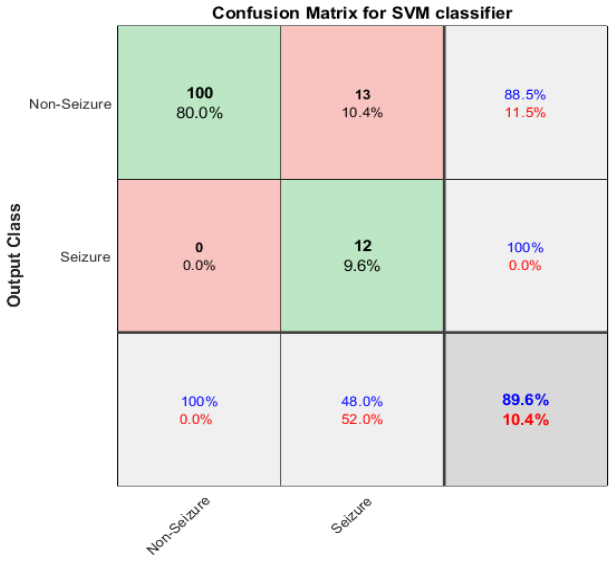


Fig. 8. Confusion matrix for SVM Classifier

Fig. 8 shows the confusion matrix for SVM. We can understand from the confusion matrixes 89.6%. From the confusion matrix, we can easily understand that our feature extraction method has better gross output than other methods [14].

IV. Conclusion

The most used techniques for EEG signal classification require much more sampled data and a wide number of extracted features. We have used EMD for feature extraction of EEG signals and identified the seizure class using support vector machine. The bandwidth features are used for classification in this method. This technique resulted in better accuracy and proficiency. The suggested method is used because of its lower requirement of data and features associated with a strong classifier.

The recommended method may be applied for examination and sorting of other non-stationary signals. Future guidelines of this research include real-time classification using the convolutional neural network (CNN) and the increment of classification accuracy using Wavelet kernel function.

References

[1] *Seizures and the Epilepsies: Hope Through Research*. National Institute of Neurological Disorders and Stroke, 2015. [Online]. Available: https://catalog.ninds.nih.gov/pubstatic//18-NS-156/15-156.pdf. [Accessed Dec. 20, 2018].

[2] K. Polat and S. Güneş, "Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform." *Appl. Math. Comput.*, vol. 187, pp. 1017-1026, April 2007.

[3] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in EEGs using time–frequency analysis." *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 5, pp. 703-710, Sep. 2009.

[4] L. Guo, D. Rivero, J. Dorado, J. R. Rabunal and A. Pazos, "Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks." *J. Neurosci. Methods*, vol. 191, pp. 101-109, 2010.

[5] S. F. Liang, H. C. Wang and W. L. Chang, "Combination of EEG complexity and spectral analysis for epilepsy diagnosis and seizure detection." *EURASIP J. Adv. Signal Process.*, vol. 2010, p. 853434, 2010.

[6] A. Venkitaraman and C. S. Seelamantula, "Fractional Hilbert transform extensions and associated analytic signal construction." *Sig. Proc.*, vol. 94, pp. 359-372, 2014.

[7] N. Rehman, and D. P. Mandic, "Multivariate empirical mode decomposition." *Proc. R. Soc. Lond. A.*, vol. 466, pp. 1291-1302, 2010.

[8] EEG Data Download [Online]. Available: http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3&changelang=3 . [Accessed Jan. 20, 2019].

[9] A. Zeiler, R. Faltermeier, I. R. Keck, A. M. Tomé, C. G. Puntonet and E. W. Lang, "Empirical mode decomposition-an introduction." in *IEEE Int. Jnt. Conf. N. Netw.*, pp. 1-8, 2010.

[10] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis." *Proc. R. Soc. Lond. A*, vol. 454, pp. 903-995, 1998.

[11] L. Cohen and C. Lee, "Instantaneous bandwidth for signals and spectrogram." in *1990 Proc. IEEE Int. Conf. Acoust. Speech, Signal Process*, pp. 2451-2454.

[12] V. Vapnik, *The nature of statistical learning theory*. New York: Springer-Verlag, 1995.

[13] V. Bajaj, and R. B. Pachori, "Classification of seizure and nonseizure EEG signals using empirical mode decomposition." in *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 6, pp. 1135-1142, 2012.

[14] H. U. Amin, W. Mumtaz, A. R. Subhani, M. N. M. Saad, and A. S. Malik, "Classification of EEG signals based on pattern recognition approach", *Frontiers in computational neuroscience,* 2017.

Detection of Abnormal Electrocardiogram (ECG) Using Wavelet Decomposition and Support Vector Machine (SVM)

Pratik Deb[1], Mohammad Nooruddin[2], Md. Shajahan Badsha[3]
Khulna University of Engineering & Technology, Khulna.
pratikdeb12@gmail.com[1], mohammadnooruddin58@gmail.com[2], shajahanb55@gmail.com[3]

***Abstract*—Electrocardiogram (ECG) is a graphical representation of the electrical activity of the heart which is obtained by placing various electrodes on some specific positions of the body surface of the subject. Abnormalities in the ECG signal of a patient may indicate cardiac diseases that need to be attended by physicians on an urgent basis. Hence, it is necessary to detect an abnormal ECG for the betterment of the patient. Such a method to classify ECG signals whether they are normal or abnormal is developed in this work. Angina, Bundle Branch Block, Cardiomyopathy Heart Failure, Dysrhythmia, Myocardial Hypertrophy, Myocardial Infarction, Myocarditis, Valvular Heart Disease: all these cardiac conditions have been classified as abnormal ECG signal in our work. First, statistical features like skewness, kurtosis, standard deviation of detail and approximation coefficients of the Daubechies wavelet (db10) of order 5 for a number of abnormal and normal ECG signals obtained in the feature extraction stage. Secondly, Support Vector Machine (SVM) was used for classification which was trained by the features extracted in the first stage. Finally, the accuracy, sensitivity, specificity of this method was checked by testing the SVM with 36 signals obtained from MIT-BIH Normal Sinus Rhythm Database and 36 signals from PTB Diagnostic ECG Database which yielded an accuracy, sensitivity, specificity 98.61%, 97.37%, 97.22% respectively.**

Keywords—ECG, SVM, sinus, signals, coefficients, features, sensitivity, accuracy, specificity.

I. INTRODUCTION

ECG signals are obtained through non-invasive measure which shows the electrical activity of the heart. Abnormalities in ECG signal can indicate a case of medical emergency which is needed to be taken care of immediately as it will deteriorate the condition of the patient under observation. So, it is necessary to develop a method to classify the ECG signals based on their shapes and characteristics whether they indicate a cardiac disease or not [1].

Cardiovascular (CVD) and circulatory diseases are the leading causes of death in the world. In 2013 there were more than 54 million deaths (95% uncertainty interval [UI], 53.6–56.3 million) globally and 32% of these deaths i.e. 17 million (95% UI, 16.5–18.1 million), were due to CVD. Most of these CVD deaths were found to be due to either ischemic heart disease (IHD) or cerebrovascular disease [2]. Hence, a system is required that will aid the physicians and diagnosticians in detecting an abnormal ECG signal which resembles a troubled cardiac condition. Such a system is developed through our proposed method. For this method to work with high efficacy it is necessary to train it with a stiff database that will enable it to separate a normal ECG from an abnormal ECG [3]. This method is fast, easy and very accurate. The entire work was carried out in MATLAB. Databases used to train the system are:

- The Physikalisch-Technische Bundesanstalt (PTB) Diagnostic ECG Database
- The MIT-BIH Normal Sinus Rhythm Databaseprovided.

The MIT-BIH Normal Sinus Rhythm Database includes 18 long-term ECG recordings of subjects referred to the Arrhythmia Laboratory at Boston's Beth Israel Hospital (now the Beth Israel Deaconess Medical Center). Subjects included in this database did not have any significant arrhythmias; they include 5 men, aged 26 to 45, and 13 women, aged 20 to 50. The Physikalisch-Technische Bundesanstalt (PTB) Diagnostic ECG Datababase contains a total of 549 records from 290 subjects (age range 17-87, mean age 57.2; 209 men, mean age 55.5, and 81 women, mean age 61.6; ages were not recorded for 1 female and 14 male subjects).

In patient monitoring system, one of the parameters monitored constantly is the Electrocardiogram of the patient [4]. It is because if there is anything wrong with the patient's heart, the ECG will convey this information. The purpose of this work is to classify a number of cardiac diseases as abnormal ECGs using a fast and congenial method with high efficacy. Such a method is needed if any kind of abnormality appears in the ECG, even a layman can understand it and notify the doctors.

Various papers have been published regarding detection of abnormal ECG signals. These techniques primarily use various classifiers to classify the abnormal and normal ECG signals. Some of them also use morphological features to classify the ECG. The methods also vary in extracting features though there are some similarities too. A strategy dependent on LabVIEW (Laboratory Virtual Instrument Engineering workbench) is described in [2]. The purpose of LabVIEW signal processing tools is to denoise the signal before applying the developed algorithm for feature extraction. Primary feature that has been extracted are the R-peaks which is done using Haar wavelet. For low resolution signals daubechies (db6) wavelets are used. Other features used in this paper are Q wave width, T wave, P wave, R wave width, ST segment and T wave amplitude and duration. Frontal plane axis scoring pattern is applied for abnormality detection. In this study Tachycardia, Bradycardia, Left Ventricular Hypertrophy, Right Ventricular Hypertrophy and

Myocardial Infarction have been considered as abnormal ECG. A wavelet coherence (WTC) technique has been proposed in [3] for ECG signal analysis. WTC is used to calculate the similarity between two waveforms in frequency domain. Parameters extracted from WTC function is used as the features of the ECG signal. These features are optimized using Bat algorithm. The Levenberg Marquardt Neural Network (LVNN) classifier is used to classify the optimized features. Atrial fibrillation is classified as abnormal disease in this work. The method proposed in [4] focuses on classifying Myocardial Infarction as an abnormal ECG signal. In the feature extraction stage, first segmentation of each ECG beat is performed. Then every ECG beat is processed by flexible analytic wavelet transform (FAWT) for decomposing into subband signals. Using these subband signals Sample entropy (SEnt) is computed. Finally in the classification stage, the sample entropy values are fed to a number of classifiers, namely: J48 decision tree, Random Forest (RF), Least-Squares Support Vector Machine (LS-SVM) and Back Propagation Neural Network (BPNN). An approach based on Cross Wavelet Transform (XWT) is described in [5]. One way to measure similarity between two waveforms in time-domain is to find out the cross-correlation between the signals. The Morlet wavelet is the mother wavelet used in this paper. When XWT is applied to a pair of data, it gives us Wavelet Cross Spectrum (WCS) and Wavelet Coherence (WCOH). If the characteristics or pattern of the ECGs are different, then WCS and WCOH would demonstrate distinguishing characteristics over two specific locales or regions R1 and R2, where R1 stands for the QRS complex and R2 stands for the T wave region. After marking discernible dissimilarities of the signals in R1 and R2, ECGs are classified as normal and abnormal based on an appropriate threshold.

The method for abnormality detection developed in this work is different from the works mentioned in this section in such a way that it does not require any kind of preprocessing stage. It's a straightforward method which begins by taking the detail and approximation coefficients from daubechies wavelet of order 10 of the normal and abnormal ECGs. The reason behind choosing 'db10' wavelets is Daubechies wavelets family provides better performance for detecting discontinuities and transients with respect to continuous wavelet transform**.** Then the features, skewness, kurtosis, standard deviations are extracted for training the SVM. Skewness, kurtosis, standard deviation was chosen as features over other statistical features because they can determine the extent of asymmetry of the distribution, measure it's tailedness and determine its dispersion. Finally, the abnormal ECG signals are detected.

We have organized the paper as follows: Section II describes the complete methodologies for the proposed method. Section III includes the results and discussion of this work and gives a comparison between other works and our proposed method. Finally, Section IV comes with the conclusion of the whole work done in this paper.

II. METHODOLOGY

For doing observation and analysis we have used ECG records which were collected from the MIT-BIH Normal Sinus Rhythm Database and were made available through Physionet. The range of age of the subjects is 20-42 years.

A. Workflow of proposed feature extraction method

A simplified workflow of the feature extraction method is described in this section. First a certain number of signals are selected for the purpose of feature extraction and training the Support Vector Machine (SVM). First, only the normal ECG signals are selected. Then they are decomposed into 5 levels using Daubechies wavelets of order 10 ('db10') and their detail coefficients and approximation coefficients are extracted. Skewness, Kurtosis, Standard Deviation of these coefficients are calculated which are the desired features [7]. Using these features, a feature matrix and a group matrix are created. A numerical value of '1' is used in the group matrix for normal ECGs. Same procedure is followed from the beginning for abnormal ECGs with the exception being that a numerical value of '0' is used in the group matrix [8]. A block diagram for the workflow is given in Figure 1.

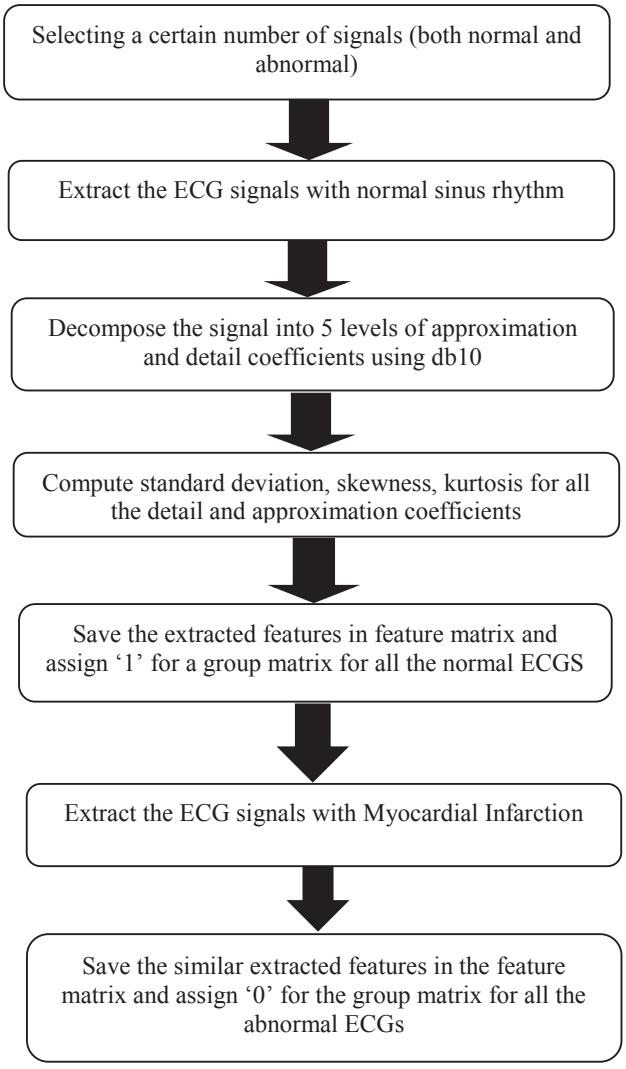


Figure 1: Workflow of proposed feature extraction method

B. Feature Extraction Method

Here, the proposed method for feature extraction has been described. All the steps for wavelet decomposition using Daubechies wavelets of order 10 i.e. 'db10' into 5 levels have been described step-by-step [9]. The relevant MATLAB waveforms are also given. Creation of feature matrix and group matrix for training the Support Vector Machine is explained too [10].

First, a number of ECG signals are selected for feature extraction which includes both normal and abnormal ECG. All the files are in mat extension format. Each file representing abnormal (diseased) ECG contains signals from 15 different leads whereas each file for normal ECG contains signals from 2 different leads. The steps of feature extraction are described below:

1. Seven (7) .mat files (105 ECG signals) with Myocardial Infarction i.e. abnormal ECG and seven (7) .mat files (14 ECG signals) having normal sinus rhythm are selected for feature extraction.

2. There are 2 dimensions (ECG signals) in a for a recording of normal sinus rhythm. Hence, in the beginning of feature extraction, the first dimension is extracted in Figure 2.

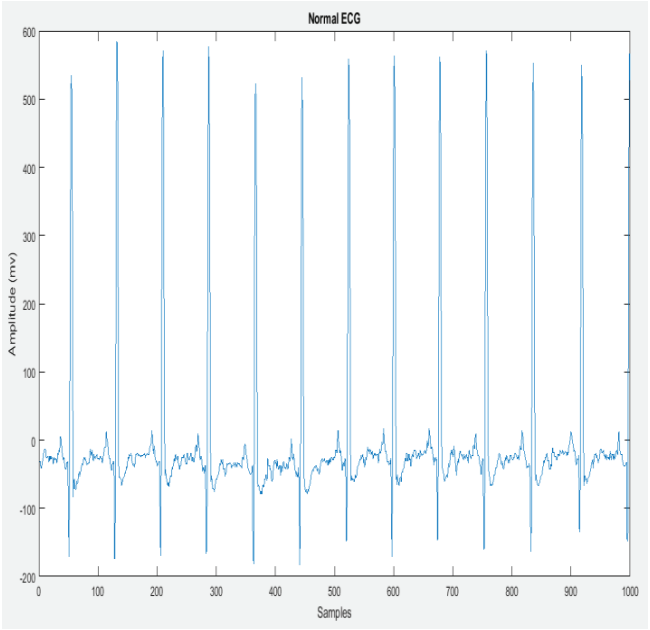


Figure 2: ECG signal of Normal Sinus Rhythm

3. The extracted ECG signal is decomposed into five levels of detail coefficients using 'db10' and saved in variables D1, D2, D3, D4, D5 in matrix form. The extracted coefficients of levels 1, 2, 3, 4, 5 are saved in D1, D2, D3, D4, D5 respectively. D1 coefficients of normal ECG in one channel, D2 coefficients of normal ECG in one channel is shown in Figure 3. D3 coefficients of normal ECG in one channel, D4 coefficients of normal ECG in one channel, D5 coefficients of normal ECG in one channel is shown in Figure 4.

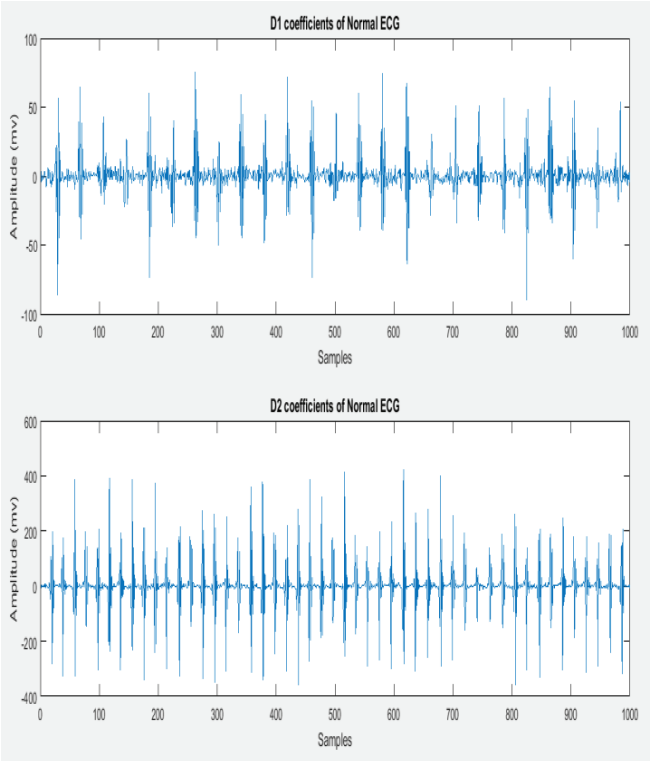


Figure 3: D1, D2 extracted from normal ECG

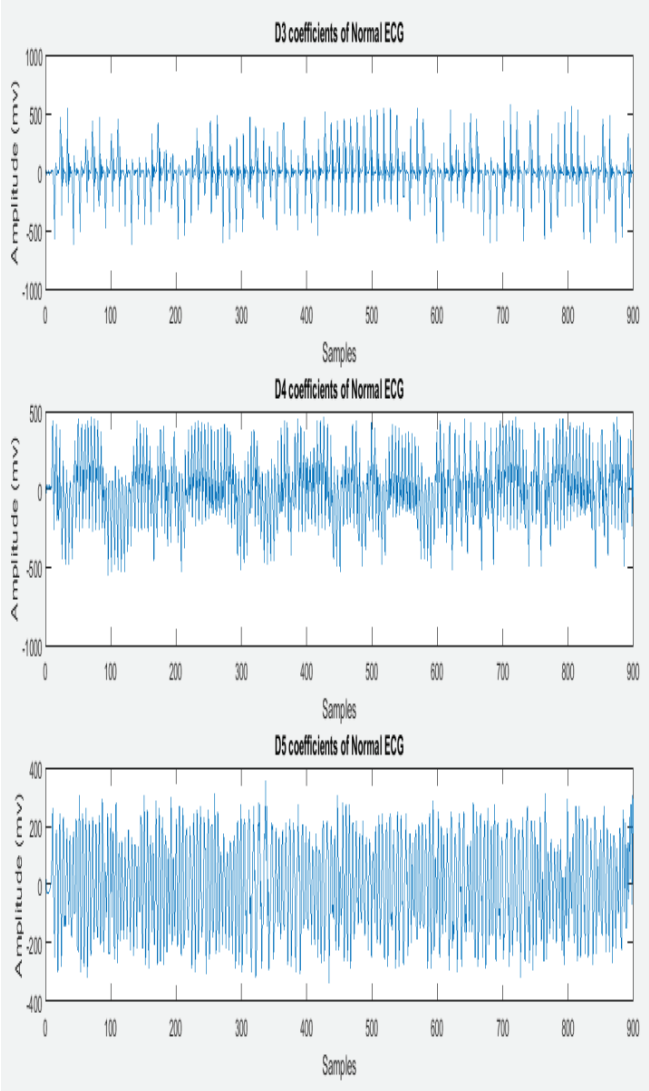


Figure 4: D3, D4, D5 extracted from normal

4. Then, extracted ECG signal is decomposed into five levels of approximation coefficients using 'db10' and saved in variables A1, A2, A3, A4, A5 in matrix form [11]. Coefficients of level 1 (A1), level 2 (A2), level 3 (A3) are shown in Figure 5 and coefficients of level 4 (A4), level 5 (A5) are shown in Figure 6.

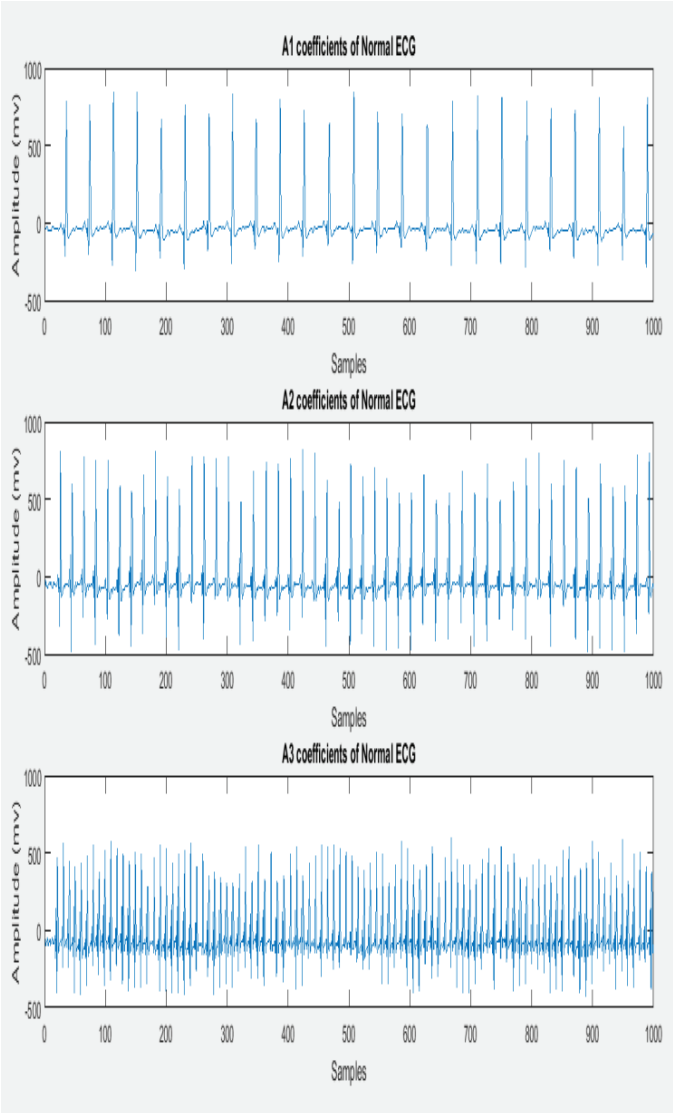


Figure 5: A1, A2, A3 extracted from normal ECG

5. Skewness, Kurtosis, Standard Deviation of each of D1, D2, D3, D4, D5 are calculated and stored in variables ske1, kur1, sd1; ske2, kur2, sd2; ske3, kur3, sd3; ske4, kur4, sd4; ske5, kur5, sd5. Therefore, a total of 15 features have been extracted.

6. Similar values for A1, A2, A3, A4, A5 are stored in variables aske1, akur1, asd1; aske2, akur2, asd2; aske3, akur3, asd3; aske4, akur4, asd4; aske5, akur5, asd5. Again a total of 15 features have been extracted. So, there are a total of 30 features.

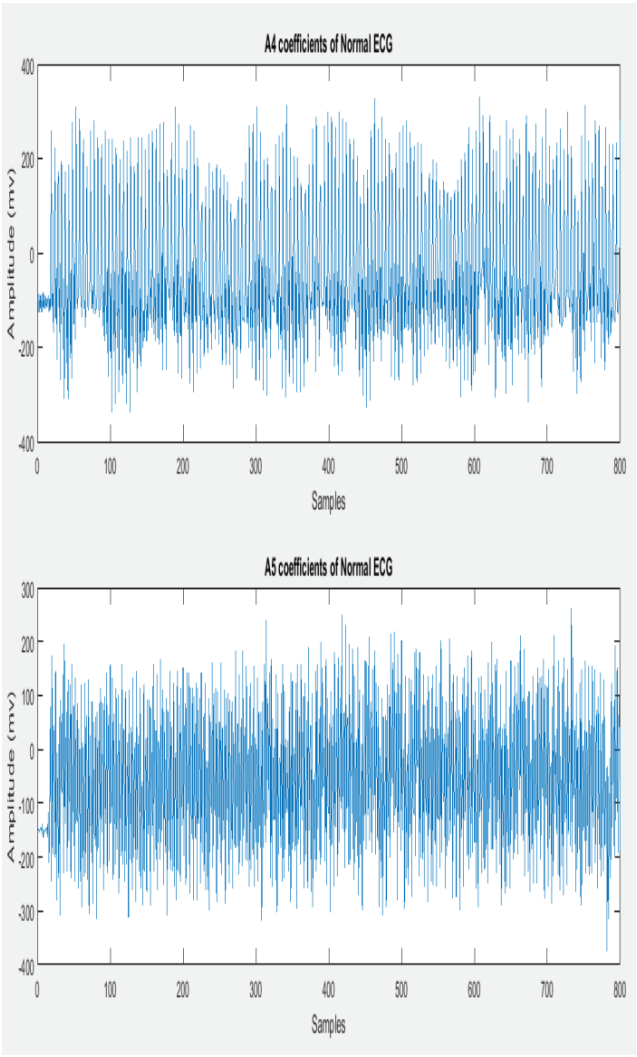


Figure 6: A4, A5 extracted from normal ECG

7. These statistical parameters of both detail and approximation coefficients are saved in the first row of a variable aa.

8. A value of '1' is assigned to the first row of another matrix (namely group matrix) for the first row of aa indicating that it was a normal ECG signal.

9. Identical operation is carried out for the other dimension of ECG signal for this record of normal ECG [12].

10. Hence, the same operation is carried out for a total of 14 times as there are 14 normal ECG signals.

11. Group matrix for normal ECG is created.

12. In the same manner group matrix for 105 abnormal ECG signals is also created with the difference being that a value of '0' is assigned for indicating abnormal ECG. This implies that same statistical features have been extracted for abnormal ECG too. So, the group matrix will have 119 rows and 1 column.

13. So, the final feature matrix will have 119 rows (since the total number of signals is 119) and 30 columns (total number of features are 30).

The discrete wavelet changes (DWT) is an execution of the wavelet transform utilizing a discrete arrangement of the wavelet scales and interpretations complying with some characterized standards. In other words, the signal is broken down into mutually orthogonal arrangement of wavelets using this transform, which is distinguishes itself the continuous wavelet transform (CWT). [15].

C. Abnormality Detection Method

1. The feature matrix and group matrix created in the feature extraction stage are used to train the Support Vector Machine.

2. Separate test matrices are created for normal and abnormal (diseased) ECG signals for testing the Support Vector Machine.

3. If the output of the SVM is '0' then it means that SVM has classified the test signal as an abnormal ECG signal. If the output is '1', then it means SVM has classified the test signal as a normal ECG signal.

4. Calculate the accuracy as it is already known which signal is normal and which one is not.

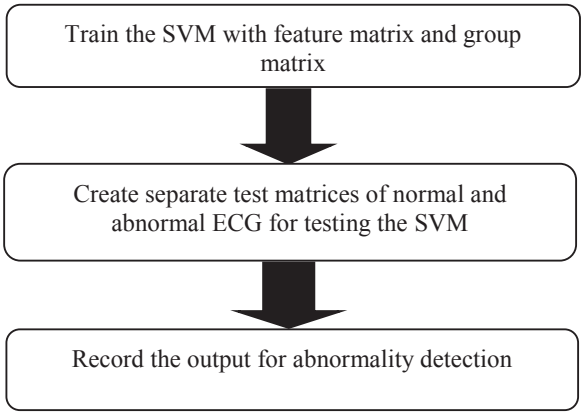


Figure 7: Workflow for abnormality detection

However, in our work we used Support vector machine (SVM) which is a supervised learning method that analyzes data and recognizes patterns, used for classification (machine learning) and regression analysis. The standard SVM is a binary linear classifier. A purpose of a binary classifier is to predict which of the two classes the input belongs. Hence, for a set of training examples, each labeled as belonging to one of two categories, an SVM training algorithm fabricates a model that predicts category or label of the of the input. Intuitively, the examples that are used to build the SVM model can be thought of as points in space, mapped in such a way that examples of the separate categories or labels are divided by a gap that can be as wide as possible. For every new example, it is then mapped into that same space and the category of the example is predicted by utilizing the information on which side of the gap they fall on.

D. Performance Metrics

The performance of the abnormality detection method is evaluated by the widely used metrics: accuracy (AC), sensitivity (SE), specificity (SP). These metrics are calculated using the following equations:

$$AC = \frac{Total\ correctly\ classified\ recordes}{Total\ number\ of\ records} \times 100\% \dots\dots\dots\dots (1)$$

$$SE = \frac{True\ Positive}{True\ Positive + False\ Negative} \times 100\% \dots\dots\dots\dots\dots (2)$$

$$SP = \frac{True\ Negative}{True\ Negatives + False\ Positive} \times 100\% \dots\dots\dots\dots\dots (3)$$

The explanation of the terms True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) are given below:

FP: Normal class classified as abnormal
TP: Abnormal class classified as abnormal
FN: Abnormal class classified as normal
TN: Normal class classified as normal

The performance metrics are explained here which are going to be used to evaluate the performance of abnormality detection method. Then SVM has been described comprehensively as it was instrumental in the completion of this proposed work. A step by step workflow of the abnormality detection method has been given too.

III. RESULTS AND DISCUSSION

Here, the important data for classifying abnormal ECG signals are presented in tabular form. A final overall result is presented in this chapter. The result and the procedure has been compared with other pre-existing methods here later.

A. Data Used For Feature Extraction

First, 7 recordings i.e. a total of 7 ECG records representing Normal Sinus Rhythm were used for feature extraction in our work. These signals were collected from MIT-BIH Normal Sinus Rhythm Database. Then the features of ECGs exhibiting Myocardial Infarction were extracted. The ECG records exhibiting Myocardial Infarction were collected from PTB Diagnostic ECG Database. The records are given in Table I and Table II.

Table I and Table II contain the records used for feature extraction and training the Support Vector Machine in order to detect abnormal ECG signals. The records given in Table I are taken from 7 subjects who exhibited normal sinus rhythm. All the ECG records are of same duration which made the wavelet decomposition convenient as all detail coefficient and approximation coefficient matrices at the same level had equal size. Skewness, Kurtosis, Standard Deviation of the aforementioned records were then calculated.

TABLE I. ECG recordings of Normal Sinus Rhythm

Record	Age	Sex	Duration (Hour:Min:Sec)
16265m	32 years	Male	2:10:12
16272m	20 years	Female	2:10:12
16273m	28 years	Female	2:10:12
16420m	38 years	Female	2:10:12
16483m	42 years	Male	2:10:12
16539m	35 years	Female	2:10:12
16773m	26 years	Male	2:10:12

TABLE II. ECG recordings of Myocardial Infarction

Record	Age	Sex	Duration
s0010rem	81 years	Female	38 min
s0014lrem	81 years	Female	1h 55min
s0015lrem	58 years	Female	1h 55min
s0016lrem	81 years	Female	1h 55min
s0017lrem	63 years	Male	1h 55min
s0020arem	69 years	Male	1h 55min
s0020brem	69 years	Male	1h 55min

Table II contains the records of abnormal ECG signals used to complete the rest of the training of the Support Vector Machine which were collected from PTB Diagnostic ECG Database. Each record includes 15 simultaneously measured signals: the conventional 12 leads (i, ii, iii, avr, avl, avf, v1, v2, v3, v4, v5, v6) together with the 3 Frank lead ECGs (vx, vy, vz). Each signal is digitized at 1000 samples per second. All the patients in the records were diagnosed with Myocardial Infarction. Age of the patients range from 58-81 years. Though duration of every record is not same, the problem was eradicated using appropriate MATLAB functions. To match the lengths of each files, zeros were appended for files that have shorter duration.

B. Overall Result

The following table includes the final output of our proposed i.e. it represents the SVM's ability to detect abnormal ECG signal.

TABLE III. Performance of the proposed abnormality detection method

Number of ECGs	Diagnosis of ECG	Number of ECGs classified as abnormal ECG	Number of ECGs classified as normal ECG
36	Abnormal	36	0
36	Normal	1	35

TABLE IV. Performance of the proposed abnormality detection method (continued)

Number of ECGs	Accuracy	Sensitivity	Specificity
36	98.61%	97.37%	97.22%
36			

From Table III and Table IV, it can be seen that a total number of 72 ECG signals were tested to evaluate the performance of the proposed abnormality detection method. Equal number of normal ECGs (ECGs that exhibit normal sinus rhythm) and abnormal ECGs (ECGs that exhibit cardiac diseases), which means 36 normal and 36 abnormal ECGs were tested by the SVM. When tested, SVM showed the numerical value '0' for all the abnormal signals. Hence, all the abnormal ECGs were correctly classified. While testing the normal ECGs, the SVM showed the numerical value '1' for one abnormal ECG which means it classified one abnormal ECG as a normal ECG. Accuracy, sensitivity, specificity is calculated too.

C. Comparison with Some Pre-existing Methods

Given in this section is a comparison with some of the pre-existing methods for detection of abnormal ECG signal. All the works have used the PTB Diagnostic Database. Two separate tables are used to draw a complete comparison with the results of our proposed method with seven different methods. The metrics chosen for comparison are the feature set used for feature extraction, method of classification, accuracy, sensitivity, specificity. All the aforementioned metrics were obtained from the works mentioned in Table V and Table VI except for the work done by Hamed and Nasser, in which the specificity was not calculated.

TABLE V. Comparison with other works

Work By	Feature Set	Classification	Accuracy
Kumar et al. [16]	FAWT, SEnt	LS-SVM	99.31%
Hamed and Nasser. [17]	QRS complex	Shannon Energy	99.846%
Swati and Mitra [18]	XWT	Heart Rate	99.31%
Kora and Krishna [19]	WTC	LVNN, Bat	99.1%
Huang and Zhou [20]	Morphological, Slope, PCA	SDA, SVM, LLR	88.52%
Rai et al. [21]	Morphological	BPN, MLP, FFN	98.52%
Proposed Method	DWT, skewness, kurtosis, standard deviation	SVM	98.61%

TABLE VI. Comparison with other works (continued)

Work By	Feature Set	Sensitivity	Specificity
Kumar et al.	FAWT, SEnt	99.62%	98.12%
Hamed and Nasser.	QRS complex	99.924%	
Swati and Mitra	XWT	97.3%	98.8%
Kora and Krishna	WTC	96.97%	99.43%
Huang and Zhou	Morphological, Slope, PCA	89.68%	84.62%,
Rai et al.	Morphological	97.92%	96.5%
Proposed Method	DWT, skewness, kurtosis, standard deviation	97.37%	97.22%

From Table V and Table VI we can see that the performance metrics, namely: accuracy, sensitivity, and specificity of the proposed method is somewhat close to the other methods mentioned here. But, our proposed method differs from others in the choice of extracted feature set and the choice of classifier. Another advantage that our proposed method has over the others, is that it eliminates the need for any kind of pre-processing on the electrocardiogram signals before feature extraction. While other methods rely on the morphological features like R-R interval, Q wave, S wave, QRS complex etc and wavelet transforms, our proposed method utilizes the ability of daubechies wavelets (db10) to detect discontinuities and transients. Then skewness, kurtosis, standard deviation of the detail and approximation coefficients are used as features for training the SVM which improves performance since they can determine the extent of asymmetry of a distribution, measure the tailedness and determine dispersion. Hence, though the methods mentioned in Table V and Table VI yield almost similar results, our proposed method distinguishes itself by extracting the appropriate statistical features after the application of daubechies wavelet decomposition which eliminates the requirement of any pre-processing. This makes our proposed method simple yet accurate.

IV. CONCLUSIONS

ECG is instrumental when it comes to delivering important information about the cardiac conditions of a patient. Abnormalities in the ECG signal might be a symptom of a severe cardiac complication which can be life threatening. It is therefore necessary to detect an abnormal ECG signal and receive treatment from concerned personnel.

The method proposed here detects classifies an ECG signal into two groups: normal and abnormal. Eight cardiovascular conditions and diseases, namely: Angina, Bundle Branch Block, Cardiomyopathy Heart Failure, Dysrhythmia, Myocardial Hypertrophy, Myocardial Infarction, Myocarditis, Valvular Heart Disease, all are classified as ECG signals with abnormalities.

The whole procedure was carried out in MATLAB. Detail and approximation coefficients were used as the primary features and SVM was used to classify the ECG signals. The system proved to be quite efficient as it yielded an accuracy of 98.61%.

REFERENCES

[1] A.K. Joshi, A. Tomar, M. Tomar, "A review paper on analysis of Electrocardiograph (ECG) signal for the detection of arrhythmia abnormalities." *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*", vol. 3(10), pp. 12466-12475, 2014.

[2] R. Huang, Y. Zhou. , "Disease classification and biomarker discovery using ECG data." *Biomed Research International",* 2015.

[3] S. Mitra, M. Mitra., "Application of cross wavelet transform for ECG pattern analysis and classification." *IEEE Transaction on Instrumentation and Measurement*", vol. 63(2), pp. 326-333, 2014.

[4] D. Pandit et al., "Improved abnormality detection from raw ECG signals using feature enhancement." Presented at *International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, 2016.

[5] Qian Zheng, Chao Chen, Zhinan Li, Anpeng Huang, Bingli Jiao, Xiaohui Duan and Linzhen Xie (2013), "A novel multi-resolution SVM (MR-SVM) algorithm to detect ECG signal anomaly in WE-CARE project" , *ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC),* 2013.

[6] Venkatesan, C., Karthigaikumar, P., Paul, A., Satheeskumaran, S. and Kumar, R., "ECG Signal Preprocessing and SVM Classifier-Based Abnormality Detection in Remote Healthcare Applications" *IEEE Access*, vol. 6, pp.9767-9773, 2018.

[7] Zining Ding, Feng Wang and Ping Zhou, "Fetal ECG extraction based on different kernel functions of SVM" *3rd International Conference on Computer Research and Development*, 2011.

[8] Raj, S. and Ray, K., "ECG Signal Analysis Using DCT-Based DOST and PSO Optimized SVM" *IEEE Transactions on Instrumentation and Measurement*, vol.66(3), pp.470-478, 2017.

[9] Melgani, F. and Bazi, Y., "Classification of Electrocardiogram Signals with Support Vector Machines and Particle Swarm Optimization", *IEEE Transactions on Information Technology in Biomedicine*, vol. 12(5), pp.667-677, 2008.

[10] Faziludeen, S. and Sabiq, P., "ECG beat classification using wavelets and SVM", *IEEE CONFERENCE ON INFORMATION AND COMMUNICATION TECHNOLOGIES*, 2013.

[11] Varon, C., Caicedo, A., Testelmans, D., Buyse, B. and Van Huffel, "A Novel Algorithm for the Automatic Detection of Sleep Apnea From Single-Lead ECG" , *IEEE Transactions on Biomedical Engineering*, vol. 62(9), pp.2269-2278, 2015.

[12] Sharma, S. and Nagal, D., "Identification of QRS complexes in single-lead ECG Using LS-SVM", *International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)*, 2014.

[13] Sharma, L., Tripathy, R. and Dandapat, S., "Multiscale Energy and Eigenspace Approach to Detection and Localization of Myocardial Infarction", *IEEE Transactions on Biomedical Engineering*, vol. 62(7), pp.1827-1837, 2015.

[14] Ebrahimzadeh, A. and Khazaee, A., "Higher order statistics for automated classification of ECG beats", *International Conference on Electrical and Control Engineering,* 2011.

[15] Khandoker, A., Palaniswami, M. and Karmakar, C. , "Support Vector Machines for Automated Recognition of Obstructive Sleep Apnea Syndrome From ECG Recordings" , *IEEE Transactions on Information Technology in Biomedicine*, vol. 13(1), pp.37-48, 2009.

[16] M. Kumar, R. B. Pachori, U. R. Acharya. "Automated diagnosis of myocardial infarction ECG signals using sample entropy in flexible analytic wavelet transform framework." *Entropy*, 2017.

[17] H. Beyramienanlou, N. Lotfivand., "Shannon's Energy based algorithm in ECG signal processing." *Computational and Mathematical Methods in Medicine*, vol. 17(5), pp. 575-580, 2017.

[18] S. Mitra, M. Mitra. , "Application of cross wavelet transform for ECG pattern analysis and classification." *IEEE Transaction on Instrumentation and Measurement*, vol. 63(2), pp. 326-333, 2014.

[19] P. Kora, K. R. Krishna. , "ECG based heart arrhythmia detection using wavelet coherence and bat algorithm." *Sensing and Imaging*, vol. 17(12), pp. 424-428, 2016.

[20] R. Huang, Y. Zhou. , "Disease classification and biomarker discovery using ECG data." *Biomed Research International,* 2015.

[21] H. Mohan, R. Anurag, T. S. Shukla. , "ECG signal processing for abnormalities detection using multi-resolution wavelet transform and artificial neural network classifier." *Measurement*, vol. 46(9), pp. 3238-3246, 2013.

Epileptic Seizure Detection and Classification using Support Vector Machine from Scalp EEG Signal

1st Sania Zahan
Dept. of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
sania.cse@gmail.com

2nd Md. Rabiul Islam
Dept. of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
rabiul_cse@yahoo.com

***Abstract*—Epilepsy is a neurological disorder originating from brain cells that can affect harshly on patients' life. Imbalance in electrical signals in the neuron cells results in involuntary partial or full body movements or other physiological symptoms. Seizure attack is unpredictable and during its occurrence patient may lose control which can cause serious injury even death. Medical facilities like medication or surgery can be done to improve living condition and life expectancy of patients. For these measures to be beneficial early and correct detection of epilepsy is crucial. However detection from scalp EEG is tough due to the presence of artifacts, the state of the brain and the frequency of seizure occurrence. Hence this study proposes a reliable model of detection system. A zero phase bandpass butterworth filter is used to extract only the EEG signal eliminating all physiological and device artifacts. Frequency distribution of brain signal in both interictal and ictal state differs from that in normal person. So statistical measurements that correctly maps these changes are used to classify the dataset. For classification, a nonlinear support vector machine is used on two sets of dataset combination. Performance of detecting epileptic signal even in the interictal state is promising for use in medical applications.**

***Index Terms*—Electroencephalogram, Bandpass filtering, Entropy, Power spectral density, Support Vector Machine.**

I. Introduction

Epilepsy is a neural disorder of the brain and can occur at any stage of life. It originates from excessive electrical discharges in some group of brain cells and is one of the world's oldest recognized abnormal conditions. For centuries people have feared this disease for its bizarre symptoms and unpredictability. For this reason people with such disorder were misunderstood and discriminated. This can impact very harshly on the quality of life for people with the disorder and their families. The main and only symptom of Epilepsy is Seizures and is defined as having two or more unprovoked seizures varying from the briefest lapses of attention or muscle jerks to severe and prolonged convulsions. Epileptic seizures have four different states: the preictal state, which is a state that appears before the seizure begins, the ictal state that begins with the onset of the seizure and ends with an attack, the postictal state that starts after ictal state, and interictal state that starts after the postictal state of 1st seizure and ends before the start of preictal state of consecutive seizure [1]. However, there are different kinds of seizures which can affect people in different ways depending on where in the brain region it originates. Seizures can occur from a few seconds to a few minutes leading to greater damage if it stays longer. The most common form of seizures are unconsciousness, muscle stiffness or convulsion. But there are some other form that are harder to notice like the patient may stare into space for a few seconds. Others may cause only a few muscle twitches, a turn of the head, or a strange smell or visual disturbance that only he/she can sense. During a seizure attack the electrical activity in the brain show abnormal behavior and when it ends brain activity starts to go back to normal. A person is diagnosed as epileptic if two unprovoked seizures attack occurs (or one unprovoked seizure with the likelihood of more) that were not caused by some known and reversible physical condition like alcohol withdrawal or extremely low blood sugar. Epilepsy may start due to various reason. Some epilepsy is genetic and the way one or more genes work in the brain may provoke seizure. Structural change in the neuron cells can also introduce seizure attack. Infection, tumors and injury in the brain may causes seizure at a later time in life. Now a days strokes is one of the most common causes of new onset seizures.

Approximately 50 million people currently live with epilepsy worldwide with an estimated 2.4 million people diagnosed with epilepsy each year. Close to 80% of people with epilepsy live in low and middle income countries among whom 70% can be cured with medication [2]. The most severe threat of this disorder is its unpredictability preventing the patient from normal activities like swimming, driving, crossing a street etc. Epileptic patients can be treated with drugs or surgical procedures [3]. This severity can be subsided with proper detection and classification of epileptic patients ensuring correct treatments to avoid dire consequences, especially for those whose condition is difficult to diagnose clinically, such as those with multiple forms of epilepsy including masked and unusual forms of epilepsy [4]. If the detection can be made properly, then measures to prevent/cure seizure can be taken resulting in prevention of severe injury and even death. However a visual inspection of brain signal recording is time consuming and can be erroneous due to fatigue and other factors. Therefore a correct automatic detection is highly convenient.

There is a massive difference between the frequency dis-

tribution of epileptic and non-epileptic Electroencephalogram (EEG) signal and it can be effectively used in the detection process. In this paper, the focus is to detect epileptic patient with single channel EEG irrespective of the presence of seizure attack and create a generalized detection system. The work is done in three steps such as preprocessing, feature extraction and classification. The binary classifier considers the two normal datasets (Eye open and eye closed) as nonepileptic signal and the other three sets interictal (Signal from epileptic zone and opposite hemisphere) and ictal as epileptic signal.

II. Literature Review

In detecting epileptic seizure preictal and ictal states can be efficiently used since they indicates whether a seizure is going to occur and the actual occurrence. For decades researchers are trying to detect epilepsy with maximum sensitivity. In [1], Usman et. al. have proposed an algorithm for seizure prediction using both statistical moments and spectral features where the signal was first decomposed into frequency bands by empirical mode decomposition. They observed average prediction time of epileptic seizure as 23.48 minutes before the onset of the seizure with average sensitivity as 92.23% and specificity as 93.38%.

Entropy has been used as a measure of the complexity or uncertainty, of an EEG (Electroencephalography) signal, where the more chaotic the signal is, the higher the entropy is [5]. Many authors agree that, during a seizure, the brain activity is more synchronous than during a normal, nonseizure phase [5] [6]. Fergus et al. [6] took four entropy measurements and other features. The EEG channels were grouped into regions and top five features per region were selected based on their rank determined by the linear discriminant backward search technique. They used KNNC classifier that achieved 93% sensitivity and 94% specificity.

A different more data adaptive approach has been adopted in [4]. Accuracy of identification system may decline drastically, if signals are recorded under different conditions and have different data distribution [4]. They used transductive transfer learning mechanism for regression and binary classification including fuzzy rules for better learning and stronger modeling of uncertainty. A. T. Tzallas [7] et. al. approached a time-frequency (t-f) analysis where features including energy distribution in the t-f plane were used. They used artificial neural network for classification and achieved 89% accuracy. In [8] features including nonlinear, time, frequency and time-frequency domain extracted from fewer subbands were analyzed. A quadratic classifier achieved an overall accuracy of 98.7%.

In this study, signal in the frequency spectrum 0.5 Hz - 30 Hz is considered only since below 0.5 Hz range contains mainly device artifacts and the major important rhythmic components of the EEG signal lies in this range. Variance and frequency distribution in the brain during seizure attack and in normal state differs vastly. Hence features such as entropy, power spectral density and root mean square are used. Segmentation of the dataset into fewer seconds enhances classification accuracy greatly as seizure may occur only for a couple of seconds. Thus a 3 second trial of the original better captures the pattern and frequency distribution.

III. Data and Proposed Method

A. Data Description

The publicly-available Bonn database [9] was used in this study. The whole dataset consisted of five sets (denoted as Z, O, N, F, and S) and each of them contains 100 single-channel EEG segments of 23.6 seconds duration with sampling frequency of 173.61 Hz. Sets Z and O were taken from five healthy subjects. The subjects were awake and relaxed with their eyes open (set Z) and closed (set O), respectively. The last three sets were obtained from five epileptic patients. Sets N and F contained activity measured during seizure-free intervals, among which set F was recorded from within the epileptogenic zone and set N from the hippocampal formation of the opposite hemisphere. Set S contained only seizure activity [10]. They were further grouped into three classes – normal (sets Z and O), interictal (sets N and F), and ictal (set S) classes. The EEG for five datasets with single channel is illustrated in "Fig. 1".

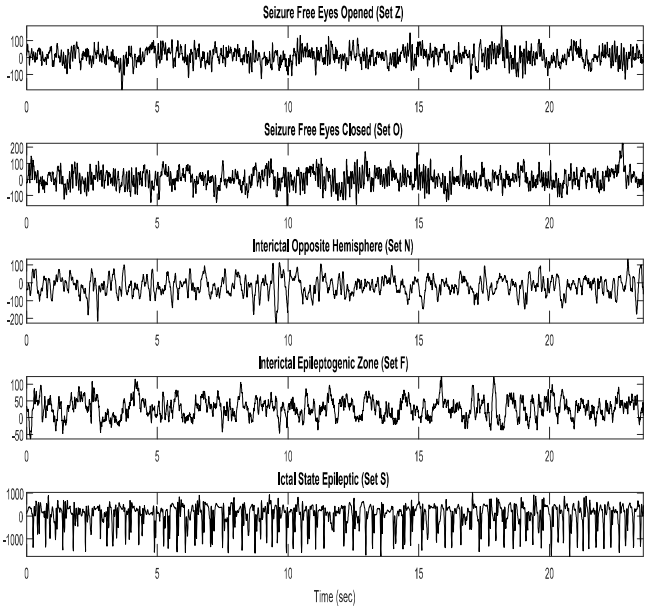


Fig. 1. Single channel EEG signal of length 23.6 sec for five datasets

B. Proposed Algorithm

This paper approaches a generalized detection system that aims to detect epileptic patient even in the interictal state. EEG signal represents neural activity thus it perfectly reflects a persons brain state. There is a distinct difference in the frequency distribution of normal and epileptic EEG. However frequency information decays vary rapidly over time, hence the window in consideration plays an important role in the detection process. Epileptic seizures may last from couple of seconds to minutes [11] and hence in this work a 3

seconds window is used to properly locate seizure features. The proposed method is depicted in "Fig. 2".

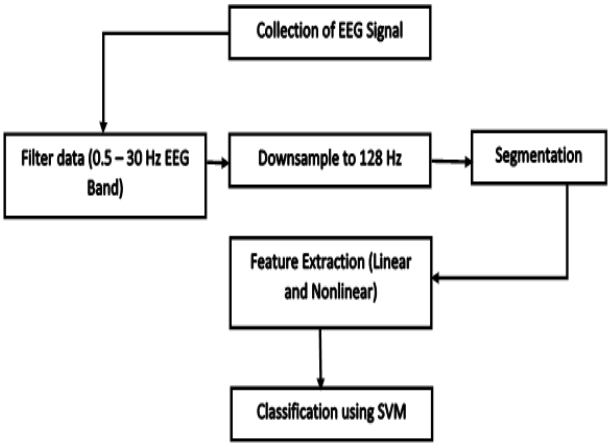


Fig. 2. Flowchart of epilepsy detection system

Steps in proposed method:

i The EEG signal is first downsampled and filtered using a zero phase band pass filter to contain only the desired rhythmic components.
ii The filtered signal is then segmented in 3 seconds trial to better capture the features.
iii Then features including approximate entropy, power of peak frequency and root mean square are calculated for each trial.
iv Finally the features are used to detect and classify the epileptic and nonepileptic signal using support vector machine.

IV. Implementation

A. Preprocessing

Each of the five sets of data are first downsampled to 128 Hz to reduce number of samples per second and to eliminate fractional numbers. It later helps perfect segmentation also. Then a zero phase butterworth band pass filter with cut off frequency at 0.5 - 30 Hz is applied. Being zero phase the filter does not affect the phase of the signal and the cut off range ensures that only the rhythmic components of the EEG signal is remained and all other physiological and electrical artifacts and noises are removed. The frequency response of the butterworth filter is best that other filters and produces cleaner cutoff in the passband and stopband. Thus this maximally flat filter works best for filtering noisy data like EEG. As epileptic seizure remains only for a couple of seconds to minutes, the data are segmented into 3t seconds trials to perfectly capture the frequency information resulting in more accurate classification. Thus the whole dataset was segmented into 7 trials for each set.

B. Feature Extraction

Entropy calculation is very popular now a days in epilepsy detection. When a seizure occurs, the brain signal becomes more synchronous than its normal state [12]. The variance right before and during a seizure attack is very high as compared to a normal brain signal. The higher the variance in a signal, the more information it carries. Thus complexity is an important feature and can be calculated using approximate entropy. In other words, approximate entropy calculates the uncertainty of a brain signal. It is unaffected by noise and robust to artifacts and outliers, hence perfect for EEG signal analysis [13]. It is described as

$$\phi^m(r) = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1} \log C_i^m(r), \quad and$$
$$ApEn(m,r,N)(u) = \phi^m(r) - \phi^m + 1(r); \quad m \geq 1 \qquad (1)$$

where $\phi^m(r)$ is relative prevalence of repetitive patterns of length m. The approximate entropy values for five datasets is illustrated in "Fig. 3".

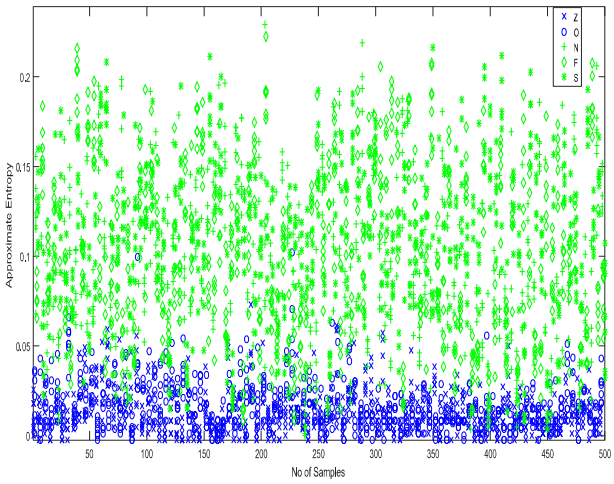


Fig. 3. Approximate entropy values of all five dataset

The power spectral density (PSD) of a signal describes the power present in the signal as a function of frequency, per unit frequency. Frequency resolution and its distribution in the signal reveals in PSD which are invisible otherwise. When a seizure occurs, a major cyclic component emerge in the EEG signal which shows itself as a dominant peak in frequency domain [14]. Thus in power spectral density representation, there is a distinct difference between the power of the peak frequency of epileptic and nonepileptic signal. Peak frequency is described as

$$f_{max} = arg\left(\frac{f_s}{N}\max_{i=0}^{N-1} P(i)\right) \qquad (2)$$

where f_s and N describe the sampling frequency and total number of samples, respectively. Dominant peak frequencies

can effectively help in the classification. In this study the power of the maximum peak in PSD is taken for each trial. The power of peak frequency in power spectral density (PSD) for five datasets is illustrated in "Fig. 4".

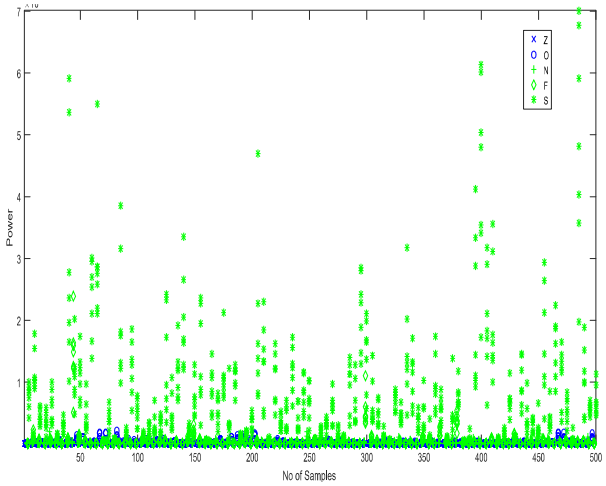


Fig. 4. Power of peak frequency in PSD for all five dataset

Root mean square measures the strength of a signal is also used in this work as a feature. During seizure brain activity is disturbed with sudden electrical discharges. This phenomena can effectively be captured with RMS value. To calculate RMS, square root of the mean square of the signal values is taken. It is described as

$$RMS = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1} x(i)^2}, \tag{3}$$

where N describes the total number of samples. The root mean square values (RMS) for five datasets is illustrated in "Fig. 5".

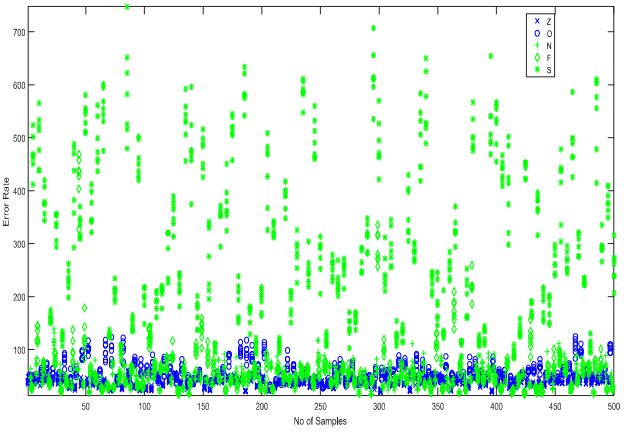


Fig. 5. Root Mean Square values for all five dataset

V. Experimental Results and Discussion

A. Classification

Support vector machine maps the training dataset in a two dimensional plane and separates these data into two parts using a hyperplane. Finally new data is categorized using this hyperplane. The hyperplane is placed in such way that maximizes distance from the data points. Hyperplane can be linear or nonlinear but since features in this work are overlapped and cannot be separated with a linear line, a nonlinear kernel is used. Kernel function transforms the dataset in a multidimensional space. Let us assume a training dataset of n points of the form

$$(\vec{x}_1, y_1), ..., (\vec{x}_n, y_n), \tag{4}$$

where y_i is either 0 or 1 indicating the label of x_i is given. There are many types of kernel functions to map the feature space in higher dimension. A second order or quadratic polynomial is best for this work since higher orders tend to overfit the training data. The polynomial kernel is defined as

$$K(\vec{X}, \vec{Z}) = (\vec{Z}^T \vec{X} + C)^n, \tag{5}$$

where n is the order of the kernel and C is a constant to trade-off the influence of higher order and lower order components. A soft margin is used to computer the classifier. The goal is to minimize this margin of the following form

$$f(\vec{w}, b) = \left[\frac{1}{n}\sum_{i=1}^{n} \max(0, 1 - y_i(w.x_i - b))\right] + \lambda||w||^2, \tag{6}$$

f is a convex function of w and b where w is the weight factor, b is bias and λ determines the trade-off between increasing the margin size and keeping x_i in the correct side of the margin.

The total dataset is divided into 2 sets: 60% for training and 40% for testing. The work is done in two sets of combination: first set Z, O and S is used where Z and O are normal and S is epileptic signal in ictal state and in the second phase all five sets are used where Z and O are labeled as normal and N, F and S are labeled as epileptic signal. A nonlinear quadratic support vector machine is used to classify the data. For combination of all five sets of data accuracy is 92.5%, sensitivity 87.5% and Specificity 100%. However the first combination of Z, O and S give an accuracy of 100%. The classification result is given in "Table. I".

TABLE I
Confusion Matrix of Proposed Method

n = 200	**Predicted: No**	**Predicted: Yes**	**Total**
Actual: No	TN = 79	FP = 1	80
Actual: Yes	FN = 11	TP = 109	120
	90	110	

The binary classifier performance at various threshold settings is determined by plotting true positive rate against false

positive rate and higher slope represents better performance in distinguishing epileptic and nonepileptic person. The receiver operating characteristic (ROC) curve is given in "Fig. 6". For classification using support vector machine, linear and nonlinear approach with different kernel functions such as RBF, polynomial, gaussian etc were tested and the quadratic classifier provided the best results. A performance comparison with other research work on the same dataset is given in "Table. II". Though accuracy degrades when interictal signal from both epileptic zone and from opposite hemisphere of the brain is added but it is noticeable that the accuracy is better with the five dataset combination than other related works. The overall accuracy is 94% with sensitivity 90.8% and specificity 98.75%.

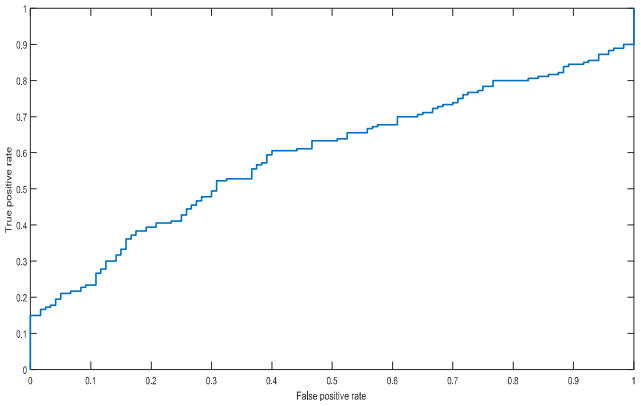


Fig. 6. Receiver Operating Characteristic (ROC) curve of Proposed Method

TABLE II
METHODS AND PERFORMANCE COMPARISON

Authors	Methods	Dataset	Accuracy
C. Yang [4] et. al.	Transductive transfer learning methods	Z,S	≥96%
		Z,F,S	≥96%
A. T. Tzallas [7] et. al.	Time–Frequency Analysis, Artificial Neural Network	Z,S	100%
		Z,O,N,F,S	89%
D. Gajic [8] et. al.	DWT, Quadratic Classifier	Non-Epileptic and Epileptic	98.7%
Proposed Method	Approximate Entropy, RMS PSD, Quadratic SVM	Z,O,S	100%
		Z,O,N,F,S	94%

VI. CONCLUSIONS

Epilepsy is a neurological disorder and in chronic state it can threaten one's life. Proper early detection can facilitate treatment. However seizures are unpredictable and presents only for a very short duration. Hence it is not practical to detect epilepsy from ictal state only. This study aims to detect epilepsy in both ictal and interictal state to enhance the efficiency of the detection system in all brain state. Observation of the feature vectors represents that the brain signal pattern in normal and epileptic patients has very different statistical distribution. Nonlinear quadratic support vector machine can effectively use these phenomena to classify the signal. While this study only aims to perform a binary classification of epileptic and nonepileptic signal, it can be applicable to predict seizure attack in advance as well. For example, entropy calculation can used to detect the pattern of a preictal brain state. In focal epilepsy, identifying the brain region of origin of epileptic seizure precisely can facilitate surgery to remove the affected brain cells. However this requires further study and analysis. In future discourse this issue will be addressed.

REFERENCES

[1] S. M. Usman, M. Usman, and S. Fong, "Epileptic seizures prediction using machine learning methods," *Computational and Mathematical Methods in Medicine*, vol. 2017, no. 7, pp. 1-10, Dec. 2017, DOI: 10.1155/2017/9074759.

[2] World Health Organization, "Epilepsy," Feb. 2018. [Online]. Available: https://www.who.int/en/news-room/fact-sheets/detail/epilepsy.

[3] M. Guenot, "Surgical treatment of epilepsy: outcome of various surgical procedures in adults and children," *Revue Neurologique*, vol. 160, no. 5, pp. S241–S250, Jul. 2004.

[4] C. Yang, Z. Deng, Kup-Sze Choi and S. Wang, "Takagi-Sugeno-Kang transfer learning fuzzy logic system for the adaptive recognition of epileptic electroencephalogram signals," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 5, pp. 1029-1094, Oct. 2016, DOI: 10.1109/TFUZZ.2015.2501438.

[5] B. R. Greene, S. Faul, W. P. Marnane, G. Lightbody, I. Korotchikova, and G. B. Boylan, "A comparison of quantitative EEG features for neonatal seizure detection," *Clinical Neurophysiology*, vol. 119, no. 6, pp. 1248-1261, Jun. 2008, DOI: 10.1016/j.clinph.2008.02.001.

[6] P. Fergus, D. Hignett, A. Hussain, D. Al-Jumeily, and K. Abdel-Aziz, "Automatic Epileptic Seizure Detection Using Scalp EEG and Advanced Artificial Intelligence Techniques," *BioMed Research International*, vol. 2015, Article ID 986736, pp. 1-17, Dec 2014, DOI: 10.1155/2015/986736.

[7] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in EEGs using time–frequency analysis," *IEEE Transactions on Information Technology in Biomedicine*, vol.13, no.5, pp. 703-10, March 2009, DOI: 10.1109/TITB.2009.2017939.

[8] D. Gajic, Z. Djurovic, J. Gligorijevic, S. D. Gennaro, and I. Savic-Gajic, "Detection of epileptiform activity in EEG signals based on timefrequency and non-linear analysis," *Frontiers in Computational Neuroscience*, vol. 9, no. 38,March 2015, DOI: 10.3389/fncom.2015.00038.

[9] EEG time series download page, Department of Epileptology, Bonn University. Sep. 23, 2014. [Online]. Available: http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3&changelang=3.

[10] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review E*, vol. 64, no. 6 Pt 1, pp. 061907, 2001.

[11] M. M. Goldenberg, "Overview of Drugs Used For Epilepsy and Seizures: Etiology, Diagnosis, and Treatment," *P & T*, vol. 35, No. 7, pp. 392-415, July 2010.

[12] H. Matsumoto, CA Marsan, "Cortical cellular phenomena in experimental epilepsy: ictal manifestations," *Experimental Neurology*, vol. 80, no. 9, pp. 305–326, April 1964.

[13] S. Pincus, "Approximate Entropy as an Irregularity Measure for Financial Data," *Econometric Reviews*, vol. 27, issue no. 4-6, pp. 329-362, 2008, DOI: 10.1080/07474930801959750.

[14] S. Sanei, and J. Chambers, *EEG Signal Processing*, New York, NY: John Wiley & Sons, 2013.

Identification of Autoregressive Systems in the Presence of Additive Noise Using the Matrix Pencil Method

Md. Shamim Hussain
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh
snirjhar@gmail.com

***Abstract*—In this paper, a novel method for estimation of the parameters of autoregressive (AR) systems in the presence of additive noise is presented. This method models the autocorrelation function (ACF) of the noisy AR signal as a sum of non-increasing complex exponentials and then estimates its parameters using the matrix pencil method. We improve upon this basic method by fitting the noisy AR signal to a higher order AR model and using its parameters to calculate a parametric ACF of the noisy signal which makes the pole extraction process using the matrix pencil method stable and more accurate. We show that this method can accurately estimate the parameters of an AR system in noise in a single pass even at a very low (-5dB) signal-to-noise ratio (SNR).**

***Index Terms*—autoregressive system, parameter estimation, matrix pencil method, low signal-to-noise ratio**

I. Introduction

Signal modeling using autoregressive systems is ubiquitous in many different fields of science and engineering such as speech analysis, economics, radar and sonar systems, seismology and biomedical signal processing. While the estimation of parameters of an AR system in absence of additive noise is straightforward with well-known methods like the Yule-Walker method or the Burg algorithm [1], in practice, the desired signal is often corrupted by observational noise. The presence of noise causes biased estimation of model parameters when the methods devised for the noise-free case are used without any modification [2]. Also, this estimation problem becomes exceedingly hard at very low signal-to-noise ratios (SNR) [3]. So, more sophisticated methods are required to estimate the AR model parameters in the presence of a high level of noise. Several methods have been proposed for parameter estimation of noisy AR systems. As an AR system in the presence of additive noise can be modeled as an ARMA system [4], a commonly used method is to use an ARMA parameter extraction method, like prediction errors methods [5] to estimate the parameters of the AR system. Another common approach to account for observational noise is to adopt a noise compensation technique such as the noise compensated lattice filter approach [6]. However, noise compensation based methods usually require an *a priori* knowledge of the noise variance. The conventional Yule-Walker method can be modified to devise the high-order Yule-Walker method for noisy AR systems which utilizes the autocorrelation coefficients at higher lags. However, this approach is prone to singularity problem of the autocorrelation matrix [7]. Among other methods, the improved least squares method with no prefiltering (ILSNP) of noisy observations reported in [8] provides accurate estimates of the AR parameters at a high SNR.

All of the methods mentioned above work well down to some positive value of the SNR. At very low SNRs, such as close to 0dB or negative SNRs, damped sinusoidal modeling of the autocorrelation function (ACF) is shown to produce a reliable estimation of the AR parameters [3], [9]. In [3] a direct approach based on the scanning of the z-plane is proposed which is very computationally expensive. In a later work, a modified empirical mode decomposition (EMD) based approach [9] is proposed to reduce the computational cost. In both cases, the poles are estimated sequentially one by one which requires that we assume the damped sinusoidal components are either orthonormal (for the scanning approach) or corresponds to separate intrinsic mode functions (for the EMD approach). Both of these approaches are heuristic in the sense that they are based on empirical observations. Although the EMD based approach does not require scanning the z-plane exhaustively, it is still computationally expensive due to the iterative nature of EMD. Both these approaches have been devised for real-valued signals only and their extension for complex-valued signals have not been explored. In this paper, the basic concept of modeling the ACF as a non-increasing sum of complex exponentials (or complex sinusoids) is adopted, while the shortcomings of the previous methods have been addressed. Several very efficient and reliable methods of estimating parameters of exponential sums have been proposed in the literature [10]. In this paper, the reliable matrix pencil [10], [11] method of estimating parameters of exponential sums is used to estimate the poles from the ACF simultaneously, with high precision and at a relatively much lower computational cost. However, this method requires relatively good estimates of the autocorrelation coefficients at higher lags, and can sometimes suffer from numerical instability or prediction errors if a naive approach is used

for estimating the ACF. So, we propose a more accurate method of estimating the ACF by first over-parameterizing the AR signal, i.e. modeling it as a high order AR signal and using the estimated parameters of the high order AR model to compute a parametric ACF which provides a much more accurate low-variance estimate of the ACF at higher lags and thus also solves the instability problem. The proposed method is principled, computationally efficient as it only uses a single pass to estimate the parameters and also directly applicable to complex signals. It successfully solves the shortcomings of [3] and [9] while also producing more accurate estimates at very low SNR levels (as low as -5dB).

II. Problem Formulation

The observed signal $x(n)$ is modeled as a superposition of a pth order autoregressive signal $y(n)$ and the observational noise $v(n)$.

$$x(n) = y(n) + v(n) \tag{1}$$

For the simplest case, we will assume that $v(n)$ is a Gaussian white noise with its samples being drawn independently from the normal distribution $\mathcal{N}(0, \sigma_v^2)$. The AR signal $y(n)$ can by modeled by the difference equation

$$y(n) = -\sum_{k=1}^{p} a_k y(n-k) + u(n) \tag{2}$$

Where, $a_k, k = 1, 2, ..., p$ are the AR parameters and $u(n)$ is a Gaussian white noise with its samples being drawn from the distribution $\mathcal{N}(0, \sigma_u^2)$. It is assumed that the input noise $u(n)$ and the observational noise $v(n)$ are independent of each other.

Another way of modeling this AR signal is as the output of a causal, infinite impulse response (IIR) system with p poles and the noise $u(n)$ as its input.

$$y(n) = h(n) * u(n) \tag{3}$$

Where the impulse response of the system $h(n)$ has the difference equation representation

$$h(n) = \begin{cases} -\sum_{k=1}^{p} a_k h(n-k) + \delta(n) & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases} \tag{4}$$

In the z-domain $h(n)$ has the representation

$$H(z) = \frac{1}{1 + \sum_{k=1}^{p} a_k z^{-k}} \tag{5}$$

Or in terms of poles as

$$H(z) = \frac{1}{\prod_{k=1}^{p} (1 - z_k z^{-1})} \tag{6}$$

Here, $z_k, k = 1, 2, ..., p$ are the poles of the AR system and are the solutions to the polynomial equation $1 + \sum_{k=1}^{p} a_k z^{-k} = 0$. The region of convergence (ROC) of $H(z)$ is $|z| > |z_m|$ where z_m is the outermost pole i.e. $|z_m| \geq |z_i|, i = 1, 2, ..., p$. Since, the system is assumed to be stable, the ROC must include the unit circle in the z-plane and all its poles must lie within the unit circle.

The problem of identification of a noisy AR system involves estimation of the AR parameters $(a_k, k = 1, 2, ..., p)$ or the poles $(z_k, k = 1, 2, ..., p)$ from the noisy signal $x(n)$. Additionally, we may be interested in estimating the variances of the input noise (σ_u^2) and the observational noise (σ_v^2).

III. Proposed Method

The autocorrelation function (ACF) of the noise-free AR signal is

$$r_y(m) = \mathbb{E}[y(n)y^*(n-m)] = \sigma_u^2 r_h(m) \tag{7}$$

Where, the autocorrelation of the impulse response $h(n)$ is

$$r_h(m) = h(m) * h^*(-m) = \sum_{n=-\infty}^{\infty} h^*(n)h(n+m) \tag{8}$$

A detailed derivation of this expression can be found in [12]. Since, the input noise $u(n)$ and the observational noise $v(n)$ are assumed to be independent, the ACF of the noisy AR signal is simply

$$r_x(m) = \mathbb{E}[x(n)x^*(n-m)] = r_y(m) + \sigma_v^2 \delta(m) \tag{9}$$

$$= \sigma_u^2 r_h(m) + \sigma_v^2 \delta(m) \tag{10}$$

So, the ACF of $x(n)$, for lags other than zero ($m \neq 0$) is (ideally) a scaled version of the autocorrelation of the impulse response. However, since we have to use finite length samples of $x(n)$ in practice, the estimated ACF will be noisy.

Our goal is to model $r_x(m)$ for $m > 0$ as a sum of non-increasing exponentials and then use the Matrix Pencil method to estimate its poles. We will make use of the following lemmas in devising our method.

Lemma 1. *Any causal and bounded sequence that has a proper function representation in the z-domain and does not have any repeated poles, can be expressed as a sum of non-increasing exponentials.*

Proof. Let $h(n)$ be a causal sequence such that

$$h(n) = \begin{cases} -\sum_{k=1}^{p} a_k h(n-k) + \sum_{l=0}^{q} b_l \delta(n-l) & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases} \tag{11}$$

Where $q < p$. Then, it has a proper function representation in the z-domain

$$H(z) = \frac{\sum_{l=0}^{q} b_l z^{-l}}{1 + \sum_{k=1}^{p} a_k z^{-k}} \tag{12}$$

$H(z)$ can be written in terms of poles $z_k, k = 1, 2, ..., p$ as

$$H(z) = \frac{\sum_{l=0}^{q} b_l z^{-l}}{\prod_{k=1}^{p} (1 - z_k z^{-1})} \tag{13}$$

Assuming no repetition of poles, by partial fraction decomposition we can write

$$H(z) = \sum_{k=1}^{p} \frac{C_k}{1 - z_k z^{-1}} \tag{14}$$

Where,

$$C_k = \frac{\sum_{l=0}^{q} b_l z_k^{-l}}{\prod_{\substack{k'=1 \\ k' \neq k}}^{p} (1 - z_{k'} z_k^{-1})} \tag{15}$$

By taking inverse z-transform we get

$$h(n) = \begin{cases} \sum_{k=1}^{p} C_k z_k^n & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases} \tag{16}$$

Lemma 2. *If a causal sequence can be expressed as a sum of non-increasing exponentials, the causal part (i.e. for non-negative lags) of its autocorrelation function can also be expressed as a sum of non-increasing exponentials with the same set of poles as the original sequence.*

Proof. The causal sequence under consideration

$$h(n) = \begin{cases} \sum_{k=1}^{p} C_k z_k^n & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases} \tag{17}$$

The causal part (lag $m \geq 0$) of the ACF of $h(n)$

$$\begin{aligned} r_h(m) &= \sum_{n=0}^{\infty} h^*(n) h(n+m) \\ &= \sum_{n=0}^{\infty} \sum_{l=1}^{p} C_l^* (z_l^*)^n \sum_{k=1}^{p} C_k z_k^{n+m} \\ &= \sum_{k=1}^{p} z_k^m \sum_{l=1}^{p} C_l^* C_k \sum_{n=0}^{\infty} (z_l^* z_k)^n \\ &= \sum_{k=1}^{p} z_k^m \sum_{l=1}^{p} \frac{C_l^* C_k}{1 - z_l^* z_k} \\ &= \sum_{k=1}^{p} \beta_k z_k^m \end{aligned} \tag{18}$$

Where,

$$\beta_k = \sum_{l=1}^{p} \frac{C_l^* C_k}{1 - z_l^* z_k} \tag{19}$$

Lemma 3. *If the causal sequence $h(n)$ can be expressed as a sum of non-increasing exponentials, its advanced and truncated form $h(n+l)\mathcal{H}(n);\ l > 0$ can also be expressed as a sum of non-increasing exponentials with the same set of poles, where $\mathcal{H}(n)$ is the heaviside step function.*

Proof. Let,

$$h(n) = \begin{cases} \sum_{k=1}^{p} C_k z_k^n & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases} \tag{20}$$

$$\text{Then, } h(n+l)\mathcal{H}(n) = \begin{cases} \sum_{k=1}^{p} C_k z_k^{n+l} & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases}$$

$$= \begin{cases} \sum_{k=1}^{p} D_k z_k^n & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases} \tag{21}$$

Where,

$$D_k = C_k z_k^l \quad \text{for } k = 1, 2, ..., p \tag{22}$$

A. Exponential Sum Model of the Autocorrelation Function

From (5) we see that $h(n)$ has a proper function representation in the z-domain. If we assume that the AR signal does not have any repeated poles, according to lemma 1, it can be expressed as a sum of non-increasing exponentials.

$$h(n) = \begin{cases} \sum_{k=1}^{p} C_k z_k^n & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases} \tag{23}$$

Now, according to lemma 2, $r_h(m)\mathcal{H}(m)$ can also be expressed as a sum of non-increasing exponentials, and according to lemma 3, so can be $r_h(m+1)\mathcal{H}(m)$

$$r_h(m)\mathcal{H}(m) = \sum_{k=1}^{p} \beta_k z_k^m \qquad \text{for } m \geq 0 \tag{24}$$

$$r_h(m+1)\mathcal{H}(m) = \sum_{k=1}^{p} \gamma_k z_k^m \qquad \text{for } m \geq 0 \tag{25}$$

Where,

$$\gamma_k = \beta_k z_k \tag{26}$$

From (10) we see that the ACF of the noisy AR signal $x(n)$ for lags greater than zero can be expressed as an exponential sum as well.

$$\begin{aligned} r_x(m+1)\mathcal{H}(m) &= \sigma_u^2 r_h(m+1)\mathcal{H}(m) \\ &= \sigma_u^2 \sum_{k=1}^{p} \gamma_k z_k^m \qquad \text{for } m \geq 0 \\ &= \sum_{k=1}^{p} c_k z_k^m \qquad \text{for } m \geq 0 \end{aligned} \tag{27}$$

Where,

$$c_k = \sigma_u^2 \gamma_k \tag{28}$$

We can estimate the poles z_k and the coefficients c_k (for $k = 1, 2, ..., p$) from this sum of exponential representation of $r_x(m+1)\mathcal{H}(m)$ by a method such as the matrix pencil method.

From (7), (24) and (26), the causal part of ACF of the noise-free AR signal $y(n)$ is,

$$\begin{aligned} r_y(m)\mathcal{H}(m) &= \sum_{k=1}^{p} \sigma_u^2 \beta_k z_k^m \qquad \text{for } m \geq 0 \\ &= \sum_{k=1}^{p} \sigma_u^2 \gamma_k z_k^{m-1} \qquad \text{for } m \geq 0 \\ &= \sum_{k=1}^{p} c_k z_k^{m-1} \qquad \text{for } m \geq 0 \end{aligned} \tag{29}$$

So, after estimating c_k and z_k we can estimate $r_y(m)$, and the variance of input noise from the first Yule-Walker equation as

$$\sigma_u^2 = r_y(0) + \sum_{k=1}^{p} a_k r_y^*(k) \tag{30}$$

The variance of observational noise can also be estimated from (9) as

$$\sigma_v^2 = r_x(0) - r_y(0) \tag{31}$$

B. Matrix Pencil Method

The matrix pencil method is a principled method for estimating the parameters of an exponential sum sequence, even in the presence of additive noise. For our purpose, we wish to determine the parameters (z_k, c_k) of the exponential sum representation of $r_x(m+1)\mathcal{H}(m)$, as presented in (27). Due to the finite length of the observed noisy signal $x(n)$, the estimated $r_x(m+1)\mathcal{H}(m)$ is noisy. The matrix pencil method is well-equipped to take additive noise into consideration. It directly estimates all poles simultaneously, in a single pass.

Two matrix pencil methods have been proposed in [10] - one based on QR decomposition and the other based on Singular Value Decomposition (SVD) of the Hankel matrix. We found that both of them performed satisfactorily for our task. But, we will discuss the rotationally invariant (ESPRIT) method based on SVD of the Hankel matrix and present our experimental results for the same. For notational convenience, let

$$s(m) = r_x(m+1)\mathcal{H}(m) \tag{32}$$

First, the rectangular $(M-L)\times(L+1)$ Hankel matrix $\boldsymbol{H}$ is formed using M samples of $s(m)$

$$\boldsymbol{H} = \begin{bmatrix} s(0) & s(1) & \dots & s(L) \\ s(1) & s(2) & \dots & s(L+1) \\ \vdots & \vdots & \ddots & \vdots \\ s(M-L-1) & s(M-L) & \dots & s(M-1) \end{bmatrix} \tag{33}$$

For efficient noise filtering, the parameter L, is chosen between $M/3$ to $M/2$. Next, an SVD of the Hankel matrix is carried out as,

$$\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H \tag{34}$$

Where, $\boldsymbol{U}$ and $\boldsymbol{V}$ are two unitary matrices containing the left and right singular vectors of $\boldsymbol{H}$ respectively and $\boldsymbol{\Sigma}$ is a diagonal matrix containing the singular values. $\boldsymbol{V}^H$ is the conjugate transpose of $\boldsymbol{V}$.

Algorithm 1 Matrix Pencil Method (with Parametric ACF) for Estimating the Parameters of A Noisy AR Signal

Input: Noisy AR signal $x(n)$; for $n = 0, 1, ..., N-1$

Output: Estimated AR parameters $\hat{a}_k$ for $k = 1, 2, ..., p$; input noise variance $\hat{\sigma}_u^2$ and observational noise variance $\hat{\sigma}_v^2$

1. Fit an AR(t) model ($t \gg p$) to $x(n)$ and calculate its parameters using the Burg method. Then calculate the impulse response $\hat{h}^{(t)}(n)$ for $n = 0, 1, ..., P$ and the parametric ACF $\hat{r}_x^{(t)}(m)$ for $m = 0, 1, ..., M$ using (39) and (40) respectively.
2. After estimating the ACF $\hat{r}_x(m)$, calculate the parameters $\hat{z}_k$ and $\hat{c}_k$ (for $k = 1, 2, ..., p$) of the exponential sum representation of $\hat{s}(m) = \hat{r}_x(m+1)\mathcal{H}(m)$ (for $m = 0, 1, 2, ..., M-1$) using the matrix pencil method.
3. Calculate the AR parameters $\hat{a}_k$ from the estimated poles $\hat{z}_k$ (for $k = 1, 2, ..., p$).
4. Estimate the ACF of the noise-free AR signal $\hat{r}_y(m)$ from (29) and use it to calculate the input noise variance $\hat{\sigma}_u^2$ using (30) and the observational noise variance $\hat{\sigma}_v^2$ using (31).

To "denoise" the Hankel matrix, the most dominant p singular values and the corresponding right singular vectors $\boldsymbol{v_1}, \boldsymbol{v_2}, ..., \boldsymbol{v_p}$ are chosen. Here, p is the number of poles of the AR system which may be known *a priori* or can be estimated at this stage from the singular values of the Hankel matrix, based on some criterion, such as the one presented in [11]. The denoised version of $\boldsymbol{V}$ is estimated from the dominant right singular vectors as

$$\boldsymbol{V}' = \begin{bmatrix} \boldsymbol{v_1} & \boldsymbol{v_2} & \dots & \boldsymbol{v_p} \end{bmatrix} \tag{35}$$

Two matrices $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are formed from $\boldsymbol{V}'$ by deleting its last and first rows respectively. Then the poles z_k ($k = 1, 2, ..., p$) are estimated by computing the eigenvalues of the matrix

$$\boldsymbol{F} = \{\boldsymbol{V}_1^H\}^\dagger \boldsymbol{V}_2^H \tag{36}$$

Here, $\{\boldsymbol{V}_1^H\}^\dagger$ is the Moore-Penrose pseudoinverse of $\boldsymbol{V}_1^H$. Once the poles have been calculated, the coefficients c_k ($k = 1, 2, ..., p$) can be determined from the least squares solution of the overdetermined linear Vandermonde-type system

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_p \\ z_1^2 & z_2^2 & \dots & z_p^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{M-1} & z_2^{M-1} & \dots & z_p^{M-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_p \end{bmatrix} = \begin{bmatrix} s(0) \\ s(1) \\ s(2) \\ \vdots \\ s(M-1) \end{bmatrix} \tag{37}$$

In practice, several steps of this method can be optimized for computational efficiency as described in details in [13]. A rigorous mathematical background of the matrix pencil method can be found in [11] and [10].

C. Parametric Estimation of the Autocorrelation Function

In the previous subsections, we have described how we can estimate the parameters of the noisy AR signal from its

autocorrelation function. A common approach of estimating the ACF is

$$\hat{r}_x(m) = \frac{1}{N} \sum_{n=0}^{N-1-k} x^*(n)x(n+m); \text{ for } m = 0, 1, ..., M \tag{38}$$

We could use this estimate to form $\hat{s}(m) = \hat{r}_x(m+1)\mathcal{H}(m)$ for $m = 0, 1, ..., M-1$, which is the approach adopted in [3] and [9]. It works well for the proposed method at high SNRs as well. However, to effectively denoise the Hankel matrix at very low SNRs and to precisely calculate the poles using the matrix pencil method we need to utilize the coefficients at higher lags (at least 6 to 8 times the number of poles). But, here, this simplistic approach to estimating the ACF often falls short, because the variance of the estimated coefficients quickly increases with increasing lag. Also, it could be hard to choose the number of coefficients (M) to use for the estimation task. If choosing too few coefficients could lead to numerical instability, and choosing too many could lead to inaccurate estimation due to high variance at greater lags.

The above approach is non-parametric and thus, it does not utilize the *a priori* knowledge about nature of the noisy AR signal. We devise a more accurate approach to estimating the ACF taking the nature of the signal under consideration. An AR(p) signal with additive noise can be modeled as an AR(p)/MA(p) (Autoregressive Moving Average) signal [4]. Again, an AR(p)/MA(p) signal can be modeled as an AR signal of higher order t [12], where $t \gg p$. We exploit the fact that there exist highly efficient and numerically stable algorithms such as the Burg method for estimating AR model parameters. So, first we model the noisy AR(p) signal as a higher order (noise-free) AR(t) signal and estimate its parameters $\hat{a}_k^{(t)}, k = 1, 2, ..., t$ and $\{\hat{\sigma}_u^{(t)}\}^2$ by an efficient method such as the Burg method. Then we can iteratively calculate the impulse response of the AR(t) signal for the first P samples from the difference equation

$$\hat{h}^{(t)}(n) = \begin{cases} -\sum_{k=1}^{r} \hat{a}_k^{(t)} \hat{h}(n-k) + \delta(n) & \text{for } P > n \geq 0 \\ 0 & \text{for } n < 0 \end{cases} \tag{39}$$

Then we can calculate the parametric ACF as

$$\hat{r}_x^{(t)}(m) = \left\{\hat{\sigma}_u^{(t)}\right\}^2 \sum_{n=0}^{P-1-k} \left\{\hat{h}^{(t)}(n)\right\}^* \hat{h}^{(t)}(n+m) \quad \text{for } m = 0, 1, ..., M \tag{40}$$

In practice, we found that $t > 3p$ works well. If the impulse response decays to ignorable values within the evaluated range we can choose similar values for P and M. Otherwise $P > 2M$ works well to avoid the windowing effect.

IV. Results

In this section, we evaluate the performance of the proposed method in estimating the parameters of noisy AR signals and compare them with other methods - more specifically the high-order Yule-Walker (HOYW), iterative least squares method (ILSNP) with no prefiltering [8] and the EMD method [9]. We will present results for three AR systems with the following sets of parameters

System 1: $\mathbf{a} = [-0.8600 \;\; 1.0494 \;\; -0.6680 \;\; 0.9592 \;\; -0.7563 \;\; 0.5656]^T$

System 2: $\mathbf{a} = [-2.7607 \;\; 3.8106 \;\; -2.6535 \;\; 0.9238]^T$

System 3: $\mathbf{a} = [0.2080 \;\; -0.6749 \;\; 0.1991 \;\; 0.9238]^T$

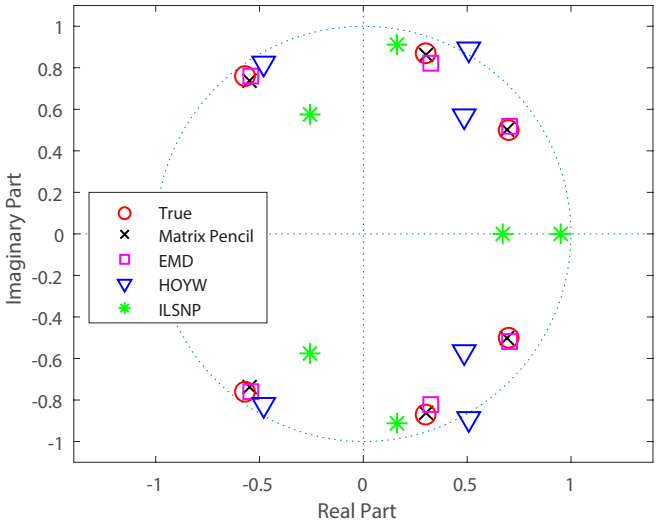


Fig. 1. Estimate of the poles of AR System 1 at SNR=-5dB by different methods

TABLE I
Performance Comparison of Different Methods for AR System 2 at SNRs of 10dB and -5dB (Under Each Estimated Parameter First SDMs and Then SDTs Are Shown)

Parameter	SNR (dB)	Proposed Method	EMD Method	ILSNP Method	HOYW Method
$a_1 = -2.7607$	10	-2.7569 ± 0.0090 ± 0.0097	-2.6997 ± 0.0105 ± 0.0619	-0.5619 ± 0.0268 ± 2.1989	-1.0123 ± 15.5852 ± 15.4881
	-5	-2.7525 ± 0.0126 ± 0.0150	-2.6799 ± 0.0514 ± 0.0954	-0.3746 ± 0.5194 ± 2.4406	-1.1215 ± 4.0272 ± 4.3012
$a_2 = 3.8106$	10	3.7993 ± 0.0189 ± 0.0221	3.6743 ± 0.0240 ± 0.2073	0.0954 ± 0.0187 ± 3.7153	-0.9818 ± 41.2368 ± 40.9992
	-5	3.7825 ± 0.0259 ± 0.0382	3.6303 ± 0.1036 ± 0.1384	0.0998 ± 0.4026 ± 3.7320	0.0449 ± 2.2651 ± 4.3799
$a_3 = -2.6535$	10	-2.6410 ± 0.0179 ± 0.0218	-2.5199 ± 0.0238 ± 0.1357	0.3728 ± 0.0201 ± 3.0264	2.2789 ± 41.9589 ± 41.7237
	-5	-2.6192 ± 0.0245 ± 0.0421	-2.4762 ± 0.0979 ± 0.2020	0.2368 ± 0.3281 ± 2.9084	0.9831 ± 4.6083 ± 5.8250
$a_4 = 0.9232$	10	0.9175 ± 0.0077 ± 0.0099	0.8658 ± 0.0099 ± 0.0588	0.2449 ± 0.0249 ± 0.6793	-1.3501 ± 18.9944 ± 18.8928
	-5	0.9057 ± 0.0104 ± 0.0209	0.0389 ± 0.0389 ± 0.0848	0.4659 ± 0.4360 ± 0.6285	0.6105 ± 5.6015 ± 5.7399
RMSE(%)	10	0.63	3.80	95.42	1137.89
	-5	1.15	5.68	96.23	183.55

TABLE II
PERFORMANCE COMPARISON OF DIFFERENT METHODS FOR AR SYSTEM 3 AT SNRS OF 10DB AND -5DB (UNDER EACH ESTIMATED PARAMETER FIRST SDMS AND THEN SDTS ARE SHOWN)

Parameter	SNR (dB)	Proposed Method	EMD Method	ILSNP Method	HOYW Method
$a_1 = 0.2080$	10	0.2081 ± 0.0071 ± 0.0071	0.2069 0 ± 0.0011	0.2624 ± 0.0078 ± 0.0549	0.2102 ± 0.0036 ± 0.0042
	-5	0.2122 ± 0.0127 ± 0.0133	0.2338 ± 0.0219 ± 0.0337	0.2738 ± 0.1194 ± 0.1350	0.1910 ± 0.0710 ± 0.0722
$a_2 = -0.6749$	10	-0.6722 ± 0.0077 ± 0.0081	-0.6771 0 ± 0.0022	-0.7187 ± 0.0078 ± 0.0583	-0.6774 ± 0.0043 ± 0.0049
	-5	-0.6663 ± 0.0143 ± 0.0167	-0.6635 ± 0.0171 ± 0.0204	0.5664 ± 0.3884 ± 0.3986	-0.7012 ± 0.0690 ± 0.0730
$a_3 = 0.1991$	10	0.1973 ± 0.0069 ± 0.0072	0.1987 0 ± 0.0004	0.1693 ± 0.0065 ± 0.0304	0.1894 ± 0.0061 ± 0.0114
	-5	0.1909 ± 0.0135 ± 0.0158	0.1676 ± 0.0264 ± 0.0409	0.1557 ± 0.1342 ± 0.1395	0.2035 ± 0.1462 ± 0.1444
$a_4 = 0.9238$	10	0.9153 ± 0.0057 ± 0.0103	0.9224 0 ± 0.0014	0.9595 ± 0.0445 ± 0.0567	0.9145 ± 0.0064 ± 0.0112
	-5	0.8997 ± 0.0110 ± 0.0265	0.8925 ± 0.0374 ± 0.0484	0.8257 ± 0.5394 ± 0.5416	0.9249 ± 0.1343 ± 0.1326
RMSE(%)	10	1.40	0.18	6.64	1.12
	-5	3.18	4.82	45.26	14.31

For each system, data were generated using their corresponding difference equations. In all simulations we used $N = 4000$ data samples. Observational white noise was added to the AR signals according to the SNR under consideration. The noisy AR signal was fitted to a high order $t = 32$ pole AR(t) signal for the estimation of parametric ACF. In our experiments, we evaluated both the impulse response and the parametric ACF for $P = M = 128$ samples/lags. For the formation of the Hankel matrix, $L = 32(\approx M/3)$ was used.

First, we present the estimated poles for System 1 by different methods at SNR=-5dB in Fig. 1. At this low SNR, both HOYW and ILSNP methods fail to satisfactory estimates of the poles. But both the EMD and the proposed matrix pencil method are able to estimate the poles with satisfactory accuracy, even at such low level of SNR. But unlike the EMD method, the Matrix Pencil Method is non-iterative and computationally much more efficient.

Next, we present comparative results for different methods in estimating AR parameters at SNR levels of 10dB and -5dB. The results for System 2 and 3 are presented in Table I and Table II respectively. These results were produced by running 1000 simulations of the corresponding AR systems. Under each estimated parameter, first standard deviation with respect to mean (SDM) and then standard deviation with respect to true value (SDT) are shown. For an overall measure of accuracy in estimation the root mean squared error (RMSE) of the estimates is also shown in the tables. We used the definition of RMSE given in [9].

The results presented in Table I and II show that the proposed matrix pencil based method achieves comparable results to the EMD based method, or even outperforms the EMD method at -5dB SNR. The values of SDM and SDT show that the proposed method consistently achieves accurate estimations. Thus, the results suggest that the proposed method is able to achieve accurate results while being computationally efficient and also overcomes the shortcomings of the EMD based method.

V. CONCLUSION

In this paper, a new principled approach for estimating the parameters of an AR signal with additive noise is presented. We showed that the autocorrelation function of the noisy AR system can be modeled as a non-increasing exponential sum and then utilized the matrix pencil method to estimate its poles. The use of parametric ACF facilitates accurate estimation of the correlation coefficients at greater lags which improves the accuracy and the numerical stability of the matrix pencil method at a low SNR. We presented a rigorous mathematical derivation of the proposed method and also demonstrated its effectiveness in AR parameter estimation by means of extensive simulations. The simulation results show that the proposed method is able to accurately estimate AR parameters even at very low levels of SNR.

REFERENCES

[1] J. P. Burg, "Maximum entropy spectral analysis," in *37th Annual International Meeting, Soc. of Explor. Geophys., Oklahoma City, Okla.*, 1967.

[2] S. Kay, "The effects of noise on the autoregressive spectral estimator," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 27, no. 5, pp. 478–485, 1979.

[3] M. K. Hasan, S. A. Fattah, and M. R. Khan, "Identification of noisy AR systems using damped sinusoidal model of autocorrelation function," *IEEE Signal Processing Letters*, vol. 10, no. 6, pp. 157–160, 2003.

[4] M. Pagano, "Estimation of models of autoregressive signal plus white noise," *The annals of Statistics*, pp. 99–108, 1974.

[5] A. Nehorai and P. Stoica, "Adaptive algorithms for constrained ARMA signals in the presence of noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 8, pp. 1282–1291, 1988.

[6] W. Sun and T. Yahagi, "Adaptive overfitting lattice filter for parameter estimation of time-varying systems," *Signal processing*, vol. 29, no. 1, pp. 45–56, 1992.

[7] L. Vergara-Dominguez, "New insights into the high-order Yule-Walker equations," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 9, pp. 1649–1651, 1990.

[8] W. X. Zheng, "A least-squares based method for autoregressive signals in the presence of noise," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 46, no. 1, pp. 81–85, 1999.

[9] M. K. Hasan, M. S. Apu, and M. K. I. Molla, "A robust method for parameter estimation of AR systems using empirical mode decomposition," *Signal, image and video processing*, vol. 4, no. 4, pp. 451–461, 2010.

[10] D. Potts and M. Tasche, "Parameter estimation for nonincreasing exponential sums by Prony-like methods," *Linear Algebra and its Applications*, vol. 439, no. 4, pp. 1024–1039, 2013.

[11] T. K. Sarkar and O. Pereira, "Using the matrix pencil method to estimate the parameters of a sum of complex exponentials," *IEEE Antennas and Propagation Magazine*, vol. 37, no. 1, pp. 48–55, 1995.

[12] M. H. Hayes, "Statistical digital signal processing and modeling." John Wiley & Sons, 2009, ch. 3.

[13] H. Ikeno, "MXPFIT: A library for finding optimal multi-exponential approximations," *Computer Physics Communications*, vol. 230, pp. 135–144, 2018.

Detection Performance Analysis of Compressive Sensing in Cooperative Cognitive Radio Network

Teethiya Datta[1], Shohely Tasnim Anindo[2] and Sk. Shariful Alam[3]
[123]Khulna University of Engineering & Technology, Khulna-9203, Bangladesh
TeethiyaDatta@gmail.com[1],anindyodip@gmail.com[2],ssa1996@gmail.com[3]

***Abstract*—Compressive Sensing(CS) is a new prototype in signal processing. CS requires significant computations and memory spaces because whole wideband spectrum is approximated to find an opportunity for a Cognitive Radio (CR) usage. In this paper, a CR receiver wideband sensing unit is presented where a notable portion of the wideband spectrum is approximated through compressive sensing. This structure reduces computational burden and improves the detection performance. In this paper, we have presented a spectrum sensing approaches where compressive sensing will be accomplished under cooperative cognitive radio network as well as non-cooperative approach in terms of receiver operating characteristics (ROC). The investigation has been carried out under Additive White Gaussian Noise (AWGN), Rayleigh, Rician, and Nakagami channel considering the fading effect. Finally, the most appropriate arrangement for spectrum hole detection has been determined through step by step comparative analysis.**

***Index Terms*—cognitive radio, compressive sensing, cooperation, fading models, decision fusion rule.**

I. INTRODUCTION

The Radio spectrum is a limited, government-led natural resource. In accordance with the current policy, all frequency bands for a particular geographical location are completely and statistically licensed for wireless networks and each system has assigned a frequency band to them. Due to the improvement of a vast number of wireless devices, it has become a challenge because all currently under developed wireless systems are not provided with sufficient spectrum. So it is necessary to identify the available vacant frequency bands across multiple dimensions and efficiently utilize them without having interference with the Primary Users (PU)s of that spectrum. The task can be accomplished by cognitive radio which aims to detect the vacant spectrum holes with high probability of detection. Several methods are suggested for identifying the presence of primary users' signal transmissions.

The entire operation of CR relies on spectrum sensing which is the first and most important task of CR. Sensing can be performed either on narrowband or wideband levels. When there is enough information about the primary users signaling, center frequency and bandwidth, narrowband sensing can be performed.

On the other hand, wideband spectrum sensing obtains better utilization of the frequency spectrum, when the occupancy details of the primary users are unknown to cognitive network users. Conventional wideband spectrum sensing techniques are based on high sampling frequency. Because functioning at or above Nyquist rates are challenging, as it can cause more computational complexity. The design and implementation of wideband sensing systems are both difficult due to high implementation complexity and large energy consumption from high-rate analog-to-digital converters (ADC). There are numerous wideband sensing techniques that make use of sub-Nyquist sampling commonly known as Compressive Sensing (CS). Tian and Giannakis firstly introduced compressive sensing theory to sense wideband spectrum in [1]. This technique used fewer samples closer to the information rate, rather than the inverse of the bandwidth, to perform wideband spectrum sensing [2]. After reconstruction of the wideband spectrum, wavelet-based edge detection was used to detect spectral opportunities across wideband spectrum [3]. In this paper, Compressive Sensing is presented under cooperative cognitive radio network that enables the efficient operation of both the PU and the CR networks. It will be easier to get rid of the problem of multipath fading and hidden node problem if spectrum sensing is performed with the assistance of multiple CR nodes.

II. METHODOLOGY

A. Compressive Sensing

In both scholarly community and industry, Sub-Nyquist techniques are drawing more and more recognition. For the purpose of detecting spectral opportunities in the wideband, sub-Nyquist wideband sensing alludes to the procedure of obtaining wideband signals/spectrums using sampling rates lower than the Nyquist rate. Compressive Sensing (CS) is an essential type of sub-Nyquist wideband sensing. Compressive sensing is a technique that can efficiently acquire a signal using relatively few measurements, by which unique representation of the signal can be found based on the signals sparseness or compressibility in some domain [2]. The wideband spectrum is fundamentally sparse as the spectrum is poorly utilized. CS is a promising candidate for wideband spectrum sensing as it uses sub-Nyquist sampling rates. The purpose of compressive sensing is to recover signals from far fewer measurements than needed for traditional sampling.

A (ultra-)wide frequency band can be considered, where the both licensed primary communication systems and unlicensed CR networks are hosted by the band. Different from the compressed sensing scheme in [4], we consider a slotted frequency segmentation structure in which the entire wideband channel is divided into M non-overlapping narrowband sub-channels centered at $\{f_m\}_{m=0}^{M-1}$ [5]. The transmitted signals of J active primary users are considered and denoted by $s_j(t), j = 1,...,J$. The signal transmits through a wireless fading channel and

$s_j(t)$ reaches the i-th CR receiver in the form $h_{ij}(t) * s_j(t)$, where $*$ denotes convolution and $h_{ij}(t)$ is the channel impulse response [5]. The received signal at CR i is thus given by [5],

$$r_i(t) = \sum_{j=1}^{J} h_{ij}(t) * s_j(t) + w_i(t) \quad (1)$$

where the ambient noise $w_i(t)$ is white Gaussian with zero mean.The first step of compressive sampling is to collect timedomain samples [5]. To reduce the sampling burden and

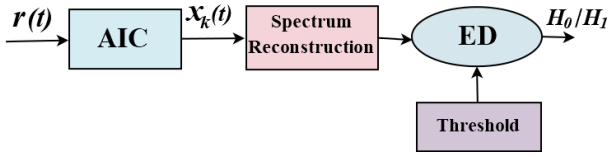


Fig. 1. Compressive Sensing by means of Analog to Information Converter (AIC)

computational complexity in the wideband regime, a hint of promising Analog to Information Converter (AIC) have illustrated for spectrum sensing architecture in Fig.1. An AIC is conceptually similar to an analog-to-digital converter (ADC) operating at Nyquist rate followed by compressive sampling which gives us the opportunity to avoid the use of high rate ADC . Therefore, this model consists of a pseudo-random number generator, a mixer, an accumulator and a low-rate sampler. The pseudo-random number generator produces a discrete-time sequence that extracts information from the modulated signal $r(t)$ by a mixer. The accumulator is used to sum the demodulated signal, while a low rate sampler is utilized to sample its output signal. After that, it is necessary to reconstruct the sparse signal which can be directly accomplished from partial measurements using CS algorithms. Let the output of the AIC is sampled signal of received signal,$r(t)$, denoted by r_k,

$$r_k = [r_{kM}, r_{kM+1}.....r_{kM+M-1}] \qquad k = 0,1..K \quad (2)$$

r_k is the time domain representation of $r(t)$, and is a $M \times 1$ vectors. The size of the measurement matrix Φ_A is $M \times N$, such that,

$$x_k = \Phi_A r_k \quad (3)$$

Here $N < M$ and measurement matrix Φ_A is a set of basis function used to collect the time domain sample from the received signal. It amounts to collect samples on the Nyquist grid but skipping randomly $(M-N)$ time instants to reduce the average sampling rate [5]. Hence, output of the AIC denoted by the size of $N \times 1$ vectors,

$$x_k = [x_{kN}, x_{kN+1}.....x_{kN+N-1}] \qquad k = 0,1..K \quad (4)$$

Now we need to estimate the spectrum of the received signal for energy detection (ED).Eventually, recovery of the compressive sampling can be done by solving l_1 norm optimization problem [6]- [7] which is based on basis pursuit algorithm [8].

For energy detection(ED) scheme, the energy of the reconstructed spectrum of received signal will be compared with a threshold, measured by Gamma Inverse Function [9]. The task of spectrum sensing is the following binary hypothesis-testing problem [10]:
H_0 : primary user is absent;
H_1 : primary user is present.

B. Cooperative Spectrum Sensing

When multiple CRs cooperate with each other in a network for the purpose of spectrum sensing, the process is called cooperative sensing. Fig. 2 illustrates a system model where compressive sensing is established for cooperative network. In a cognitive radio network (CRN), every cognitive node

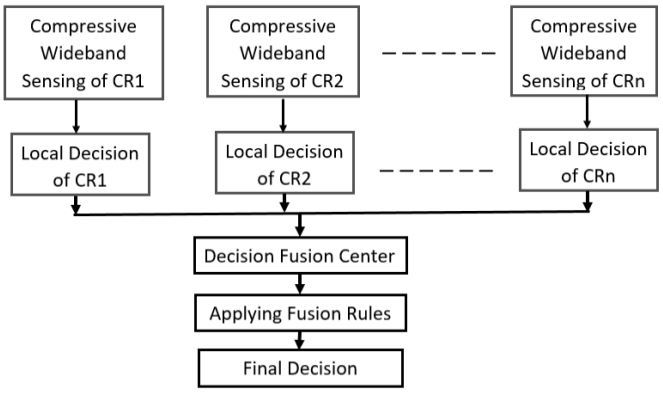


Fig. 2. System Model for Compressive Sensing in Cooperative Network

separately performs compressive spectrum sensing utilizing energy detection scheme. Then to settle the conclusion they sends the local decision (binary or decision statistic) through control channel to the common receiver called Decision Fusion Center. The received spectrum at individual CR is compared with a predetermined threshold value to construct the local decision. In order to minimize the controll channel overhead, CRs only share their final 1-bit hard decisions (H_0 or H_1) rather than their decision statistics [10]. A cognitive manager (CM) or decision fusion center is accountable for collecting the local sensing decisions from all the member nodes presented in a CRN. Then the local decisions are fused and combined by employing different fusion rules. Thus the final decision made by the fusion center will enhance the detection performance in a cooperative CRN. We have applied four decision fusion rules at Fusion Center:

1) AND-rule: In this rule, if all of the local decisions sent to the decision maker are one, the final decision made by the decision maker is one [11]. The fusion centers decision is calculated by logic AND of the received hard decision statistics. For simulation, we have used the following formula for AND-rule [11]:

$$P_{d,AND} = P_{d,i}^{N} \quad (5)$$

Where $P_{d,i}$ is the probability of detection for i-th CR node and N is the total number of CR nodes.

2) OR-rule: In this rule, if any one of the local decisions sent to the decision maker is a logical one, the final decision made by the decision maker is one [11]. For simulation, we have used the following formula for OR-rule [11]:

$$P_{d,OR} = 1 - \Sigma(1 - P_{d,i})^N \tag{6}$$

3) Majority-rule: In this rule, if half or more of the local decisions sent to the decision maker are one the final decision made by the decision maker is one [11].The following formula from [11] describes the rule:

$$P_{d,MAJ} = \sum_{l=\lfloor N/2 \rfloor}^{N} \binom{N}{l} P_{d,i}^{l}(1 - P_{d,i})^{N-l} \tag{7}$$

Here $\lfloor . \rfloor$ represents floor operator. In mathematics and computer science, the floor function maps a real number to the greatest preceding or the least succeeding integer, respectively. More precisely, $floor(x) = \lfloor x \rfloor$ is the greatest integer less than or equal to x.

4) Optimal Fusion rule: Though the logical operations can provide a satisfactory performance in many practical scenario but the optimality of the fusion is not guaranteed. The optimal fusion (OF) rule comes out which is employed by the following likelihood ratio test [12]:

$$\Lambda = \sum_{i=1}^{n} [u_i \log \frac{P_{di}}{P_{fi}} + (1 - u_i) \log \frac{1 - P_{di}}{1 - P_{fi}}] + \log \frac{P(H_1)}{P(H_0)} \tag{8}$$

Where P_{fi} is the probability of false alarm for i-th CR node and when $\Lambda \geq 0$, then the Primary User is present; otherwise, Decision Maker decides the absence of Primary Users. In line with optimal fusion rule or simply optimum rule, we can set a weighted value for each CR node depending on the distance between cognitive radio user and PU. The weight of a node determines the priority of its decision to affect the final one. The highest priority goes to the CR user which is more near to the PU than other CR user.

C. Fading Models

Fading models are used to reduce the fading effect and improve the performance of the communication channel. The additive white Gaussian noise (AWGN) is a fundamental commotion used in information hypothesis to emulate the impact of numerous irregular procedures that happen in nature. AWGN channel model is characterized by the linear addition of white noise with a constant spectral density and a Gaussian distribution of amplitude. The model does not represent blurring, frequency selectivity, obstruction, non-linearity or scattering. Rayleigh fading channel is introduced where signal is received from multipath. When the impulse response is modeled as a zero mean complex-valued Gaussian process, the channel is characterized as a Rayleigh fading channel [13]. The main difference between the model of Rayleigh fading and Rician fading is that in Rician fading a predominant segment is present. Rician K-factor has been used in the simulation which is defined as the ratio of signal power in dominant component over the (local-mean) scattered power [13]. The Nakagami distribution matches some data which is verifiable by observation or experience rather than theory or pure logic, which makes it superior to other models. Nakagami parameter, m is used to model signal for excessive to mild fading issue. If the signal amplitude follows m, a gamma distribution is followed by the PDF [14].

III. System Model Simulation & Performance Analysis

In this section, a simulation study is performed to validate the analytical results. First we examine the single user case, subsequently we consider cooperative sensing. We consider, at baseband, the wideband signal is in the range of $[1 - 64]$ MHz, that is partitioned into M = 32 equal-bandwidth sub-channels of 2 MHz. In this section, simulations are performed along with the description of obtained graphical layouts which are acquired during the sensing of primary user signals applying compressive sensing scheme and different decision fusion rules in cooperation. Results are compared between different fading channels. All simulations in this work is executed using MATLAB version R2016a. Monte Carlo (MC) method, is a stochastic technique (based on the use of random numbers) which forms the basis of these simulations. BPSK modulated signal has been applied in the spectrum sensing method for different fading channel (AWGN,Rayleigh,Rician,Nakagami). The detection performance is analyzed based on the Receiver Operating Characteristic (ROC) curve. ROC curve is a fundamental plot in signal detection theory where the advantage of ROC curve is utilized to capture all the aspects of signal detection theory in one graph.

At first the performances of different fusion rules should

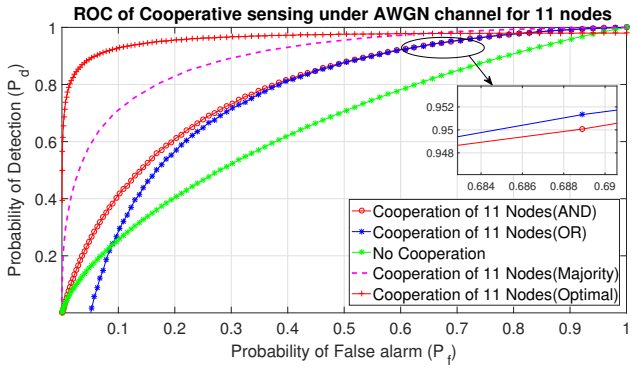


Fig. 3. Comparison of different fusion rules under AWGN channel

be compared employing a specific fading channel. The ROC curves for the decision fusion rules (AND, OR, Majority, Optimal Fusion) under AWGN channel has been shown in

Fig.3 for 100000 Monte-Carlo simulation. The figure indicates that the performance of cooperation process (11 nodes) is more improved than the non-cooperative one, because probability of false alarm may rise due to some hidden node problem when primary user's transmitted signal could not be detected by one node because of the location of the node. But whenever a number of nodes (11 nodes in this case) cooperate their detection result the probability of detection become higher than the probability of false alarm and thus cognitive radio device can avoid unwanted interference to the primary user (PU). For more better detection and less interference it is needed to compare different fusion rules under definite fading or non-fading channel. The application of AND rule under AWGN channel enhances the probability of detection in the region where probability of detection, P_d increases more rapidly than probability of false alarm, P_f. In this case, there is the possibility of causing interference with PU is the least as the presence of PU can be detected by all of the nodes. However, in the region where P_d increases less rapidly than P_f, OR-rule surpasses the ROC of AND-rule which can be observed from the zoomed part of the figure. It can be said that the region where the probability of false alarm become higher gradually, the OR-rule gives better detection performance than AND-rule as the presence of PU can be detected by any one of the nodes. The detection performance by application of majority-rule is superior to both the AND-rule & OR-rule because the PU detection depends on the greater part of the cooperative nodes. Cooperative sensing for Optimal Fusion rule provides a very sharp and abrupt curve as illustrated in the figure which indicates that the probability of detection is higher for a lower probability of false alarm. It is known that a higher value for P_d and a lower value for P_f is desirable at the same time for effective spectrum sensing. Therefore, it can be said that the optimal fusion rule offers the best performance among the four fusion rules under AWGN channel. But in the case, when we did not know the distance between primary user and cognitive radio user we should prefer majority rule for cooperation.

The ROC curves for the decision fusion rules (AND, OR, Majority, Optimal Fusion) under Rayleigh channel has been shown in Fig.4 for 100000 Monte-Carlo simulation. The ROC curve of majority-rule deviates more than the ROC curves of the other three rules. The region of high threshold value, the ROC curve of the majority rule is superior than the ROC curve of AND-rule and OR-rule momentarily. After a while, the ROC curve of AND-rule & OR-rule, become superior than the ROC curve of the majority rule, when the threshold value decreases and P_f rises. However, in the region where P_d increases less rapidly than P_f, OR-rule surpasses the ROC of AND-rule which can be observed from the zoomed part of the figure. It can be said that in that region, the OR-rule gives better detection performance than AND-rule as the presence of PU can be detected by any local decision received by the Fusion Center. Cooperation done by Optimal Fusion rule provides a very sharp and abrupt curve as illustrated in the figure which indicates that the probability of detection is higher for a lower probability of false alarm. Therefore, the optimal fusion rule presents the greatest performance among the four fusion rules under Rayleigh channel.

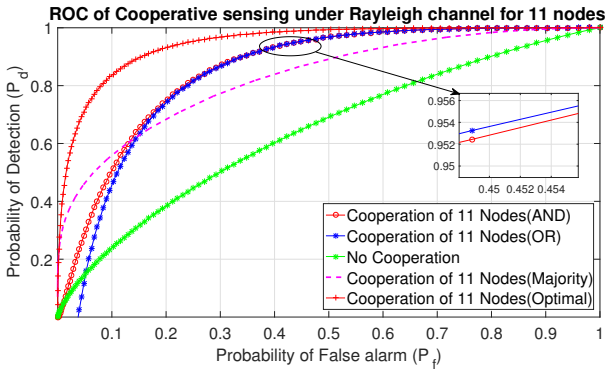


Fig. 4. Comparison of different fusion rules under Rayleigh channel

The ROC curves for the decision fusion rules (AND, OR, Majority, Optimal Fusion) under Rician channel has been shown in Fig.5 for 100000 Monte-Carlo simulation. The Rice factor (k) has been set as 2 in case of cooperative spectrum sensing. Among the four rules, optimal fusion rule performs the best detection performance while the performance reduction is most noticeable for OR-rule according to the analysis. The performance of majority rule is inferior than optimal fusion rule but superior than AND operation & OR operation. The ROC of AND-rule surpasses the ROC of OR-rule because AND-rule gives better detection performance here. However, in the region where P_d increases less rapidly than P_f, OR-rule surpasses the ROC of AND-rule which can be observed from the zoomed part of the figure. The ROC curve rises almost linearly for the non-cooperative Compressive Sensing approach. Cooperative approach has analyzed for 11 CR nodes.

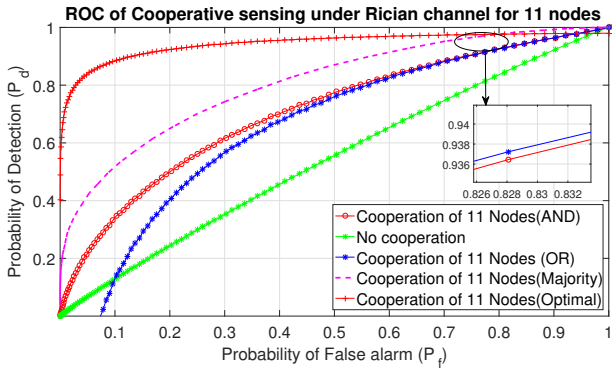


Fig. 5. Comparison of different fusion rules under Rician channel

The ROC curves for the decision fusion rules (AND, OR, Majority, Optimal Fusion) under Nakagami channel has been shown in Fig.6 for 100000 Monte-Carlo simulation. The Nakagami model represents an analytical evaluation where Nakagami distribution, m is applied for analysis of fading data. The Nakagami distribution (m) has been considered as

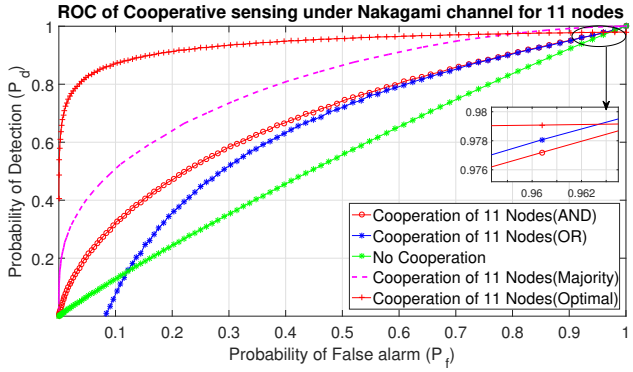


Fig. 6. Comparison of different fusion rules under Nakagami channel

unity. The figure indicates that using the optimal fusion rule provides better detection than the other rules under Nakagami fading effect. The reason behind it is that when it is used, the cooperation is performed based on the weighted value of the CR nodes. The performance of OR-rule reduces the most, yet the majority-rule offers better performance than other non-cooperation approaches, AND-rule, and OR-rule. However, in the region where probability of false alarm rises, logic operation OR surpasses the ROC of logic operation AND which can be observed from the zoomed part of the figure.

The ROC curves for different channel (AWGN, Rayleigh,

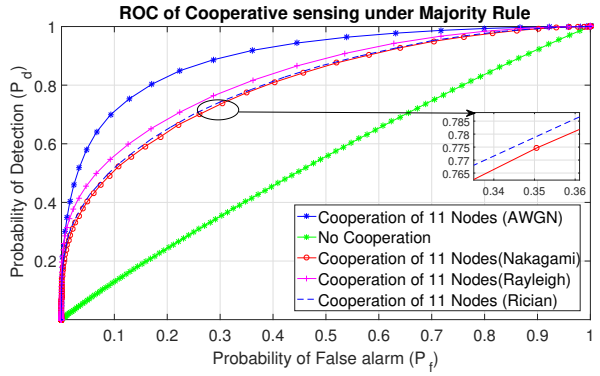


Fig. 7. Cooperation using Majority-rule under different fading channels

Rician, Nakagami) under majority rule has been shown in Fig.7 for 100000 Monte-Carlo simulation. When the cooperation of 11 nodes is done in non-fading AWGN it shows the best performance as indicated in the figure. For Rayleigh channel, the P_d is lower in comparison with the AWGN channel because of the multi-path fading characteristics of the Rayleigh channel. In this figure, the performance of cooperative sensing in Rician channel is more inferior than Rayleigh fading channel & superior than Nakagami fading channel. The ROC curve of majority-rule for Nakagami channel deviates more than the ROC curves of majority-rule for the other two fading channels. Hence for Nakagami fading channel, the performance of cooperative sensing degrades mostly than other two fading channels, due to rapid fading effects. The Nakagami distribution (m) has been considered as unity.

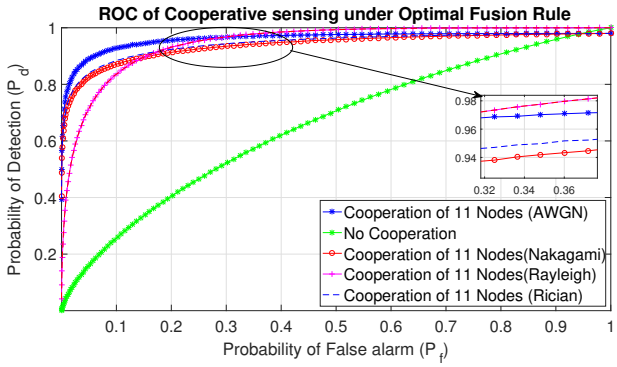


Fig. 8. Cooperation using Optimal Fusion-rule under different fading channels

Now we need to know for more effective spectrum sensing which fading model will be appropriate for optimal fusion rule, which has the most superior performance than the other three fusion rules. The ROC curves for different channel (AWGN, Rayleigh, Rician, Nakagami) under optimal fusion rule has been shown in Fig.8 for 100000 Monte-Carlo simulation. When the cooperation is done in non-fading AWGN it shows the best performance as indicated in the figure. From the figure comparing with the Rician fading channel, the performance degrades more under Nakagami channel, although optimal fusion rule provides almost equal performance under both fading channel. At the region where P_d increases more rapidly than P_f, the performance of Rayleigh fading channel is inferior than the performance of Rician & Nakagami channel momentarily. After a while, as the value of P_f rises gradually, the performance of Rayleigh channel become superior than the other two fading channel which can be observed from the zoomed part of the figure. The Rayleigh fading channel, is suitable for transmission in the absence of Line of Sight (LoS) between the transmitter and the receiver due to multi-path fading. Whereas Rician channel provides a strong dominant Line of Sight component & Nakagami channel characterizes rapid fading. Therefore, it can be said that, for non-fading case, cooperative approach using optimal fusion rule under AWGN channel is preferable. But for fading case, Rayleigh channel is the optimum condition for spectrum sensing in CR network.

IV. Conclusion

Spectrum sensing performed by a stand-alone CR is critical and challenging while dealing with multi-path fading and hidden node problems and reliability of performance is questionable. Thus, cooperative spectrum sensing is introduced to get reliable detection performance as well as to minimize the interference to PUs and in consequence, the false alarm probability is reduced. We have studied cooperative spectrum

sensing over non-fading and different fading channels in CR and compared among them. It has been found that probability of detection is increased by using different fusion rules for cooperation which can reduce hidden node problem. It's been also observed that the Optimal Fusion rule has better performance than AND, OR and MAJORITY rule. After Optimal Fusion rule, the priority goes to MAJORITY rule for co-operation due to the limitation of Optimal Fusion rule where distance between primary user and cognitive radio user is needed to be known. It's also observed that spectrum sensing is harder in presence of Nakagami fading channel and performance of detection reduces more in Nakagami channel than other fading channel. Comparing different channels it can be said that best detection performance can be achieved under non-fading AWGN channel. But a communication channel generally experiences fading as transmitted signal cross a different path to reach to receiver due to presence of reflection in the environment. Therefore, for effective spectrum sensing the next priority goes to Rayleigh fading channel according to the analysis. The output signals were simulated successfully and their characteristics were recognized. Their cooperation performance were analyzed based on the several decision fusion rules in the presence of AWGN, Rayleigh, Rician and Nakagami fading. At the end, the compressive sensing played a vital role to improve the sensing performances by cooperation. In future, the analysis of throughput and the computational burden can also be performed as a part of performance analysis of Compressive Sensing in cooperative network.

REFERENCES

[1] Tian, Zhi and Giannakis, G.B., "A Wavelet Approach to Wideband Spectrum Sensing for Cognitive Radios", *IEEE 1st International Conference on Cognitive Radio Oriented Wireless Networks and Communications, CROWNCOM,july,2006*

[2] Hongjian Sun, Arumugam Nallanathan, Cheng-Xiang Wang, Yunfei Chen, "Wideband Spectrum Sensing for Cognitive Radio Networks: A Survey", *IEEE Wireless Communications,April 2013*

[3] Zhi Tian, "Cyclic Feature Based Wideband Spectrum Sensing Using Compressive Sampling", *Proceedings of IEEE International Conference on Communications, ICC 2011, Kyoto, Japan, 5-9 June, 2011*

[4] Z. Tian, and G. Giannakis, "Compressed Sensing for Wideband Cognitive Radios", *Proc. of IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Vol. IV, pp. IV.1357-1360, Honolulu, April 2007.*

[5] Zhi Tian, "Compressed Wideband Sensing in Cooperative Cognitive Radio Networks", *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference,30 Nov.-4 Dec. 2008*

[6] D. L. Donoho, "Compressed Sensing", *IEEE Transactions on Information Theory, vol.52,no.4, pp. 1289-1306,2006.*

[7] E. J. Cands and M. B. Wakin, "An introduction to compressive sampling", *IEEE Signal Process. Mag., vol. 25, no. 2, pp. 2130, Mar. 2008.*

[8] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit", *SIAM J. Sci. Comput., vol. 20, no. 1, pp. 33-61, 1999.*

[9] Mitsuru Uchiyama, "The principal inverse of the gamma function", *Proc. Amer. Math. Soc. 140 (2012), 1343-1348,August 3, 2011*

[10] Jiaqi Duan and Yong Li, "Performance analysis of cooperative spectrum sensing in different fading channels", *2nd International Conference on Computer Engineering and Technology,16-18 April 2010.*

[11] Md. Shamim Hossain , Md. Mahabubur Rahman, Md. Ibrahim Abdullah and Mohammad Alamgir Hossain, "Hard Combination Data Fusion for Cooperative Spectrum Sensing in Cognitive Radio", *International Journal of of Elektrical and Computer Engineering (IJECE),Vol.2, No.6, December 2012.*

[12] P. K. Varshney, "Distributed detection and data fusion", *Springer-Verlag,1st ed., Book chapter, 1996.*

[13] G. A. Bhalerao, R.G. Zope, "BER improvement of DSCDMA with rake receiver using multipath fading channel *International Conference on Recent Trends in engineering & Technology 2013, 2013.*

[14] F. Digham, M. Alouini and M. Simon, "On the energy detection of unknown signals over fading channels", *IEEE Transactions on Communications, vol. 55, no. 1, pp. 21-24, 2007.*

Design of EEG Based Wheel Chair by Using Color Stimuli and Rhythm Analysis

Nafiul Hasan[1],Md. Mahmudul Hasan[2], Md. Akramul Alim[3]
[1,2,3]*Department of Elictrical and Electronic Engineering,*
[1,2,3]*Bangladesh Army University of Engineering & Technology*,
Qadirabad, Bangladesh
[1]nafiulhasan71@gmail.com, [2]mubin.eee.kuet@gmail.com,[3]eaar@ymail.com

***Abstract*—A novel methodology of brain controlled intelligent wheelchair by using color stimuli is proposed here. A general methodology is applied to find out most effective rhythm for color classification. Primary colors RGB and secondary color yellow were chosen for left, right, forward and stop command. Alpha, Beta, Theta, Delta rhythms were analyzed for three different subjects. Using dissimilar features of time and frequency domain twelve artificial neural network were built to decide the best rhythm. Principal component analysis was made for each EEG signal to remove the background effect of color stimuli. Comparing the findings it is visualized that beta rhythm is the most efficient rhythm with minimum mean square error of 4.845×10^{-9} among all designed ANN for color classification.**

Keywords— Electroencephalogram (EEG), Principal component analysis (PCA), Rhythm, Color stimuli, Artificial Neural Network (ANN).

I. INTRODUCTION

About hundred billion neurons are in human brain. Thoughts and emotional states makes interaction among these neurons. As a result different amplitude and frequency based EEG are produced. There are five rhythms namely alpha, beta, delta, theta and gamma [1]. Their frequency is eight to thirteen, thirteen to thirty, one to four, four to eight and thirty six to forty four Hz respectively.

Researches are going on to make an efficient and effective intelligent wheel chair for Quadriplegic and paralyzed patients. Most of recent works are based on eye blink such as described in [2]. Eye blinking is a natural phenomenon and uncomfortable command feature. By using one and two eye blink and attention level, a BCI was developed. Attention and eye blinks were also used in [3] but they have made it wireless and interfaced with RS-232C. SSVEP based wheelchair control is analyzed in [4]. Eye blink based another work is discussed in [5]. However an eye blinking pattern was used by them. An cost effective automated wheelchair with using conventional eye blinking feature is discussed in [6]. By recognizing eye movement such as blinking, glancing, eye opening a model is discussed in [7]. Researches on the basis of steady state visual evoked potential with different flickering frequency showed up good performance [8]. For classification support vector machine was used with satisfactory result.

In most of the cases either eye blinking feature or EEG raw signal was used. Some used internet of things. But there is lack of effectiveness. An innovative approach by using color stimuli is used here. To remove the background effect of the color stimuli the acquired EEG signals were gone through principal component analysis. Later band analysis was done. Though this work is depended on color stimuli but it can introduce more command features.

This paper presents a methodology to drive an EEG based wheelchair by using color stimuli with most effective rhythm. The primary colors red, green, blue were used for direction command (left/right/forward). A secondary color yellow was used to pause. The data acquisition was done by BIOPAC and Acqknowledge 4.1 software [9]. PCA and attribute extraction was done. Artificial neural network were built for classification in MATLAB 2014a.

The following portion of this article is designed as follows: part II describes about suggested method, section III describes about analyzing, section IV and V represents data analysis and classification. It ends with a opinion and coming up scope in portion VI.

II. PROPOSED MODEL

A. Projected Model

For designing EEG based wheel chair by color stimuli with PCA feature the model is described in Fig. 1. After getting the brain signal by BIOPAC and MP 36 , noise and artifact were removed. The rhythms were obtained by FIR band pass filter. To remove background effect of the color stimuli principal component analysis was done by Acqknowledge 4.1 software. The features were also extracted in this platform. Total twenty features were extracted in different frequency bands. Comparing the mean square error of artificial neural network which were based on extracted features a decision is taken about the proposed model efficiency and most effective rhythm.

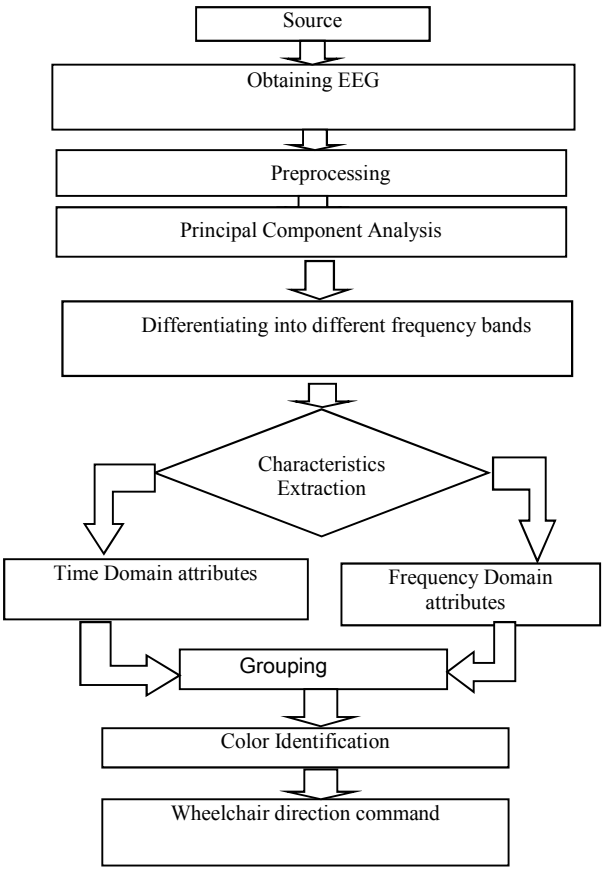


Fig. 1. Flow chart of suggested scheme

A. *Electroencephalography Determination and Exploratory Condition*

The test was conducted in Biomedical Engineering Lab of Khulna University of Engineering & Technology, Bangladesh. The test was conducted on three subjects, not suffering from any psychological disease or color blindness. The color stimuli was shown in personal computer monitor and instructed to focus on indefinite color. In real life patients may focus on monitor or colored board. Total ten trials were taken. First nine trials were used for training and validation purposes and the rest for test purpose.

B. *Experimental arrangement*

1) *Hardware description*: The BIOPAC data acquisition unit MP 36 was used in this experiment. Fig. 2 defines acquiring of EEG while conducting the experiment

2) *Diagonistic software*: Varioud softwares like BIOPAC student Lab Pro was used. For PCA and feature extractaion Acqknowledge 4.1 software was utilized. Principal component analysis was neceassry for removing background effect as the obtianed raw data was so much noisy. Artificial neural network was used in MATLAB 2014a platform.



Fig. 2. Representational view while conducting experiment in BME lab, KUET

III. PREPROCESSING AND ANALYSIS

A. *Filtering , Artifact seperating, smoothing*

The raw electroencephalogram was saved while experiment. There was noise may be due to EMG, EOG, hand movement, background effect of color stimuli. As 0.5 Hz to 44 Hz only necessary, band pass FIR filter was used. The line frequency (50 Hz) was removed. Smoothing was done for averaging the signal values.

B. *Principal Component Analysis*

Source is decomposed to a new signal by PCA. It is very effective as a data reduction tool. As the raw data was so much noisy although after removing artifact it was hard to classify individual color. So PCA was made for better color classification.

C. *Frequency bands analysis*

After filtering, smoothing, principal component analysis EEG was split up in different frequency bands by Acqknowledge 4.1 software. FIR band pass filter was used.

D. *Feature Extraction*

In time and frequency domain total twenty features were extracted. Maximum value, standard deviation, skew, kurtosis, Linear regression, moment of order from one to five, Lyapunov exponent, mutual information, correlate, Nonlinear model, power spectrum density mean , PSD max, PSD skew, fast Fourier transform mean, FFT max, FFT skew were extracted for three subjects in alpha, beta, delta, theta rhythms. All attributes were extracted in Acqknowledge 4.1 software.

IV. ATTAINED DATA

A. *Contrast of attributes of a individual rhythm for an individual color*

In time and frequency domain total twenty attributes were analyzed. In frequency domain PSD mean, PSD max, PSD skew , FFT mean , FFT max, FFT skew were analyzed. These attributes of subject 1 beta band for blue color is given in Table I.

TABLE I. RESPONSE OF BETA BAND IN FREQUENCY DOMAIN OF SUBJECT 1 FOR BLUE COLOR.

Trial	PSD (μV^2)/Hz			FFT(μV)		
	Mean	*Max.*	*Skew*	*Mean*	*Max*	*Skew*
1	0.038	0.748	4.085	0.016	0.341	4.024
2	0.038	0.769	4.185	0.017	-8.160	1.260
3	0.044	0.700	3.342	0.017	0.315	3.601
4	0.042	0.622	3.400	0.019	0.338	3.639
5	0.033	0.512	3.544	0.016	0.280	3.625
6	0.050	0.647	3.176	0.019	0.393	4.214
7	0.040	0.546	3.023	0.016	0.360	4.354
8	0.025	0.200	2.125	0.016	0.177	2.439
9	0.030	0.394	3.119	0.018	0.241	2.960
10	0.038	0.490	2.900	0.018	0.310	3.133

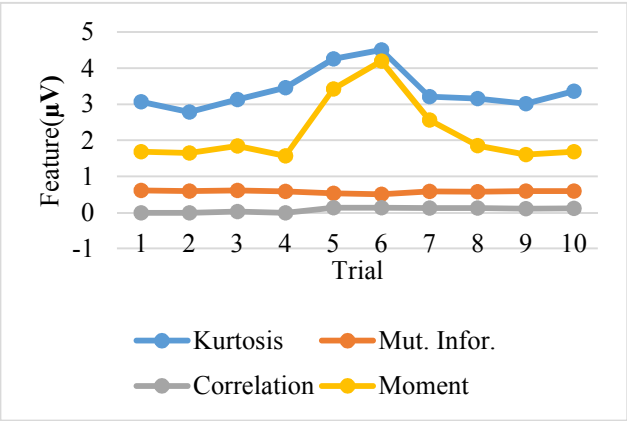


Fig. 3. Different attributes of beta band of subject 2 for red color.

Fig. 3 shows features of beta rhythm of subject 2 for red color. The kurtosis, mutual information, correlation, moment varied very small for same color and subject. Similarly twenty features were obtained for red, yellow, green, blue color and alpha, beta, delta, theta rhythm.

B. *Contrast of attributes of definite rhythm for various color*

There were three subjects and instructed at same condition. After PCA raw EEG was separated in frequency bands. Twenty features for each rhythm and every color were extracted. Fig. 4 shows PSD mean value of subject 1 of beta rhythm for different colors. For different colors the features varied greater comparatively.

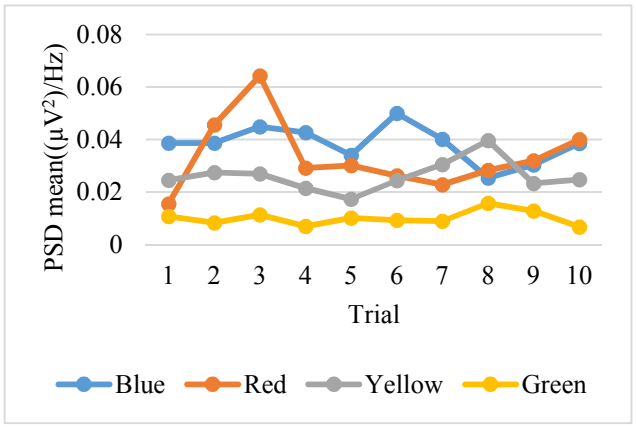


Fig. 4. Comaprison of PSD mean of subject 1 using beta rhythm for different colors

V. CLASSIFICATION RESULTS AND DISCUSSION

The categorization was required for analyzing the effectiveness of color classifying so that individual command can be identified. Moreover the effectiveness of rhythm also analyzed. By comparing the results of ANN most effective rhythm for color classification as well as command detection is determined. MATLAB built in nntool GUI was used. Feed forward backprob. , TRAINLM training function, LEARNGDM adaption learning function, two layer of Tan-Sigmoid transfer function, twenty neurons, Levenberg Marquardt learning algorithm were used. The target value was 10,20,30,40 for blue, red, yellow, green respectively.

A. *Scenario 1: Alpha Rhythm*

The ANN designed for alpha rhythm was successful to classify the colors having mean square error of 4.845×10^{-9} , regression of 0.98157 , gradient of 5.889×10^{-4} . This ANN showed worst performance.

B. *Scenario 2: Theta Rhythm*

The ANN designed for theta rhythm was able to classify colors as well as command with mean square error of 1.1028×10^{-7}, regression of 0.996 , gradient of 1.10731×10^{-3} . It showed better performance than previous.

C. *Scenario 3: Delta Rhythm*

The ANN designed for delta rhythm was able to classify colors as well as command with mean square error of 3.56×10^{-9} , regression of 0.94164, gradient of 1.2478×10^{-5}. Relatively better performance obtained.

D. Scenario 4: Beta Rhythm

The ANN designed for beta rhythm was able to classify colors as well as command with mean square error of 4.845×10^{-9} , regression of 0.98157, gradient of 5.889×10^{-4}. It showed the best performance. Performance of the ANN designed for beta rhythm of subject 1 is shown in Fig. 5.

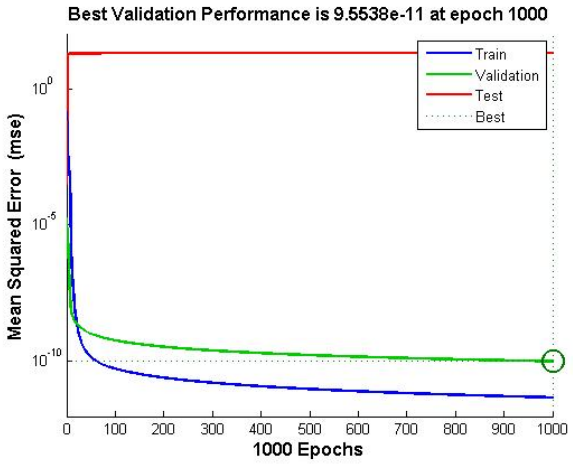


Fig. 5. Performance of ANN designed for Beta rhythm of subject 1

TABLE II . COMPARISON OF TWELVE DESIGNED NEURAL NETWORK.

Rhythm	Subject	Mean Square Error	Overall Regression	Gradient
Beta	1	9.5538×10^{-11}	0.9887	1.0290×10^{-6}
	2	1.4438×10^{-8}	0.9560	0.0017657
	3	1.5100×10^{-12}	1	4.1267×10^{-10}
	Average	4.8450×10^{-9}	0.9815	5.8890×10^{-4}
Delta	1	8.4434×10^{-9}	0.9997	2.0042×10^{-6}
	2	2.2354×10^{-9}	0.8260	1.2010×10^{-5}
	3	1.7720×10^{-12}	0.9991	2.3422×10^{-5}
	Average	3.5600×10^{-9}	0.9416	1.2478×10^{-5}
Theta	1	2.4561×10^{-7}	0.9999	0.0029
	2	5.9513×10^{-8}	0.9989	2.5050×10^{-5}
	3	2.5746×10^{-8}	0.9896	0.0003
	Average	1.1028×10^{-7}	0.9960	1.1073×10^{-3}
Alpha	1	5.6423×10^{-7}	0.9768	0.0518
	2	1.3553×10^{-5}	0.9938	0.1130
	3	3.2050×10^{-7}	0.9990	0.0227
	Average	4.8125×10^{-6}	0.9898	0.0625

All the findings of each designed neural network are showed in Table II. As total twelve ANN were built and there were three subjects so results of ANN for all subjects for same rhythm were calculated to an average value for generalization. The overall accuracy for color classification of Beta, Delta, Theta, Alpha rhythm based ANN were 100%, 99.8%,96.3% and 95% respectively. Data obtained in this work showed best accuracy for color classification as well as command for direction of wheel chair movement with Beta rhythm. Delta rhythms showed better performance and alpha rhythm showed worst performance.

VI. Conclusion

From this offline analysis of brain wave for the improvement of brain controlled wheel chair by using color stimuli, it is crystal clear that color stimuli can be used for this purpose effectively. The beta rhythm provided best performance for color classification of individual with minimum mean square error of 4.845×10^{-9} . Contradictory alpha rhythm showed worst performance with mean square error of 4.81257×10^{-6} .This method is more effective than conventional eye blinking based model because the command feature can be increased with increase in number of color stimuli. The feature of rotation at a certain angle, forward, backward, pause,left, right can be introduced through this proposed method. However there is a drawback of this proposed method is subject's visual weakness. The efficiency can be increased with increase in number of channels.

The future work can be implementation of this method with an increased number of color stimuli and channel. Making a fast online analysis system of EEG is under progress.

ACKNOWLEDGMENT

This work was assisted by Higher Education Quality Enhancement Project (HEQEP), University Grant Commission, Bangladesh under Subproject "Postgraduate Research in BME", CP#3472, Khulna University of Engineering & Technology, Bangladesh.

REFERENCES

[1] M. M. Hasan, M. H. A. Sohag, M. E. Ali, and M. Ahmad, "Estimation of the Most Effective Rhythm for Human Identification using EEG Signal", 9th International Conference on Electrical and Computer Engineering (ICECE),2016.

[2] A. Dev, M. A. Rahman, and N. Mamun3 , "Design of an EEG-based Brain Controlled Wheelchair for Quadriplegic Patients", 3rd International Conference for Convergence in Technology (I2CT),2018

[3] J. S. Lin , K. C. Chen, and W. C. Yang, "EEG and Eye-Blinking signals through a Brain Computer Interface Based Control for Electric Wheelchairs with Wireless Scheme", 4th International Conference on New Trends in Information Science and Service Science, 2010

[4] K. Kim and S. W. Lee, "Towards an EEG-basedIntelligent Wheelchair Driving System with Vibro-tactile Stimuli", IEEE International Conference on Systems, Man, and Cybernetics, October 9-12, 2016.

[5] P. Lahane, S. P. Adavadkar, S. V. Tendulkar, B.V. Shah, S.Singhal, "Innovative Approach to Control Wheelchair for Disabled People Using BCI", 3rd International Conference for Convergence in Technology (I2CT), 2018.

[6] A. Maksud, R. I. Chowdhury, T. T. Chowdhury, S. A. Fattah, C. Shahanaz,S. S. Chowdhury, "Low-cost EEG Based Electric Wheelchair with Advanced Control Features", Proceedings of the IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, 2017.

[7] H. T. Nguyen1, N. Trung, V. Toi and V. Tran, "An autoregressive Neural Network for Recognition of Eye Commands in an EEG-Controlled Wheelchair", The 2013 International Conference on Advanced Technologies for Communications (ATC'13), 2013.

[8] R. Singla, A. Khosla,and R. Jha, "Design of EEG Based Wheel Chair by Using Color Stimuli and Rhythm Analysis", J Med Eng Technol, vol. 38(3), pp. 125–134, 2014

[9] Reference manuals of AcqKnowledge software guide and MP system Hardware guide. (2008). BIOPAC Systems Inc., 42 Aero Camino,Goleta, CA – 0067, USA.

Contrast Enhancement using Mode-Separate Histogram Equalization

Sadia Khan Rupa
Information and Communication Technology
Bangladesh University of Professionals
Dhaka, Bangladesh
sadiakhanrupa1996@gmail.com

Md. Faishal Yousuf
Samsung R & D Institute
Dhaka, Bangladesh
faishal.yousuf@gmail.com

Shanto Rahman
Information and Communication Technology
Bangladesh University of Professionals
Dhaka, Bangladesh
shanto.rahman@bup.edu.bd

***Abstract*—Due to the inferior quality of image capturing devices and acute environmental condition, digital images may suffer from poor contrast. Several histogram equalizing methods have been proposed to enhance these images but most of them either create over enhancement with artifacts or cannot fulfill consumers' desired contrast. However, due to having improper handling of the pixels, histogram equalization (HE) based techniques often fail. Hence, a new image enhancement approach has been introduced here named as Mode-Separate Histogram Equalization (MSHE) which finds a separation point using most frequent intensity of an image and afterwards HE is applied. In experiments qualitative and quantitative assessments exhibits that the performance of MSHE is better than other contrast enhancement with brightness preservation techniques.**

***Index Terms*—Contrast enhancement, Histogram equalization, Brightness preservation**

I. Introduction

Digital images often suffer from low contrast due to the poor quality of image capturing device, light variation, surface reflection and so on. In that case, it demands the process of contrast enhancement, brightness preservation, image correction by denoising and sharpness enhancement. For making a good quality image, several methods have been already proposed. Of those, Histogram Equalization (HE) is most widely used technique [1]. In spite of its popularity, it fails to preserve mean brightness and causes over enhancement at times. To overcome this problem, Brightness preserving Bi-Histogram Equalization (BBHE) proposes mean intensity of an image to separate the histogram [2], whereas Dualistic Sub Image Histogram Equalization (DSIHE) uses median for partitioning the histogram [3]. Meanwhile, Bilateral Histogram Equalization with Pre-processing (BHEP) uses harmonic mean [9] and Minimum Mean Brightness Error Bi-histogram Equalization (MMBEBHE) computes a threshold based on absolute mean brightness error to separate the histogram [8]. However, these mean based histogram separation techniques fail for different kind of images, for example BBHE doesn't work well when there is no symmetrical distribution and DSIHE sometimes causes over enhancement due to an outlier. For dark images, BHEP fails to enhance the contrast. Again computational complexity for MMBEBHE is very large.

To mitigate above mentioned problems several recursively separated histogram equalization based techniques have been proposed. Among those, Recursive Mean Separate Histogram Equalization (RMSHE) uses the mean to separate the histogram recursively [4] and Recursive Sub Image Histogram Equalization (RSIHE) considers the median value for the same [5]. Recursively Separated and Weighted Histogram Equalization (RSWHE) uses both mean and median value to separate the histogram recursively and additionally it modifies the probability density function *(pdf)* to use as a normalized power law function [6]. Though brightness is preserved more in these recursion techniques, larger recursion level means no enhancement, output image is same as the input image. To mitigate these problems, it is essential to give answer of the following research questions.

- How to separate histogram for making equal distribution?
  To separate histogram, the most frequent intensity of an image with a weighting factor can be proposed.
- How the intensities can be spreaded all over the intensity range?
  For spreading the intensities histogram equalization should be applied over the sub parts. And finally these two sub parts are merged to get the desired one.

The rest of this paper is organized as follows. Section 2 presents an overview of the existing works while Section 3 presents proposed solution. Section 4 demonstrates the efficacy of the proposed algorithm and other existing methods. Finally, Section 5 concludes the findings and future remarks.

II. Literature Review

As contrast enhancement is a crucial demnad, different methods have already been proposed. HE enhances the contrast by flattening histogram distribution of an image [1]. However, HE creates over enhancement and causes mean shift problem. Different extensions of HE exists for enhancing image contrast mostly based on mean, median of the image. Details are being discussed in the following.

The first mean separation technique proposed by Kim et al. titled as BBHE where an input image histogram is decomposed into two based on their mean [2]. Then HE is applied over two parts individually and combined into one single image. BBHE works well when the input image follows symmetric distribution but it fails to perform well when input images have non-symmetric distribution [9].

Another algorithm that uses mean as the criterion for histogram division is RMSHE, proposed by Chen et al., is a generalization of both HE and BBHE [4]. RMSHE is proposed to separate the histogram recursively and it produces 2^r number of sub histograms where recursion level *r* is specified by the user. Although this algorithm preserves more brightness, it produces some side effects too. Again if the value of *r* is too large, the output image is same as the input image [22].

Chen et al. propose to use absolute difference between input and output image mean to decide a threshold level in MMBEBHE [8]. A threshold for the lowest absolute mean brightness error (AMBE) is selected and using this input image is separated into two and equalized individually as in BBHE. Nevertheless, the repetitive calculation of AMBE creates large computational complexity [6].

Harmonic mean is also used in BHEP, proposed by Amil et al. [9], which sharpens input image to get the edges correctly using high boost filter and divides image histogram using harmonic mean. After combining the sub images into one image, it is smoothed using Gaussian filter to remove the extremely sharp edges. This technique may not enhance the contrast for dark images [15].

Another extension of histogram equalizing technique is median based. First image median to separate histogram is used in DSIHE, proposed by Wang et al. [3]. Although DSIHE does not shift the input mean, sometimes it fails to preserve original image mean brightness. At times it creates over enhancement problem because in an image the median value can be a high peak or impulse noise [26].

RSIHE is proposed by Sim et al. which also uses median recursively as the separating factor [5]. The sub images will be decomposed, if the recursion level *r* is specified as *2*. Although RSIHE performs better than RMSHE, it also faces the same problem as RMSHE when the value of *r* is too large [23].

Kim et al. propose RSWHE which uses both mean and median image intensity to separate histogram recursively [6], similar to RMSHE and RSIHE. It also introduces histogram weighting module which smooths the sub histograms using normalized power law function. However, statistical information might lose and quality can be degraded [7].

Besides these histogram equalizing algorithms, some techniques such as adaptive gamma correction with weighting distribution (AGCWD) and adaptive gamma correction for image enhancement (AGC) follow gamma correction. AGCWD overcomes the problem happening in RSWHE where cumulative distribution function (*cdf*) is directly utilized for setting up the parameter [7]. AGCWD works well except when input images having a lack of bright pixels, then enhancement may not be like as desired [19]. Rahman et al. propose AGC which also uses gamma correction where images are classified into four classes based on their mean and standard deviation [14]. Although AGC works well for almost all kind of images, for some high contrast images it needs further improvements increasing the gamma value [21]. Again this method can not preserve edge details of original image [20].

Focusing on low contrast images Celik et al. propose an algorithm named as Contextual and Variational Contrast Enhancement (CVC) which increases the gray level differences between neighboring pixels [11]. 2-D input image histogram is constructed here, so the computational complexity of this algorithm becomes larger [7].

Layered Difference Representation (LDR) is proposed by Lee et al., considers full dynamic range of an input image to improve contrast [12]. Unfortunately, LDR treats all the pixels equally where foreground pixels should be given more importance [24].

Tsai et al. propose decision tree-based contrast enhancement for various color images where luminance variance and mean are used to classify input image into six groups [13]. Then image enhancement is done based on the peak and valleys, and piece wise linear transformation function is applied. The predefined values might not work well in all the cases [14].

Using a local transformation method Hussain et al. propose histogram specification technique for dark images, which applies a sobel operator to separate edges and non edges and computes histograms for both [15]. Segments are found and used to have two target histograms which are lastly merged to get one; that is used in the histogram specification technique to enhance the whole image. It fails for bright images and create artifacts in the bright regions [27].

From the above discussion, it is found that among the existing image enhancement algorithms, histogram equalizing methods show superior outcome compare to others for low contrast images. Separating histogram and applying HE on both sub parts is the main idea of HE based techniques. However, dividing the histogram imprecisely is a major reason of failure to enhance the image for these techniques. Therefore, for improving low contrast images, accurate separation of histogram is required in HE oriented algorithms.

III. Proposed method

The main objective of this paper is to transform low contrast image into a visually pleasant one by enhancing its contrast. Hence, a Mode-Separate Histogram Equalization (MSHE) is proposed where a trade off has been done between over enhancement and low contrast.

In MSHE, at first image mode is calculated which is the most frequent intensity of an image and adjusted with a weighting factor and a bias which gives an accurate separation point to divide the histogram. Afterwords, histogram equalization is applied individually on each part of the histogram. The sub histograms are combined into one to get the desired output image. The overall procedure is shown in Fig. 1.

At the beginning of MSHE, a histogram separation point (X_m) has been identified to separate image perfectly which is calculated by Equation 1.

$$X_m = c + \alpha \times X_{mode} \tag{1}$$

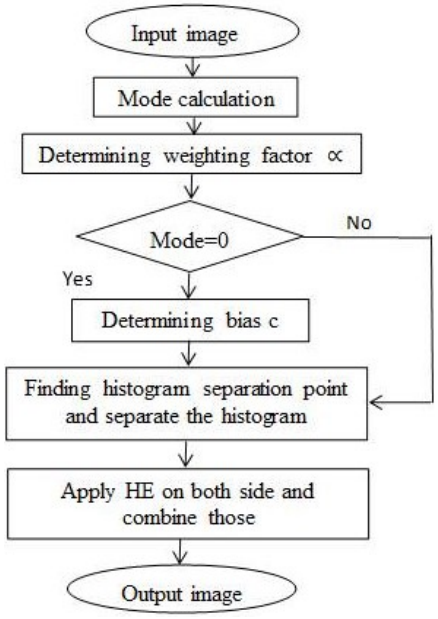


Fig. 1. Flowchart of MSHE

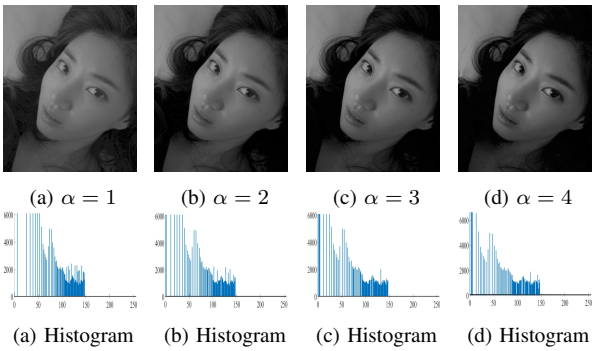


Fig. 2. Impact of using α greater than 1 over 'Girl' image

Here, X_{mode} is the image mode which is adjusted with a weighting factor (α) and a bias (c) where α should be as follows

$$\begin{cases} 0 < \alpha < 1, & X_{mode} = \wp \\ \alpha \geq 1, & X_{mode} < \wp \end{cases} \tag{2}$$

Here, $\wp$ is the highest intensity of an image i.e., 255. When mode is exactly $\wp$, α should be in $[0, 1]$. Again sometimes α needs to be greater than or equal to 1 for getting better contrast when image mode is less than $\wp$. From Fig. 2 and Fig. 3, it is shown that the image contrast is improved with the increase of α where it exhibits that best result when $\alpha = 2$. This weighting factor α is determined using tangent of the difference between two ratios of some statistical parameters calculated by the following equation.

$$\alpha = \mid \tan(\frac{X_{max}}{X_{mode}} - \frac{\sigma}{2}) \mid \tag{3}$$

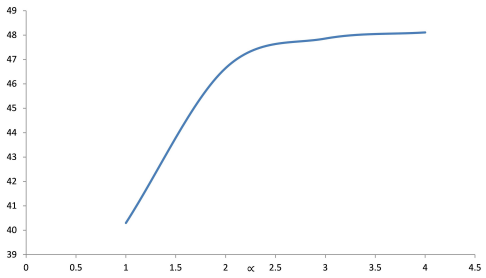


Fig. 3. Root-mean-square contrast of 'Girl' image is increasing with α

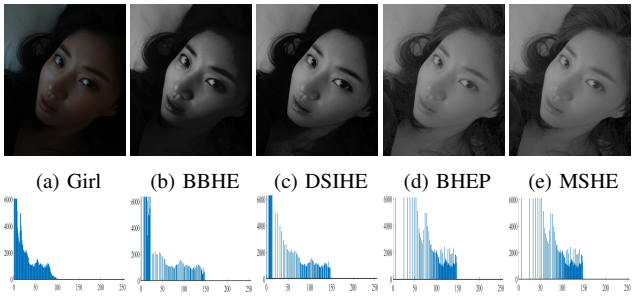


Fig. 4. Rational behind choosing Mode as the separation point

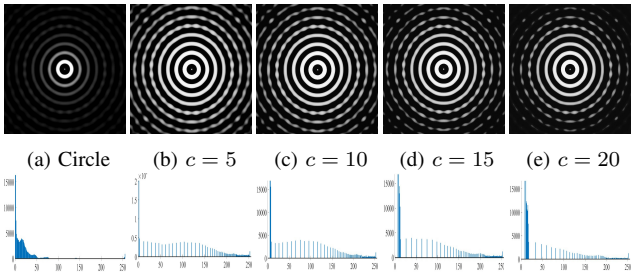


Fig. 5. Rational behind choosing c

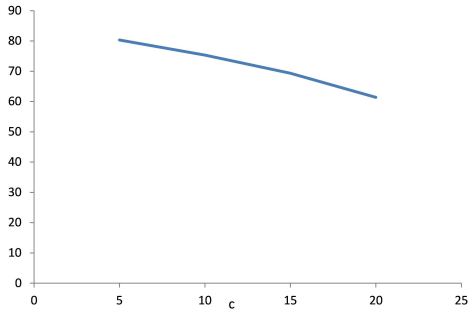


Fig. 6. Root-mean-square contrast of 'Circle' image is decreasing with c

Here, σ, X_{max} are the standard deviation and the maximum intensity respectively. tan function returns the output of input variable that is expressed in degrees. The reason of using tan function is given in Lemma 1 and Lemma 2.

Another factor c is calculated using absolute difference

between standard deviation (σ) and image mean (μ), follows specific condition similar to the Dirac delta (δ) function [25]. c only contains value when most frequent intensity of an image is zero. The impact of c is shown in Fig.5 and Fig.6.

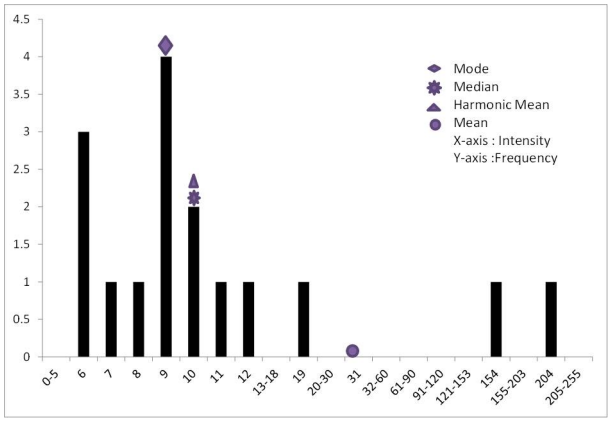


Fig. 7. Impact of choosing image mean, median, harmonic mean and mode

$$c(X_{mode}) = \begin{cases} \mid \sigma - \mu \mid, & X_{mode} = 0 \\ 0, & otherwise \end{cases} \tag{4}$$

Now input image is decomposed into two sub images X_L and X_U based on X_m followed by Equation 5 and 6.

$$X_L = X(i,j) | X(i,j) \leq X_m, \forall X(i,j) \in X \tag{5}$$

$$X_U = X(i,j) | X(i,j) > X_m, \forall X(i,j) \in X \tag{6}$$

Here, X is an input image which is separated using X_m. X_L and X_U are the lower and upper part of the image respectively. The sub image X_L is composed of $\{X_o, X_1, ..., X_m\}$ and the other part X_U is composed of $\{X_{m+1}, X_{m+2}, ..., X_{n-1}\}$. Later, histogram equalization is applied on the two parts individually and finally the output image is achieved by combining the two sub parts.

A. Rational behind choosing tan

Lemma 1 α remains less than 1, if X_{mode} = 255.
proof For $X_{mode} = 255$, the maximum value of α in proposed method will be $\alpha_{max} = \mid \tan(1 - \frac{\sigma}{2}) \mid$. For a choice of $\sigma = 10$, we get $\alpha_{max} = \mid \tan(1 - \frac{10}{2}) \mid = .06992 < 1$

Lemma 2 $\alpha \geq 1$, iff tan is used.
proof When $sine$ or $cosine$ function is used instead of tan, the maximum value of α_{max} can never be greater than one. For a choice, using statistical parameters of 'Girl' image, given in Fig. 2, we get $\alpha_{max} = \mid \sin(\frac{148}{2} - \frac{25}{2}) \mid = \mid \sin(61.5) \mid = .879$, where σ=25, X_{max}=148 and X_{mode}=2. In case of $cosine$ function we get $\alpha_{max} = \mid \cos(61.5) \mid = .477$ where $\alpha_{max} = \mid \tan(61.5) \mid = 1.84$

B. Rational behind choosing Mode

Histogram equalization uses the highest intensity of an image to equalize it, while its extensions like BBHE, DSIHE, BHEP and even MSHE changes its highest intensity by separating into two parts. From Equation 5 and 6, the upper limit X_m is the highest intensity for the separated part X_L and it is used to equalize X_L where X_{n-1} is used for the other part X_U.

$$\begin{bmatrix} 9 & 10 & 12 & 6 \\ 8 & 10 & 9 & 11 \\ 6 & 19 & 204 & 154 \\ 9 & 6 & 7 & 9 \end{bmatrix} \tag{7}$$

The effect of choosing mode instead of mean, median and harmonic mean is described using a 4×4 matrix (Equation 7) which is similar to BHEP [9]. From Fig.7, it is shown that the mean of the matrix is 31 which denotes that 87% pixels (14 out of 16) are on the left side and rest only 13% pixels stays in the right side. As a consequence, the 87% pixels are not enhanced more than intensity 31 in BBHE. But the aim is to divide histogram where pixels are equally distributed on both parts because only this can ensure maximum contrast of an image. Median and harmonic mean are used to divide histogram in DSIHE and BHEP respectively and for this matrix both median and harmonic mean is 10 which signifies 69% pixels (11 out of 16) are in the left side and 31% pixels are in right side. Equal distribution of histogram on both parts is not even ensured in DSIHE and BHEP too. Meanwhile, MSHE proposes mode based contrast enhancement which provides equal distribution of histogram. Here, mode is 9 and α is 1, returns a separation point at 9. This divides histogram into two parts, containing 56% (9 out of 16) and 44% pixels, respectively(according to Fig.7). The impact of selecting mode is also illustrated in Fig. 4.

IV. Experimental result

In this paper, the proposed method is compared with some other existing contrast enhancement techniques namely HE [1], BBHE [2], DSIHE [3], RMSHE [4], RSIHE [5], AGCWD [7], AGC [14]. This comparison is conducted followed by qualitative and quantitative approaches.

A. Qualitative assessment

In this section visual assessment has been performed over a large number of images and the data set has been taken from [1]. Most of the cases, it has been found that MSHE shows better outcome than other techniques.

Fig. 8 (Stonewall) contains a dark image where the main challenge of this image is to read the message written over the wall. It is found that the message can be read in HE but artifacts are also introduced. BBHE does not enhance contrast at all and DSIHE creates accepted enhancement where at least the message can be seen properly. RMSHE and RSIHE produce similar enhancement as in BBHE and DSIHE respectively and this is because for both cases the recursion level have been used one. Although AGC enhances contrast

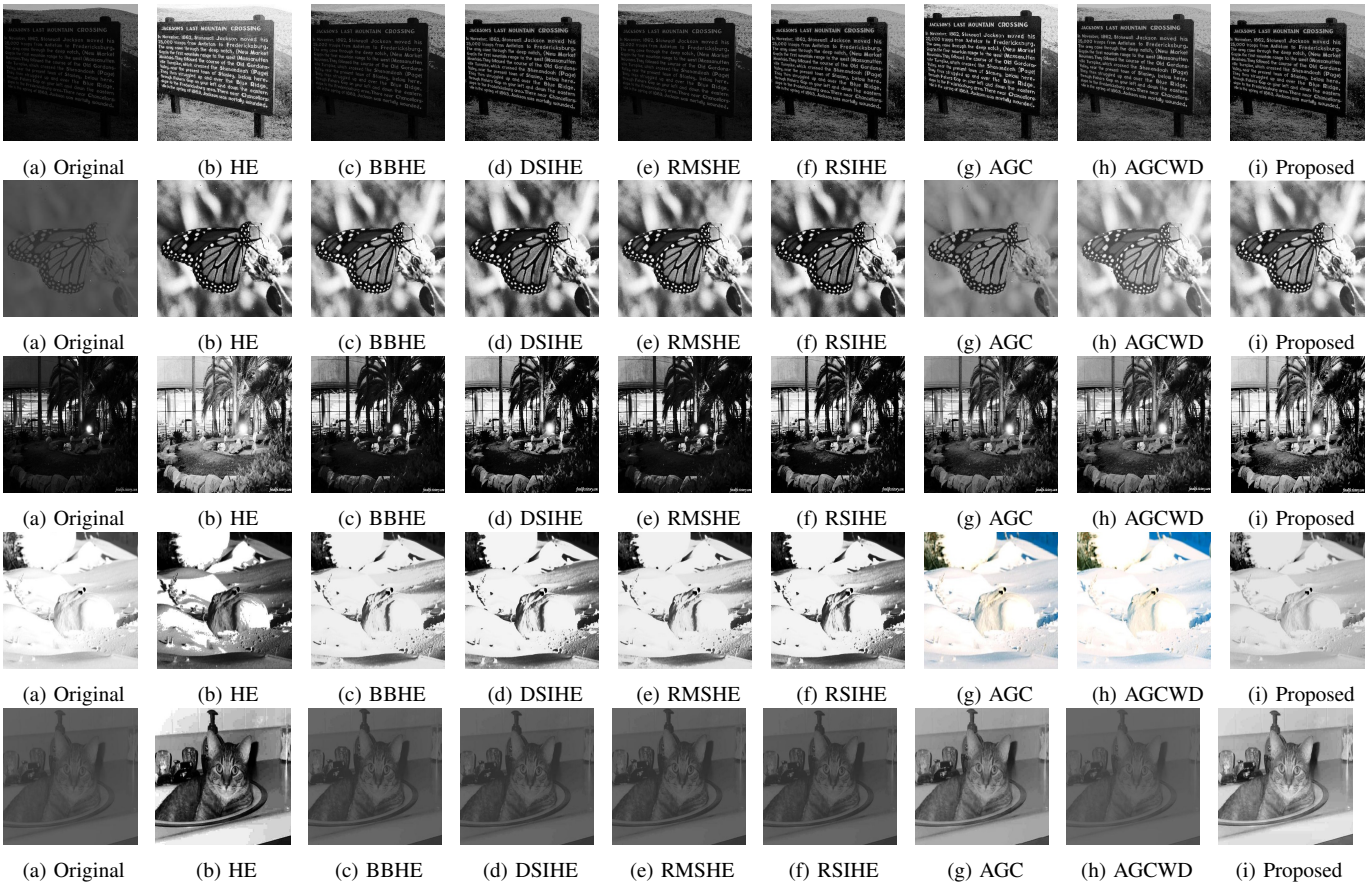


Fig. 8. Enhancement results of different methods (top-down → 'Stonewall', 'Butterfly', 'Fountain', 'Hare', 'Cat')

but it creates a blurry image where the message cannot be read at all whereas AGCWD exhibits the message quite clearly. However, MSHE enhances the image with appropriate contrast where the message is clearly readable without any artifacts.

Another dark image with low contrast is in Fig. 8 (Butterfly) where the main difficulty is to find the image details including the flower. HE, BBHE, DSIHE, RMSHE and RSIHE create over enhancement. Although AGC does not produce over enhancement, it fails to enhance the brightness. AGCWD enhances the contrast well but it produces over brightness problem. Meanwhile, the proposed method shows output with perfect contrast with details. Fig. 8 (Fountain) is also a dark image where brightness enhancement is needed. HE produces desired brightness enhancement. BBHE and RMSHE fail where DSIHE and RSIHE show insignificant impact. AGC and AGCWD show comparative better result but more enhancement is needed for clear visualization, while MSHE is successful for corroborating brightness enhancement.

Fig. 8 (Hare) is a low contrast image where HE shows over enhancement as usual, DSIHE and RSIHE increase contrast more than BBHE and RMSHE respectively. However, all of them need further improvement. In the case of gamma correction, AGC exhibits better result than AGCWD. MSHE enhances without creating over enhancement.

Another dark image with low contrast is in Fig. 8 (Cat) where the major concern is both the cat and its background glasses. HE enhances the image but the background suffers from over enhancement problem. BBHE, DSIHE, RMSHE, RSIHE show no change in outcome. AGC produces accepted enhancement for this image whereas AGCWD almost creates no enhancement. But among all the techniques, MSHE enhances the image properly where both cat and background glasses are equally visible.

B. Quantitative measurement

The quality judgment of an image can vary from person to person. Through quantitative evaluation, it is possible to prove numerically the quality of images i.e., enhanced or not. Here, root-mean-sqaure (rms) contrast and discrete entropy (DE) are used to evaluate the performances of different enhancement techniques along with the proposed method.

1) Root-mean-square (rms) contrast: Measuring *rms* contrast is a well known technique to describe the contrast [17], [18]. Low *rms* refers to low contrast. Table I shows value of *rms* contrast of different methods. From this table, it is shown that HE creates larger *rms* value for most of the images. However, here, over enhancement is severely observed within the images. AGC also produces larger *rms* values. Among the histogram equalization methods, MSHE performs better.

2) Discrete entropy: Entropy refers to uncertainty of an image. An image has more entropy depending on the randomness

TABLE I
Root-mean-sqaure contrast for each algorithm

Image Name	HE	BBHE	DSIHE	AGCWD	AGC	Proposed
Stonewall	74.93	49.48	62.54	60.67	79.05	62.99
Fountain	74.60	68.94	80.26	67.40	63.55	82.05
Butterfly	74.89	72.05	75.44	46.55	35.7	64.92
Hare	74.93	68.10	78.40	45.56	51.0	44.72
Cat	74.70	16.34	16.34	9.86	38.25	48.58

of the variable [16]. Low entropy symbolizes low contrast and higher entropy refers the reverse in image processing. Table II demonstrates the discrete entropy (DE) for each algorithm and it shows that AGC creates larger entropy for all the cases except the 'Cat' image. MSHE holds second larger entropy. Hence, it can be concluded that after AGC, MSHE performs better than other histogram equalizing techniques.

TABLE II
Discrete entropy for each algorithm

Image Name	HE	BBHE	DSIHE	AGCWD	AGC	Proposed
Stonewall	4.92	5.29	5.38	5.40	5.40	5.36
Fountain	5.42	5.91	5.97	7.08	7.24	5.95
Butterfly	4.60	4.78	4.76	4.80	4.80	4.78
Hare	4.92	5.21	5.23	5.09	5.57	5.33
Cat	5.06	4.97	4.97	4.71	5.19	7.20

V. CONCLUSION

As failure of proper partition is a major limitation of histogram equalization based techniques, in this paper we propose a simple and effectual technique for contrast enhancement to abolish this problem. The proposed MSHE determines separation point using mode which splits almost equal area over histogram and produces better enhancement. Its efficacy of producing satisfactory results has been proved over other brightness preserving contrast enhancement techniques by qualitative and quantitative analysis. This motivates to collaborate MSHE with medical image processing, digital photography and other implementations in consumer electronics.

REFERENCES

[1] Gonzalez, Rafael C., and Paul Wintz. "Digital image processing(Book)." Reading, Mass., Addison-Wesley Publishing Co., Inc.(Applied Mathematics and Computation 13 (1977): 451.

[2] Kim, Yeong-Taeg. "Contrast enhancement using brightness preserving bi-histogram equalization." IEEE transactions on Consumer Electronics 43.1 (1997): 1-8.

[3] Wang, Yu, Qian Chen, and Baeomin Zhang. "Image enhancement based on equal area dualistic sub-image histogram equalization method." IEEE Transactions on Consumer Electronics 45.1 (1999): 68-75.

[4] Chen, Soong-Der, and Abd Rahman Ramli. "Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation." IEEE Transactions on consumer Electronics 49.4 (2003): 1301-1309.

[5] Sim, K. S., C. P. Tso, and Y. Y. Tan. "Recursive sub-image histogram equalization applied to gray scale images." Pattern Recognition Letters 28.10 (2007): 1209-1221.

[6] Kim, Mary, and Min Gyo Chung. "Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement." IEEE Transactions on Consumer Electronics 54.3 (2008).

[7] Huang, Shih-Chia, Fan-Chieh Cheng, and Yi-Sheng Chiu. "Efficient contrast enhancement using adaptive gamma correction with weighting distribution." IEEE Transactions on Image Processing 22.3 (2013): 1032-1041.

[8] Chen, Soong-Der, and Abd Rahman Ramli. "Minimum mean brightness error bi-histogram equalization in contrast enhancement." IEEE transactions on Consumer Electronics 49.4 (2003): 1310-1319.

[9] Amil, Feroz Mahmud, et al. "Bilateral histogram equalization with pre-processing for contrast enhancement." 2016 17th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD). IEEE, 2016.

[10] Coltuc, Dinu, Philippe Bolon, and J-M. Chassery. "Exact histogram specification." IEEE Transactions on Image Processing 15.5 (2006): 1143-1152.

[11] Celik, Turgay, and Tardi Tjahjadi. "Contextual and variational contrast enhancement." IEEE Transactions on Image Processing 20.12 (2011): 3431-3441.

[12] Lee, Chulwoo, Chul Lee, and Chang-Su Kim. "Contrast enhancement based on layered difference representation of 2D histograms." IEEE transactions on image processing 22.12 (2013): 5372-5384.

[13] Tsai, Chun-Ming, Zong-Mu Yeh, and Yuan-Fang Wang. "Decision tree-based contrast enhancement for various color images." Machine Vision and Applications 22.1 (2011): 21-37.

[14] Rahman, Shanto, et al. "An adaptive gamma correction for image enhancement." EURASIP Journal on Image and Video Processing 2016.1 (2016): 35.

[15] Hussain, Khalid, et al. "A histogram specification technique for dark image enhancement using a local transformation method." IPSJ Transactions on Computer Vision and Applications 10.1 (2018): 3.

[16] Gull, Stephen F., and John Skilling. "Maximum entropy method in image processing." IEE Proceedings F (Communications, Radar and Signal Processing). Vol. 131. No. 6. IET Digital Library, 1984.

[17] Peli, Eli. "Contrast in complex images." JOSA A 7.10 (1990): 2032-2040.

[18] Wang, Zhou, et al. "Image quality assessment: from error visibility to structural similarity." IEEE transactions on image processing 13.4 (2004): 600-612.

[19] Rahman, Shanto, et al. "Image enhancement in spatial domain: A comprehensive study." 2014 17th International Conference on Computer and Information Technology (ICCIT). IEEE, 2014.

[20] Kallel, Fathi, and Ahmed Ben Hamida. "A New Adaptive Gamma Correction Based Algorithm Using DWT-SVD for Non-Contrast CT Image Enhancement." IEEE transactions on nanobioscience 16.8 (2017): 666-675.

[21] Olugbara, Oludayo O., Tunmike B. Taiwo, and Delene Heukelman. "Segmentation of melanoma skin lesion using perceptual color difference saliency with morphological analysis." Mathematical Problems in Engineering 2018 (2018).

[22] Wang, Chao, and Zhongfu Ye. "Brightness preserving histogram equalization with maximum entropy: a variational perspective." IEEE Transactions on Consumer Electronics 51.4 (2005): 1326-1334.

[23] Ibrahim, Haidi, and Nicholas Sia Pik Kong. "Brightness preserving dynamic histogram equalization for image contrast enhancement." IEEE Transactions on Consumer Electronics 53.4 (2007): 1752-1758.

[24] Lee, Jun-Tae, et al. "Depth-guided adaptive contrast enhancement using 2D histograms." 2014 IEEE International Conference on Image Processing (ICIP). IEEE, 2014.

[25] Dirac, Paul Adrien Maurice. "The physical interpretation of the quantum dynamics." Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character 113.765 (1927): 621-641.

[26] Arici, Tarik, Salih Dikbas, and Yucel Altunbasak. "A histogram modification framework and its application for image contrast enhancement." IEEE Transactions on image processing 18.9 (2009): 1921-1935.

[27] Oloyede, Muhtahir, et al. "A new evaluation function for face image enhancement in unconstrained environments using metaheuristic algorithms." EURASIP Journal on Image and Video Processing 2019.1 (2019): 27.

Cloud-Based Automated Power Consumption Optimization, Power management, and Appliance Control

Juglul Hasan, Tanjim Ul Haque, Sabab Hasan
Department of Computer Science & Engineering
Ahsanullah University of Science & Technology
Dhaka, Bangladesh
e-mail: {juglulhasan.11,tuhaque}@gmail.com, sabab_hasan @yahoo.com

***Abstract*—In the world of modernization, most of the societies are adapting smart technologies for different household, production and professional works. For that, the use of electric power is significantly increasing. In the developing countries, limited energy sources has created an enormous lack of power. Therefore to cope up with the modern world of digitalization and technological advancement, optimizing power consumption has become a very important aspect in this area. So to optimize the usage of power from both national power grid and renewable energy sources to its optimum usage can lead us to a better progression. As most of the modern appliances and systems are being developed nowadays under Internet Of Things, we have used IoT to develop a user-friendly power management and appliance control system.**

Keywords—Internet of Things; power control; power optimization; appliance switching; remote device management; energy distribution.

I. INTRODUCTION

In this era of digitization managing energy is essential in all categories of lifestyle and production whether it's home appliances or industrial machineries. In the developing countries power deficiency is a very common problem. Also the cost of electricity is very high due to scarcity and the scarcity of electric power is increasing rapidly for power wastage and lack of proper usage management. Now a days many countries including the developing countries are using renewable energy sources but that is not sufficient to cover the complete need yet. So power usage optimization is a must for both managing devices and handling the expenses. An intuitive meter powered by a smart controller and an AI modulator that optimizes the electric power usage can solve this matter for most of the cases in an efficient way.

The objective of this paper is to create the guideline for developing a self-regulating power meter that can manage, analyze and optimize the power usage and lets the user assess the power usage situation through a software application. The user will also have control access to the appliances through the software application which is connected to a cloud server that would store all the necessary power usage data sorted by date, time and appliances. The user would have full data and control access to the cloud server, appliances and application system through NFC and RFID. An AI modulator will manage power system optimization using voltage regulator and a complete sensor network system including sensors for temperature, humidity, ambience, lux, noise etc. For example; the controller would automatically turn off air conditioner when temperature is below human comfort level or turn off light at night when need of light is low. The AI modulator would find comfort level, ambience requirement etc. by derived machine learning.

II. RELATED WORKS

Recent researches on the sector of Internet of Things and the power control systems have provided greater advancement in reliable infrastructure for optimizing power in household activities. More research on this aspect have created much efficient smart sensor technologies which allocate us with advanced tools to build up smart home appliances. [1]. As different projects have been done to create a better smart home where the use interactions and securities are given more priority than controlling the power. In the developing countries scarcity of electricity and renewable energy is decreasing modernization in this areas. Also in some places there can be seen renewable energy source but optimization is not up to the marks. In [2] the authors have created cloud based home management system where they incorporated home appliance control, security management through separate medium. In which the expense will be a lot higher and time consuming. The goal of our model is to use renewable energy as much as it can and use limited powergrid to use less electricity. In [3] the author have researched energy management on a city based on canadian condo environment. Where it can be an optimal solution for a particular country but there stands a difference based on availability of the resources.[4] have researched over smart meters for developing countries.

We have proposed a design that can be accessed through smart device of a user to control home appliances and also see optimal usage of power. Where the meter will try to use renewable energy as much as it can and use from power grid when it's necessary. User can access his data using NFC from his smart device and also handle bills through his device. The proposed method can be more cost efficient and can optimize the energy more significantly.

III. Proposed System

The basic outline of our system is devised to orient around a 'smart meter' powered by a customized microcontroller. All the electronic appliances required and installed in any settlement (household, corporation or industrial) are first passed through numbered digital output ports of the microcontroller. Hence any and every device is controllable by the controller. The power output and signal switching are both displayable in the central control monitor of the controller.

The 'smart meter' is connected to the service web server and dedicated cloud server through a dedicated private IP connected PAN (personal area network) by zigbee-IEEE 802.15.4 [5] specification suite, GSM module and basic TCP/IP protocol suite. The GSM module works as the intermediary for the 'smart meter' and the servers (cloud and data) whereas, a combined module suite of zigbee and GSM are used to connect the controller and servers to the controller mobile phone application (android, IOS and windows platform) so that the data access, appliance control and the 'smart meter' all can be managed and controlled by the user.

.

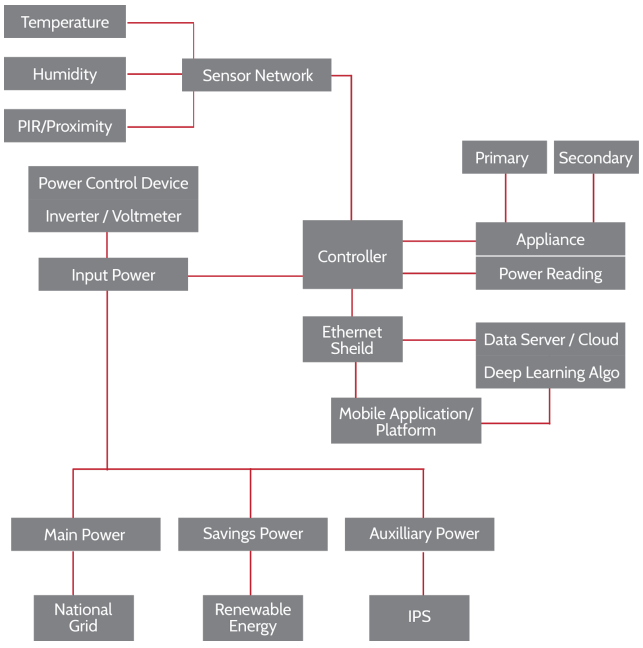


Figure 1: System Workflow Diagram

The main input power for the system and the appliances is provided by a renewable power source like a smart solar power grid or a wind based power turbine system included in the hardware system which is marked as the primary power source while the national power grid line is the secondary power source. There should also be a backup substation or an IPS for emergency power supply.

All the power usage data for every appliance and the switching control information is to be stored in the cloud server, and maintained and analyzed by the AI modulator. The modulator analyzes the usage data per time per day per season per device and optimizes the further power management and switching automation for future usage. The AI modulator generates automated power usage reports and gives warning for excess power usage on heavy load.

IV. Methodology

The overall method of our system is based on a customized microcontroller and interconnection between cloud server. There are several important aspects in our proposed system. As our propose system smart power meter optimizes power usage, it tries to use as much renewable energy it can before using power grid. For that a inverter is used. An AI modulator controls the home appliance based on the results from different sensors (e.g. temperature sensor, humidity sensor). For a user to access it requires a security measure also. The system architecture can be seen for our proposed system below.

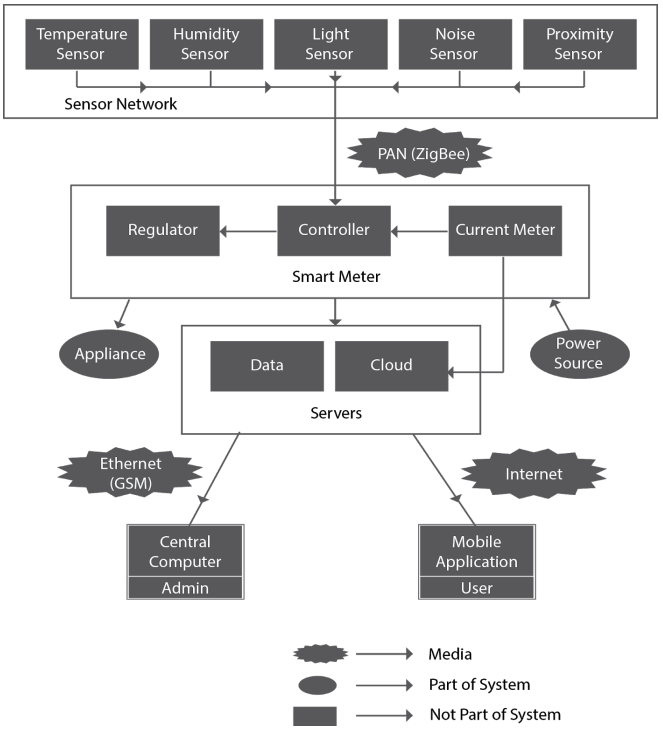


Figure 2: System Architecture

Our proposed approach can be divided in to three significant parts:

- Sensor Networks
- Smart Meter
- Cloud Server & Security

A. Sensor Network

Our proposed system provides a smart solution for controlling the home appliances based on an AI modulator. Which works depending on sensor outputs. Sensor networks includes a combination of temperature sensors LM35 [6] and humidity sensor [7] connected through a Zigbee to control the home appliance if the temperature changes from a threshold value. It also includes light sensor for controlling the lights. Passive Infrared sensor (HC-SR501) [8] and sound sensor [9] are also included for accurate monitoring of the home appliance control. If the one or two sensor value was more than a threshold value but unexpectedly the home appliance failed to shut off the PIR sensor will ensure its situation to shut it off to the smart meter.

B. Smart Meter

The dominant part of our model is the smart meter. It is based on a customized microcontroller which is connected with the power sources. The meter is powered with three types of connection where one is renewable energy, power grid and finally auxiliary power. The first objective of the smart meter is to use renewable energy as much it can then using the national power grid as needed. If a load shedding or a mishap occurs the third auxiliary unit (IPS) takes over the line. Smart meter counts the usage of power grid and inform the users how much units of electricity it have used and calculate the cost. It will also predict and advise the user about how much units of electricity he should use for most optimal results. Other than the AI modulator the user can manually control the home appliance through his smart devices. The power distribution of the smart meter relies on a inverter which connects to the necessary power source as described. An user is needed to authenticate themselves before getting the data from the smart meter. In our proposed model we have used nfc which is available in every phone. After authenticating through the nfc the user can control, see the usage and also pay the bills through local mobile transaction system.

A flow chart for the smart meter can be seen on figure-3 where the power distribution process is described briefly. All the data of energy consumption and optimization are then send in to the server where it all gets stored. The AI modulator interacts with the server data to take more proper decision optimizing the energy. Which provides user with more energy efficiency and cost effectiveness.

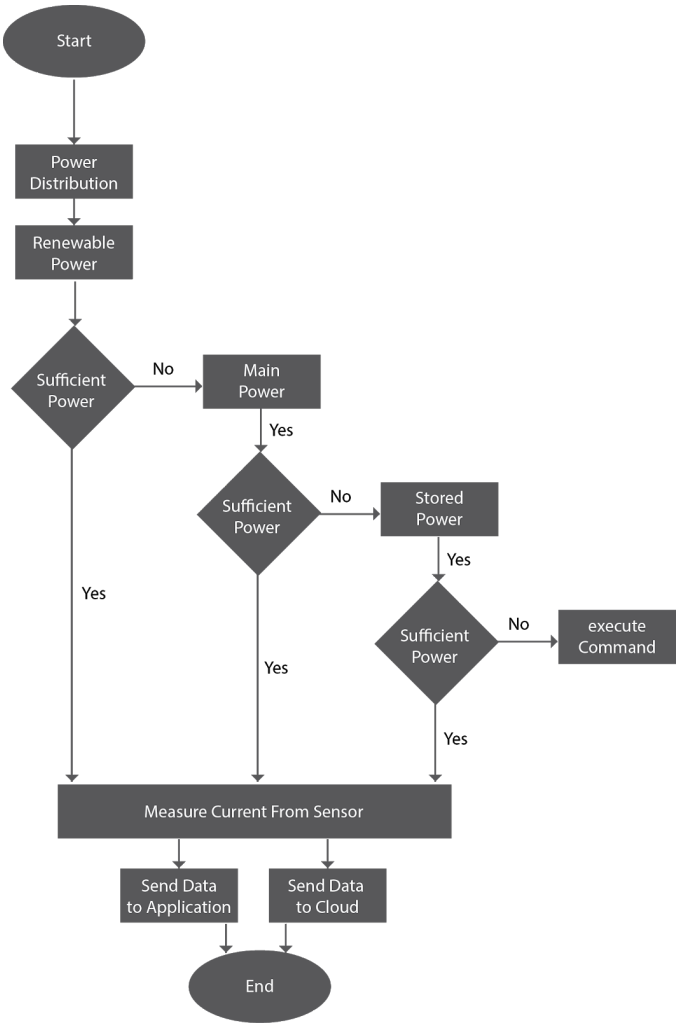


Figure 3: Flow chart for smart meter

C. Cloud Server & Security

Like any other management and optimization system or automation; the proposed power management system also contains a data management server system. But to improve the efficiency of the IoT themed system the model is planned to have cloud server instead of a central data center. As the power usage data is quite sensitive the cloud server (preferably AWS) is designed including CCSK test version 3 cloud security system by implementing 6 layer 'onion server' technology and a dedicated private IP based PAN.Only the meter system and AI modulator can access the server to analyze usage data for further optimization.

Application for proposed System

We have created user friendly application that acts as the bridge for communicating with the smart meter and user acce-

-ssibility. The application provides user with the options to collect the power usage and consumption reports from the smart meter through the cloud server. The user can also control all the home appliances connected with the smart meter. Application also provides billing facilities through local payment gateway. The application can be accessed by the nfc from the user devices. An interface prototype of our application can be seen below.

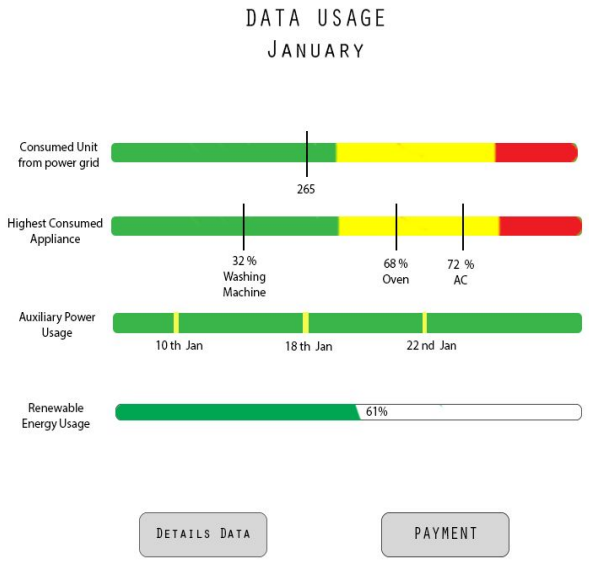


Figure 4: Application interface prototype for smart meter

Emergency Cut off

The proposed system also includes an emergency shut down system where if the microcontroller assigned to smart meter detects any kind of massive temperature variation while distributing power and changing power source through inverter the system cuts off the power until the temperature is back to normal.

V. Collation and Analysis

The proposed prototype is evinced to have a capability of around 18% reduction in power consumption in a closed environment; be it a residential or commercial settlement.The reduction includes switching off unused appliances, load transfer for heavy duty machineries, power switching between grid line and secondary power source for usage optimization and voltage regulation of common appliances.Also the system can ensure a better settlement security through the three layer monitoring sensor network. User can also compare power usage data for different electronic devices based on data & time. This power usage data in cloud database can also train the application AI to future enhance its decision making capabilities; i.e. whether to on or off AC during winter or whether to on or off lights during night-time on its own.

Most of the existing systems like EHS , Insteon, KNX, Z-Wave, X10, UPB, Lon Works , Zigbee etc. are based on short wave and local area connection network for IOT based appliance control. These standards connect the central controller to the required devices in a wireless PAN. Except Zigbee, all other protocols are prone to external hacking and network attacks, as they utilize the open bandwidth of the wireless frequency band.

All the existing systems only serve the purpose of appliance control and remote device switching with a minimal monitoring option. Our designed system model utilizes a dedicated cloud server for data storage and power usage data analytics for future application references. The integrated AI in the cross-platform mobile application would generate alerts and warnings based on unnecessary switched on devices and power over-usage. Thus unlike most existing systems this model can effectively control and minimize over usage and wastage of power. Also the user retains all authorization and full control over both the system and the application.

VI. Conclusion & Future works

Our plan for the smart meter is to incorporate it to the national grid line and all end user power consumer settlements. That way the whole power line for an area can be effectively organized and optimized. Also the power wastage can be reduced which is a very significant aspect in the developing countries or countries with scarcity of power resources. We have tried to incorporate most of the aspects that contributes maintaining a more optimized use of power. More use of renewable energy not only helps the environment but also lessens the expense of a user. Smart technologies combining with artificial intelligence can guide the user towards substantial energy efficiency.

In the near future our primary plan is to assimilate our own deep learning algorithm in the AI modulator; so that it can analyze and synthesize more data with faster speed. Also by deep learning the modulator will be able to optimize the power automation system further. We are already working on the deepmind api an open source platform by google inc. In near future we hope that about 70% of our daily power needs will be met by renewable energy, which can be solar power , wind power or any other renewable energy source. We are also trying to incorporate smart solar power grid and also enhancing the custom microcontrollers to be more precise so that energy efficiency becomes easier which can lead to a better environment and better living.

REFERENCES

[1] Farhangi, H. The path of the smart grid. IEEE Power Energy Mag. 2010, 8, doi:10.1109/MPE.2009.934876.

[2] Lee, Y.T.; Hsiao,W.H.; Huang, C.M.; Seng-cho, T.C. An integrated cloud-based smart home management system with community hierarchy. IEEE Trans. Consum. Electron. 2016, 62, 1–9.

[3] Ciancetta, F.; D'Apice, B.; Gallo, D.; Landi, C. Plug-n-Play Smart Sensor Based on Web Service. IEEE Sens. J.2007, 7, 882–889.

[4] Keelson, Eliel, K. O. Boateng, and Isaac Ghansah. "A Smart Retrofitted Meter for Developing Countries." International Journal of Computer Applications 90.5 (2014).

[5] Adams, Jon T. "An introduction to IEEE STD 802.15. 4." Aerospace Conference, 2006 IEEE. IEEE.

[6] Poonam, Y. Molge, and Yusuf Mulge. "Remote temperature monitoring using LM35 sensor and intimate android user via C2DM service." International Journal of Computer Science and Mobile Computing 2.6 (2013): 32-36.

[7] Tomer, Hemlata, and Kapil Mangla. "Study and Development of Temperature & Humidity monitoring system through Wireless Sensor Network (WSN) using Zigbee module Hemlata Tomer*, Kapil Mangla." International Journal of Engineering Research and Applications 5.7 (2015): 115-120.

[8] Yun, Jaeseok, and Sang-Shin Lee. "Human Movement Detection and Identification Using Pyroelectric Infrared Sensors." Sensors (14248220) 14.5 (2014).

[9] Hakala, Ismo, et al. "Design of low-cost noise measurement sensor network: Sensor function design." Sensor Device Technologies and Applications (SENSORDEVICES), 2010 First International Conference on. IEEE, 2010.

An IoT Based Green Waste Management System For Bangladesh

Mazhar Ibna Zahur, Md. Saidur Rahman, Aysha Akther*, Kazi Masudul Alam
DGTED Lab, Computer Science and Engineering Discipline
Khulna University, Khulna, Bangladesh
{mazhar1322, said1419, aysha*, kazi}@cseku.ac.bd

***Abstract*—Internet-of-Things (IoT) has introduced new avenues of application development towards the Smart City vision. Urbanization is a common trend of current society and its rapid growth demands smarter solutions for waste management in the cities. A probable solution of this problem is the adoption of IoT based waste management systems. In this case, we have to build smart bins that will be connected with the cloud setup of the city authority and will provide constant updates of the waste status of the bins. A critical issue in this case is to provide power to this smart bin establishments. This problem can further be handled by incorporating reusable energy sources such as solar panels with the smart bin setups. In this paper, we design and develop an IoT and cloud based framework for waste management. We have designed and developed smart bins that are intermittently connected to the cloud data storage to share their waste status. City authority is given required web interfaces to manage smart bin installations. Garbage collectors are given mobile applications to help planning garbage collection routes dynamically.**

***Index Terms*—Waste Management, Green IoT, Smart Bin, Arduino Microcontroller, Reusable energy**

I. Introduction

New technologies are sweeping the market every day with one specific goal, improving our daily living experience. IoT is an enabling technology to solve issues related to rapid urbanization. One key problem of urbanization trend is waste management of the cities. Bangladesh is also facing the same problem. Adopting IoT technologies for waste management can provide some solutions to this problem.

Bangladesh is a country with small land but large population. The number of city dwellers in Bangladesh is constantly increasing in regards to the management capacity of the city authority or the city infrastructure [13]. As a result, there is more waste, which requires more man power to collect for disposal. Due to the lack of man power and proper planning waste management procedure is not scalable. This has caused tremendous distress for the overall health of the citizens as well as caused severe damage to the environment.

Implementation of Smart City through the adoption of IoT offers a viable option to overcome this problem. In the event of CES-2018 several technology leaders have demonstrated the practical use of the Smart City Concept [3]. If we can connect all types of citizen services in a city with the internet, then citizens and the administration can communicate better, which is important to improve the quality of city living. Government of Bangladesh also has the vision to improve the digital services of the country. The main issue of adopting new technologies in Bangladesh is the requirement of power. Due to lack of adequate power generation, utilization of renewable energy for IoT technologies is a viable option for Bangladesh. This however is termed as Green Internet-of-Things i.e. Green IoT.

Green IoT focuses on energy efficiency from all perspective of IoT deployment [2]. It is believed that billions of IoT devices by 2025 will consume enormous amount of power that will contribute tremendously to greenhouse effect [5]. Adoption of green power source will ensure that IoT deployment is environment friendly. With this objective, we have designed and developed IoT based waste management system that can scale as well as utilize green energy to power the service. History data from the system will be used to identify the areas that require larger bins as well as increased number of bins. City authority can use this data to rearrange bins locations as well as improve necessary garbage collection attempts to ensure cost efficient health city living.

Rest of the paper is organized as Section II describes state of the art works from smart dustbin, smart waste management and IoT based waste management perspectives. In Section III, we describe the architecture of our proposed Green IoT based waste management system for Bangladesh. Section IV later presents detailed implementation approach of our developed prototype, which leads to Section V that concludes the paper.

II. Related works

Smart City applications are the latest trend in IoT research. Smart waste management in the cities is a significant application of this genre. There are some research works related to smart waste management that are utilizing IoT technologies. Most of the smart waste management models tackle the problem related to any specific country. Following sections describe some state of the art solutions of smart waste management.

In [7] authors describe about smart dustbin for Smart Cities. In this system there are integrated GPS and Wi-Fi enabled Smart Bins. The system has three layers – Dustbin Layer, Server Layer and Client layer. Dustbin Layer consists of some hardware including Arduino UNO R3, Sonar sensors and Wi-Fi module. Server Layer is basically data layer where Wi-Fi module updates dustbin position and current state. Client Layer receives data from the server to observe the status of the bins.

In this paper [8], authors demonstrate how IoT can be enabled with data access networks, geographic information system and electronic engineering works to improve traditional waste management system. In their solution, waste is collected by gathering intelligence about trashcans, where the prototype is embedded with sensors that can collect and transmit trash volume data over the Internet. Spatio-temporal context of the data is considered in this data collection model that further uses graph theory optimization procedure to handle waste efficiently.

Raina et al. [10] presents a concept about smart garbage collector that uses microcontroller to watch the status of the garbage bin and report the status in an HTML page. They use different types of sensors to collect the information about the garbage, which is shared using wireless systems. HTML page is also accessible through smartphones. Here they have created a server in Wi-Fi module to send data in HTML page, later they access data from any Internet-enabled devices.

In this paper, we propose an IoT based green waste management system. The proposed system is divided in three parts: Bin layer, Data layer and User layer. In the bin layer we have smart bins that are loaded with different sensors to collect garbage level information from a bin, communicates the data with the data layer. Two types of users waste management monitor or garbage collector can access this data for garbage collection route finding or for statistical analysis.

III. Proposed System Architecture

In this section, we provide details of the proposed IoT based green waste management system. Fig. 1 shows the proposed architecture of our system. We have divided the proposed system into three parts: Bin layer, Data layer and User layer. Bin layer consists of Smart Bin that is designed to collect garbage status data. The data layer contains a centralized server that stores garbage related data collected from different smart bins installed in different parts of the city. For the user layer we have two types of users: garbage collector and system monitor. Both of these users can access the garbage status data of various smart bins as required.

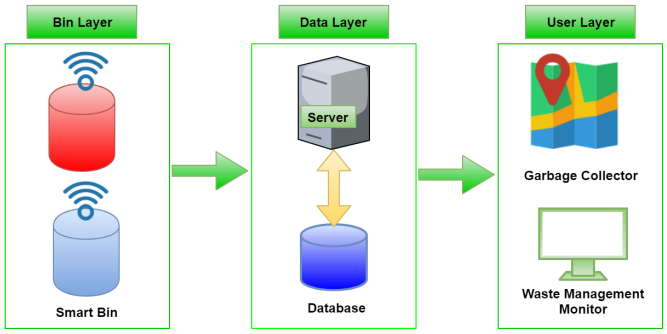


Fig. 1. Architecture of proposed IoT based green waste management system

Both the bin layer and the user layer are connected to the data layer for their operation. Our system is a one-way communication system which only collects data about the physical world. Bin layer directly sends its data to the data layer. Data layer stores information in the cloud and the user can collect information from the cloud whenever required. Garbage collector and system monitor individually collect information from data layer. The proposed smart bin depends on solar power for energy source. As a result they can be installed in any new places as long as solar energy is available. GPS location helps to locate the bins automatically from the cloud server. Following sections describe the system in details.

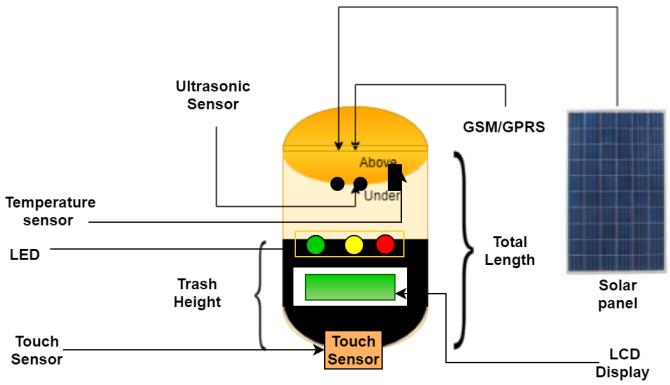


Fig. 2. System design of proposed smart bin

A. Bin layer

Bin layer of the proposed system consists of all the smart bins deployed in a target area. Every smart bin (Fig. 2) is built of a list of sensors and a microcontroller (e.g. Arduino AVR) that hosts the software to operate these sensors. In our system we have used Ultrasonic sensor, Temperature sensor and Touch sensor to measure different aspect about the smart bin. An ultrasonic sensor is used for measuring the height of the bin's waste level. The number of ultrasonic sensors will vary based on garbage bin's size. For example, a small bin can work with one sensor whereas a larger bin will require more sensors placed in different areas of the bin. Generally, ultrasonic sensor uses ultrasound wave that reflects, when gets to any barrier. We use this sensor to measure the level of the garbage in a bin from the following equation:

$$Garbage\ level(meter) = (Velocity\ of sound)/Time\ delay \quad (1)$$

We have used temperature sensor to sense internal temperature of the bin. Since, different waste has a different temperature, we can assume the types of waste. As we mentioned earlier every smart bin can be placed in remote places, we use a GPS module to know the location of our bin. A wireless module (e.g. GSM/GPRS) is used to transfer garbage observation data to the central server. Our proposed system has indicator LEDs to display the level of garbage. We have added a Touch sensor to open bin's lid automatically. And finally, a solar panel is added to the smart bin along with a

battery to store the energy. Capacity of the solar panel may vary considering the number of ultrasonic sensors based on the size of the bin. Every bin is assigned a unique ID. Whenever a new bin is added to the centralized system it gets initialized.

B. Data Layer

Data layer stores continuous garbage update in the cloud database. Every bin from the bin layer communicates with the central data layer and delivers the garbage status intermittently. Fig. 3 shows the ER diagram of the data layer. Every smart bin uses wireless module to communicate with the data layer. From the database, user can find the history information about a particular smart bin and its current and previous installations. Every bin has a mode: active and inactive. If an existing bin needs to be relocated then the mode changes to inactive and then the central control system informs the administrator of the system about this action. Data layer manages history and current data for both active and inactive smart bins.

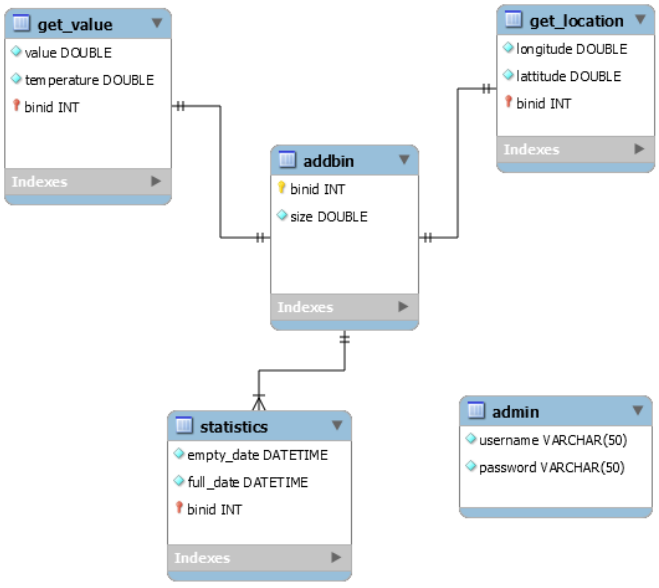


Fig. 3. Entity relationship diagram of the data layer

C. User layer

In our proposed system users are of two types: waste management system monitor and garbage collector. Following sections provide details about these user types.

- **Waste management system monitor:** This user is the administrator of the system. For each city authority there can be a list of monitors who will continuously monitor the current status of the overall system. An administrator can add, update and delete a smart bin from the central system. Also, statistical data is only available for the system monitor. Smart bins can be positioned as tags on online maps such as Google Map. System monitor can operate various types of data analytics on the history data, which will be provided to the city authority to make important decisions about waste management policy update.
- **Garbage collector:** In our proposed system, garbage collector is an user who is responsible to find the shortest route of available full garbage bins and pickup the garbage. This part of the system improvises the current model, where a garbage collection track travels to every bin whether it is full or not. In our model, only full/almost full smart bins are considered as candidates for pickup which reduces the energy cost and as well as keep the environment of the bins healthy.

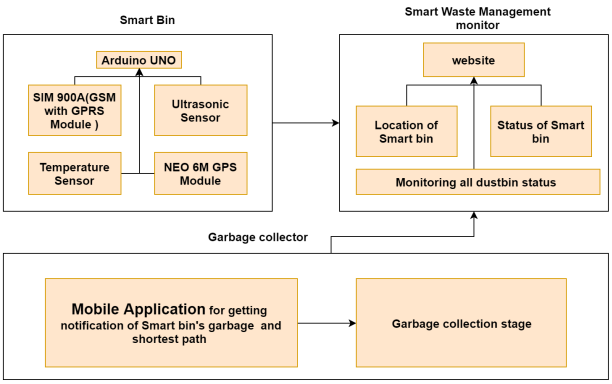


Fig. 4. Prototype application development procedure

IV. Prototype Implementation

In this section we describe the prototype implementation procedure of the proposed IoT based green waste management system. Fig. 4 details the implementation procedure of the proposed system. In order to prove the efficacy of the proposed model, we have developed a smart bin, a cloud based data server that provides user interface to both garbage collector and waste management system monitor. For the usage of garbage collector we also have developed an Android based mobile application. Location of the bins are visible in Google map that is available in the web application.

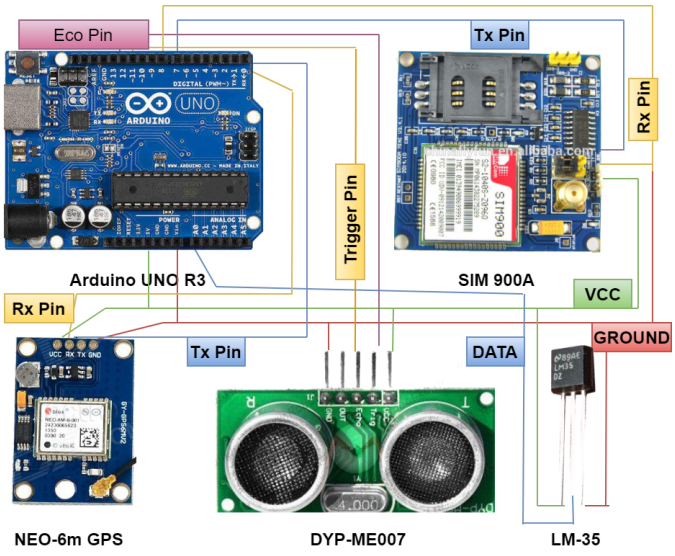


Fig. 5. Smart bin circuit diagram

A. Smart bin development

For smart bin development we have used the circuit diagram mentioned in Fig. 5. Our central processing unit is an Arduino UNO R3 board. A Ublox NEO-6m GPS module is connected with RX and TX pin of the Arduino UNO R3 board. On the other hand, SIM 900A GSM/GPRS Module is connected with a custom made RX, TX pin of the Arduino board. Ultrasonic sensor's trigger and echo pin are also connected with Arduino. The ultrasonic trigger pin is used for sending ultrasound wave and echo pin for receiving the signal. Temperature sensor's data pin is connected to the Arduino's analog pin. Temperature sensor works on voltage level, so we have connected it with an analog pin. All devices work in 5V DC power. A short description of each of this module is provided in the following sections.

- **Arduino Uno R3:** Arduino UNO R3 is an Arduino AVR based physical programmable circuit board, which is largely used for prototype hardware development. Arduino UNO R3 is based on ATmega328P microcontroller that provides 14 digital pins and 6 analog pins for hardware integration. Arduino board supports only DC current. As a result DC adapter is required, where recommended voltage is 9-12V DC, and minimum current supply is 1 Amp [4].

- **SIM 900A GSM/GPRS Module :** SIM modules are widely used for GSM and GPRS services. In Bangladesh SIM900A is easily available as it cheaper and more reliable. It works on two data frequency, 900MHz and 1800MHz. This device supports AT command and requires 5V DC and minimum 1Amp current supply. To prevent this module from restart more than 1 Amp current supply is recommended [9].

- **DYP-ME007 V2 Ultrasonic Sensor:** DYP-ME007 V2 ultrasonic ranging module is a high performance cost-effective non-contact distance measurement module with temperature compensation. It has two integrated receiver and sender. One is for sending ultrasound wave and the other for receiving. Measurement range for this device is 3 400cm, resolution 1cm and measures the target object without direct contact. Its operating voltage is DC 5V and operating current is only 10mA and response time is 100ms [1].

- **Ublox NEO-6m GPS Module:** The Ublox NEO-6M GPS engine has high precision binary output. It has also high sensitivity for indoor applications. The GPS module has a battery for power backup and EEP ROM for storing configuration settings. The antenna is connected to the module through a ufl cable which allows mounting the GPS in such a condition that the antenna can always have the open sky for best performance [12].

- **LM35 Temperature sensor:** The LM35 series are precision integrated-circuit temperature sensor that outputs a voltage which is linearly proportional to the Celsius (Centigrade) temperature. The LM35 sensor works on the analog pin. It receives voltage and returns voltage in bit pulse. Then this bit can be converted in Celsius or any units [6].

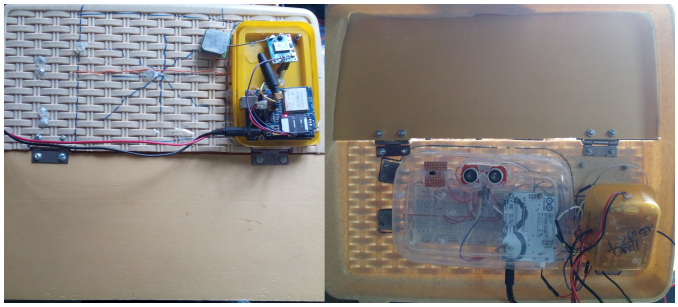


Fig. 6. Smart dustbin top and sensor view



Fig. 7. Smart dustbin side view

In our implementation of power supply we have used 12V, 9A dry cell battery and 10W solar panel. A 10W panel is enough to produce power for the system. When sunlight is available the system takes power directly from the solar panel. During insufficient sunlight, the system takes power from the battery. We have used dry-cell because it gives more efficiency than a lead-acid battery. In our test, battery back-up of our system was 19 hours, without solar light. In Bangladesh, we have sunlight approximately 11 hours in a day around the year [11]. To be fully charged up, the battery needs approximately 10 hours of sunlight. During the first boot, the system may take 15 to 30 seconds, depending on GPS signal. For having a better signal, we have placed it in an open field. Sometimes the ultrasonic sensor may provide some error data, generally 10 to 20 millimeters. For bigger size bin, it does not make any effect. In the test case, we have used a square bucket height

of 42 centimeters. We have also provided a lid on top of the bin. Fig. 7 shows the complete view of our smart bin.

B. Web application of waste management system

Our website provides full user interaction with the system. Anyone can see our website. Our website has two feature, one admin service and another general user. General user can see all Smart bin location in Google map shown in Fig 9. They can also see, individual Smart bin's status, temperature, and location. Anyone can search smart bin with an individual BIN ID. A user can also see individual smart bins last one month fill up and empty time and date. From this administrator can also monitor garbage collector efficiency. Administrator has features to Add Smart Bin, Delete or Edit if necessary. From smart bin fillup and empty statistics, admin can decide whether bin size needs to be changed. We used reverse geocoding to get location dynamically from longitude and latitude. So, if any Smart bin's longitude and latitude is changed, it automatically updates its location. We used Ajax to load page dynamically. Fig. 8 and Fig. 9 show our web application's home and map views. On the other hand Fig. 10 shows smart bin's live status.

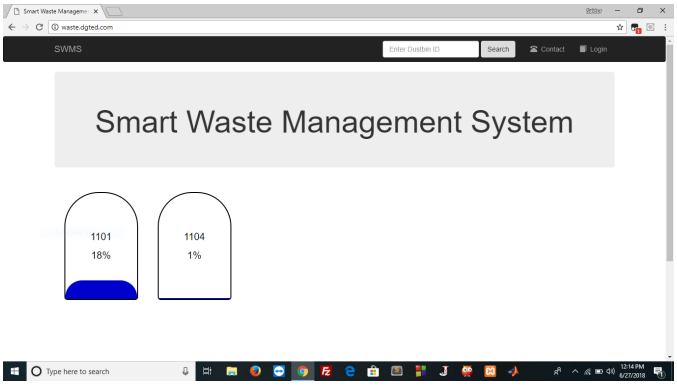


Fig. 8. Web application view

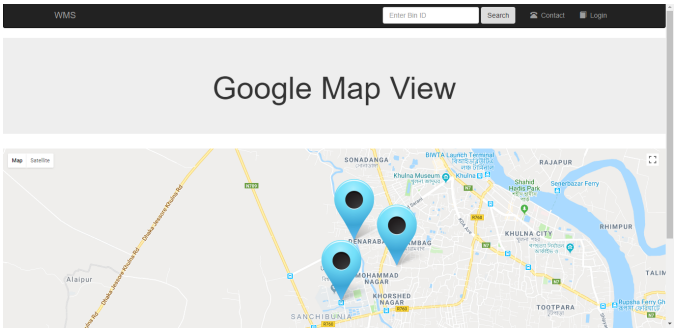


Fig. 9. Map view of smart bins

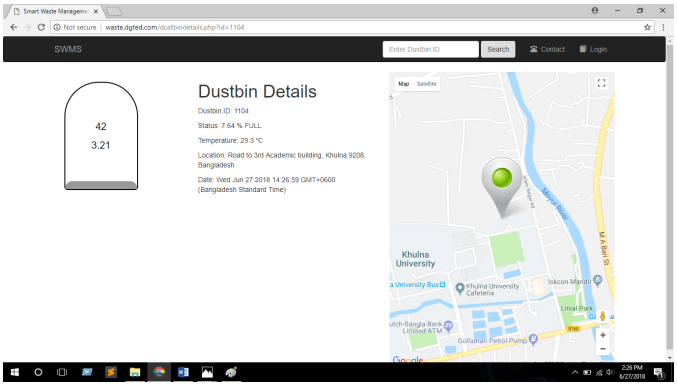


Fig. 10. Smart bin home page

C. Smartphone application

We have also developed android based smartphone application keeping garbage collector user role in mind. In our Android application we find a route to pickup all the full smart bins so that fuel efficiency is improved for the garbage truck. Generally Android app shows every smart bin's location with a marker in Google map. Marker color changes depending on Smart bin's waste status. It only generates the shortest path to the filled bins. Fig. 11 shows the pickup route of our deployed bins.

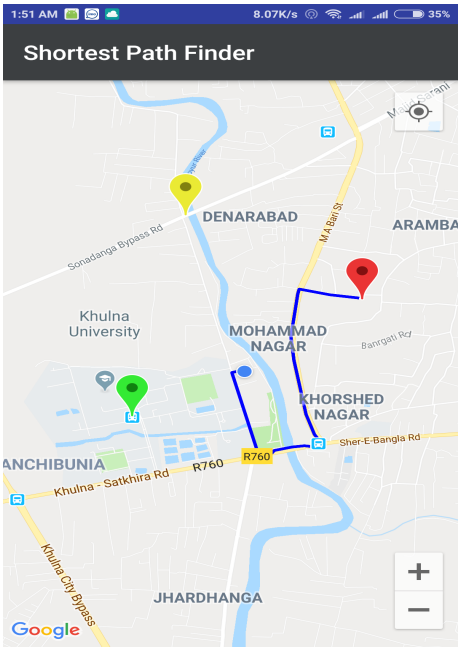


Fig. 11. Android based application for garbage collector

D. Limitations

Hardware implementation is critical challenge for the proposed system. The main issue is the scarcity of different required components and their quality of service. Specially in Bangladesh it is difficult to get many kinds of hardware components in due time. Some hardware components give faulty reading. For example, DYP-ME007 Ultrasonic sensor is

not capable of giving actual height. Sometimes it returns 10-20 cm error. SIM 900A module supports only two bandwidths. Also, Neo-6m GPS gives some error. Integration of the smart bin with the cloud based web application is also challenging due to time delay of the mobile operator's internet service.

V. Conclusion

Bangladesh is a developing country facing critical growth for city life expansion. More people are moving towards city based life and it is becoming difficult to plan waste management for the city authorities. With the application of IoT in the field of smart city management, we can now explore the avenues of smart waste management using the power of technology. In this paper, we propose an IoT based green waste management model, where smart bins equipped with sensors can continuously update the city central server about their garbage status. In our proposed model, smart bins use solar power for the power generation, which makes it more viable for Bangladesh based deployment. As city authority possesses current data about the overall garbage bin status of the city at any time, it can make efficient plan. We hope our efforts will help to eradicate waste management problems of different Bangladeshi cities.

References

[1] ALLNET. Dyp serial ultrasonic sensor, technical specification. https://www.allnet.de/fileadmin/transfer/products/111901.pdf. accessed 2018-09-20.

[2] R. Arshad, S. Zahoor, M. A. Shah, A. Wahid, and H. Yu. Green iot: An investigation on energy saving practices for 2020 and beyond. *IEEE Access*, 5:15667–15681, 2017.

[3] C. T. Association. Ces press release. https://www.ces.tech/News/Press-Releases/CES-Press-Release.aspx?NodeID=3f276d17-f619-4274-8ade-186d6b6e5adf. accessed 2018-09-20.

[4] djmatic. Arduino playground. https://playground.arduino.cc/Learning/WhatAdapter. accessed 2018-09-20.

[5] I. Hong, S. Park, B. Lee, J. Lee, D. Jeong, and S. Park. Iot-based smart garbage system for efficient food waste management. *The Scientific World Journal*, 2014, 2014.

[6] T. Instruments. Lm35. http://www.ti.com/product/LM35. accessed 2018-09-20.

[7] K. Monika, N. Rao, S. Prapulla, and G. Shobha. Smart dustbin-an efficient garbage monitoring system. *International Journal of Engineering Science and Computing*, 6(6):7113–7116, 2016.

[8] M. V. Patil and S. M. Gajbhiye. A review on internet of things based garbage bins detection systems. *International Journal of Science and Research*.

[9] R. Pujar. Difference between sim900 and sim900a gsm modems. https://www.raviyp.com/embedded/174-difference-between-sim900-and-sim900a-gsm-modems. accessed 2018-09-20.

[10] S. Raina, P. Gupta, A. Kaw, S. Pandita, and S. Parwana. Iot based smart garbage and waste collection bin. *International Journal of Engineering Science*, 10723, 2017.

[11] TIMEBIE. Daylight hours of bangladesh. http://www.timebie.com/sun/bangladesh.php. accessed 2018-09-20.

[12] R. N. Tutorials. Guide to neo-6m gps module with arduino. https://randomnerdtutorials.com/guide-to-neo-6m-gps-module-with-arduino/. accessed 2018-09-20.

[13] Worldmeters. Bangladesh population. http://www.worldometers.info/world-population/bangladesh-population/. accessed 2018-09-19.

Real-time Water Vehicle Monitoring and Tracking System for Bangladesh

Mohammad Salah Uddin and Md. Naim Khan
Department of Computer Science and Engineering
East West University, Bangladesh
{akash.bangla,naim1248}@gmail.com

***Abstract*—This paper presents a monitoring/tracking system for water passenger vehicle such as vessels or ferries. Bangladesh is a land of rivers. Each and every year peoples die due to accidents in water vehicles. The rescue teams are unable to reach there on time due to the communication barriers which causes huge damages in life. On the other hand, some captains of ferries/vessels do not maintain the weather notifications as well as overload the vessels before leaving the terminal, which are the main reasons for accident listed by BIWTA. In this regard, we have proposed real time monitoring/tracking system for water vehicle. The system (continuously) monitors the status(such as geo-location, load information, the orientation etc.) of the vehicles and sends reports to the ground/control station. As a result, the ground station will always be aware of vehicle status. The details design and implementation of proposed system are also discussed in this paper. The proposed system increases the safety by reducing the number/chances of accidents.**

***Index Terms*—Emergency Device, Water transportation, Monitoring and tracking, GPS tracker, Embedded Systems, internet of things etc.**

I. Introduction

BANGLADESH is a land of rivers that crisscrossed throughout the mostly flat territories of the country. Padma, Meghna and Jomuna are the three largest rivers of Bangladesh. It has about 24,000 km. of rivers, streams and canals that together cover about 7% of the country's surface. Water transportation is comparatively cheaper than road and air. So, Inland Water Transport (IWT) is a prime mode for the transportation of people as well as goods. Therefore, several thousands number of ferries/vessels, country boats are traditional transport which have been playing a key role in inland and coastal water for hundreds of years as a major form of transportation of this country. Though it supports in a wide range in transportation for general but it is claimed that it is notoriously dangerous due to the existence of a big number of accidental reports in this sector. Some accidental data are shown in Table-I.

According to Dhaka Metropolitan Police(DMP) and BIWTA(Bangladesh Inland Water Transport Authority), The main reason for accidents are:

- Unauthorized overloading with passengers or cargo;
- lack of skills and awareness of masters;
- Rough Weather;
- Sketchy design vessels (without following proper specifications);

Indeed, this is the suitable time to come forward to modernize and develop the inland water transport sector of Bangladesh. To do so we are presenting a system that will monitor the vessels/ferry or country boat to track the location continuously and notify to the authority for taking actions immediately. Thus, it may help to reduce the unwanted accidents and sufferings of the people in water transportation sector and save valuable lives.

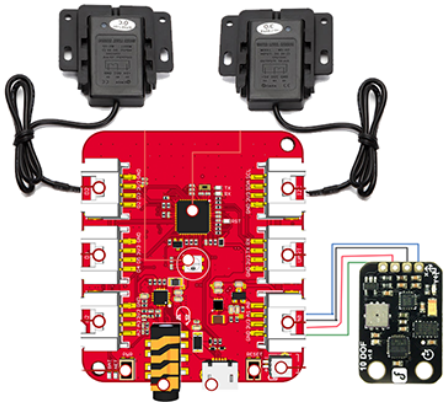


Fig. 1. The hardware design of the proposed system. The device is installed on ferries/vessels.

TABLE I
WATER VESSELS/FERRIES ACCIDENTAL DATA. SOURCE: [1]–[8].

Year	Name of the Vessels	Place of Accident
2017	Unpublished	Panguchi river, Bagerhat
2016	Unpublished	Padma river, Manikgonj
2015	MV Mostafa	Padma river, Paturia, Manikgonj
2014	ML Pinak-6	Meghna River, Mowa, Munshigonj
2014	MV Miraj-4	Meghna River, Rasulpur, Munshigonj
2014	MV Shathil-1	Ramnabad River, Patuakhali
2013	ML Sharos	Meghna River, Gozaria, Narayangonj
2012	MV Shariatpur-2	Meghna River, Gozaria, Munshigonj
2011	MV Bipasha-1	Meghna River, B-Baria
2009	ML Shahporan	Meghna River, Gozaria, Munshigonj

This paper is organized as bellow: The state of the art is presented in Section II. Section III deals with hardware requirements. The proposed system is described in Section

IV and result is presented in Section V.Finally, conclusion is given in Section VI.

II. State of the Art

The popularity of Internet of things(brief description is provided in section-VI) are increasing day by day in the area of remote monitoring system. The remote monitoring systems include, vehicle or assets monitoring, kids/pets monitoring, fleet management, parking management, water and oil leakage, energy grid monitoring etc.

Real time smart anti-theft vehicle monitoring system based on internet of things is presented in [9].The system allows users/vehicle owners to monitor their vehicle movements information (location log data). Google maps based mobile application has been used for visualizing the vehicle's movement. Additionally, NFC (near field communication) based small tag is used for proving extra layer of security.

Robi-tracker [10] and Finder [11] are other type of vehicle tracking system commercially available in Bangladesh. Those system are able to perform real-time location tracking of the vehicles. Some additional features such as geo-fencing, remote engine Off/On, speed violation etc. are also added with them.

Uber and pathao are most popular ride sharing system in Bangladesh. Driver and rider track their vehicle's movement/location in real-time by using a mobile application. A snapshot of mobile application is available in Fig. 2.

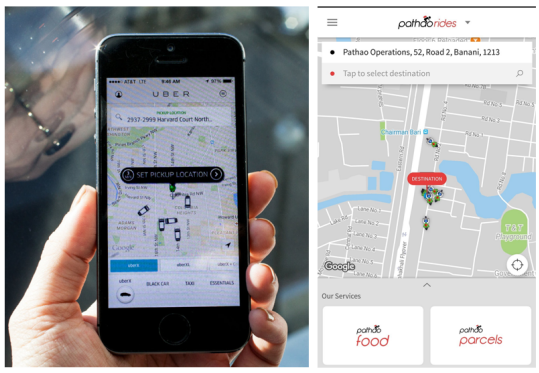


Fig. 2. Screen-shot of Uber(left) and Pathao(right) application.

A fixed route, simple bus tracking system has been proposed by Lau in [12]. The status of the bus after specified time interval has been provided using LED panel and a Smartphone application. The bus tracking systems are beneficial to college students who attend colleges with large campuses. They (students) can spend more time studying, sleeping, or relaxing rather than waiting for a delayed bus. Nowadays, almost all the European countries are using similar type of tracking system for public transportation. Some other embedded system based vehicle tracking system are presented in [13]–[15] Nowadays, traditional satellite based fleet management system has been using by malitries and airlines companies. Sky Connect(TM)Tracking System [16] one of the powerful tracking systems which has been used by these companies. Some developed countries are using intelligent transportation system for monitoring inland vehicles.

The above discussed system are only designed for ground vehicles. Those type of system are not perfectly suitable of water vehicle because of missing some additional information such as vehicle orientation, vehicle weight after loading,sinking level etc., which is important for water vehicle monitoring system. With this regard, this proposed system will played an important role for real-time communication for water vehicle.

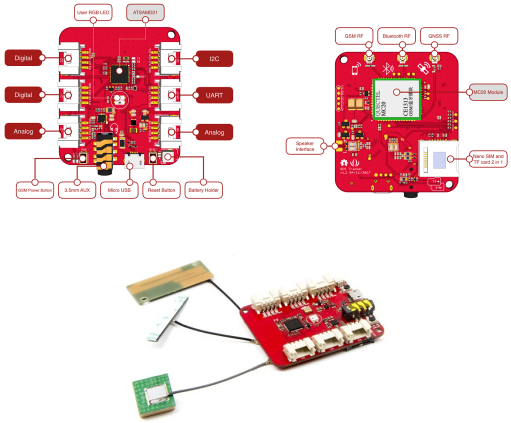


Fig. 3. Wio Tracker IoT Development Board. Source: SeeedStudio [17].

III. Hardware Requirements

A. Wio Tracker Development Board

Wio Tracker (Fig. 3) is an Arduino compatible IoT development board. The board was developed by SeeedStudio [17], a Chinese open source hardware manufacturing company. They have launched the Wio Tracker early in June 2017. The development board consists of 32-bit ARM Cortex-M0+ microcontroller. GSM & GPRS as well as GPS (Global Positioning System) & BeiDou [18], are also integrated with this board. SeeedStudio provides a full documentation of the board which is available in [19]. The board provides an all-in-one solution for tracking based IoT project.

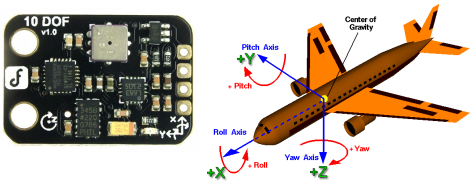


Fig. 4. The IMU sensor and the orientation measurement method.

B. Inertial Measurement Unit [20]

Inertial measurement unit (IMU) in Fig. 4 is used for determining the orientation of an object. IMU sensor are widely used in Air craft to determine the orientation. Its

also used in Human-computer interaction(HCI), Navigation and in the field of robotics. In this work, we used the IMU sensor for calculating the orientation of the water vehicle. The orientation information is useful for determining the stability of the vehicle. The stability information help us to decide whether the vehicle is in danger or not;which may reduce the risk of unwanted accident.



Fig. 5. The Water level sensor and its application.

C. Water level Sensors [21], [22]

Water level sensors are used for determining the presence of water or measuring the amount of water. Two types of water level sensors are available contactless and contactable. Contactable uses some copper wire which placed parallel to each other and contactless sensor uses hall magnetic effect for determining the presence of water. The details about contactless water sensor(in Fig. 5) are discussed in [21].

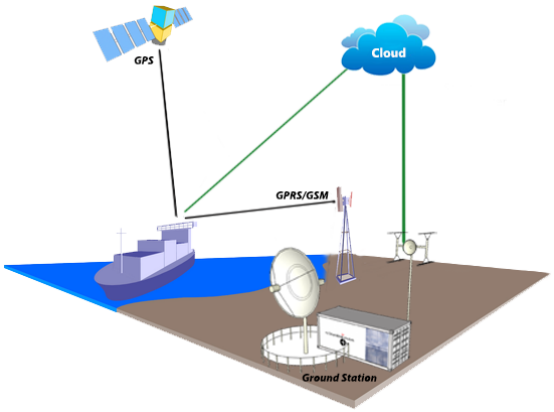


Fig. 6. Architectural view of our system. The Device(in-vehicle) is mounted on the vessels/ferries. It collects data(location, orientation and load information from sensors) and transmits them to the cloud by using GSM/GPRS or 3G/4G(if available) network. The ground station visualize those data from cloud.

IV. System Architecture

The architecture of our system is shown in Fig. 6 and the hardware design in Fig. 1. We choose Wio tracker open source development board as a main controller unit for this system. Multiple water level sensors are used here for determining the water level accurately. We calculate the load of the vessel/ferry by using displacement tonnage (details in [23]). Those information are helpful for determining the danger level of the water vessels/ferries. The water level sensors array and the IMU sensor are directly connected with development board. All the hardware/sensors for developing this system is discussed in section-III. A source program for controlling dataflow and managing the sensors data written in C compiled and flashed on microcontoller's (Wio Tracker) memory. A data SIM is mandatory for maintaining communication between device (in-vehicle) and IoT-cloud.

We have developed a desktop application for monitoring the status of the vehicle (vessels/ferries) in real time. The application collects data from IoT-cloud and visualize it. The snapshot of the prototype application is available in Fig. 7 and dataflow diagram shown in Fig. 8. The application provides all the information about the vehicle (such as Location, load info, orientation etc) through pop-up sliding notification. It also provide alert notification if any vehicle's is in danger (disconnected from the system or unstable orientation or overloaded information etc.). An user/admin account Id and credential is mandatory for accessing the monitoring system.



Fig. 7. The Desktop application for ground/control station. The application collects the data from cloud and visualize them continuously.

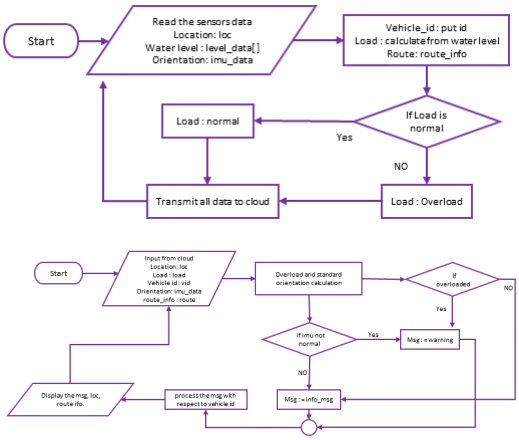


Fig. 8. The dataflow diagram of the system both device(top) and desktop application(down).

V. Result and Discussion

We have successfully tested our prototype version of the system. The resulting sensors data are shown in Fig.9. The

TABLE II
PAIRWISE COMPARISON TABLE.

	Cost	System Nature
Two Way Radio Communication	High	Manual
Cellular voice based	Low	Manual
Proposed System	Medium	Automated

development of the system requires financial support from local Government authorities.

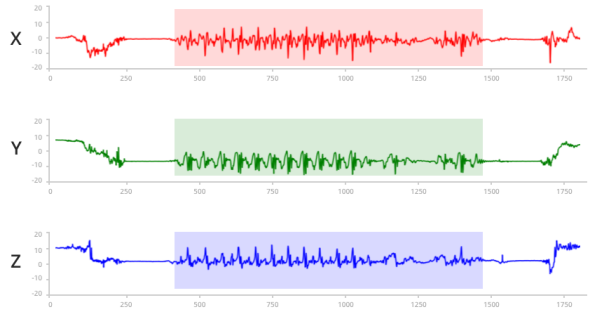


Fig. 9. The accelerometer data of in-vehicle deceive.

he Table.II depicted the pairwise comparison among existing system with our prototype system.

VI. Conclusion

River ferry capsizes in Bangladesh, killing people every year. Sinking of a capsized vessel which results in worst incidents or accidents are always unwanted and unexpected. Most of time vessels are overloaded, and they continue to move during rough weather without obeying the weather notification. Hundreds of people die each and every year in ferry accidents.

This paper presents a monitoring system which may reduce such incidents by continuously tracking the location of vessels and to keep informed about the situation so that any kind of emergency as well as causes that was marked for occurring accidents like collision and grounding to the authority. This aims to help from destroy of goods as well as lives of people drastically.

Appendix

What is IoT

The IoT or internet of things which is an emerging technology. IoT was first proposed by Dr. Ashton (MIT professor) in the middle of 1999. Professor Ashton discovered the IoT (also known as emerging technology)during his research study on RFID (Radio frequency identification). Any electronic devices or embedded system performs the communication among them through IoT. The most popular three layer architecture of IoT is presented in Fig. 10. IoT has three layers namely, ***Perception layer, Network layer and Application layer***. Electronic devices,embedded systems or things all are connected to the through perception layer. The main task of network layer is data transmission among things or devices. The bridge between IoT and human is established by application layer. More details technical descriptions along with applications domain of IoT are nicely listed in [24]–[26].

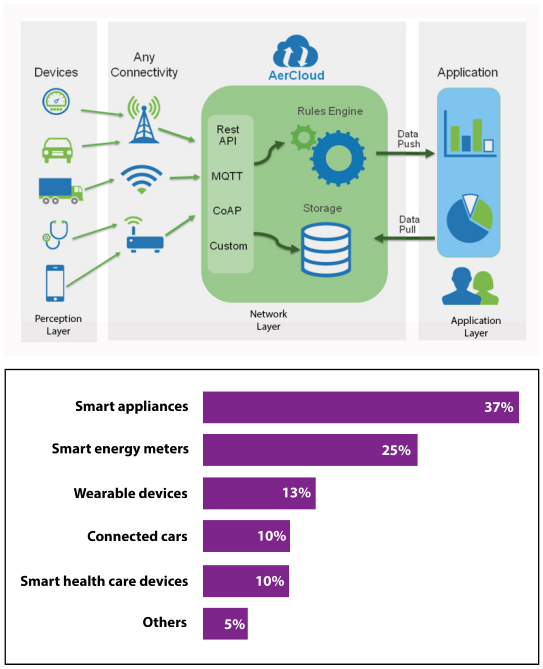


Fig. 10. The architecture of internet of things (top) and its application domain (button)(source:GSMA report [26]).

Acknowledgment

The authors would like to express their special thanks of gratitude to Bangladesh Inland Water Transport Authority for their historical data. Secondly they would like to thanks seeedstudio.com for their logistic support for building this prototype.

References

[1] BIWTA, "Annual report of biwta, 2016-2017," *Bangladesh Government Press*, 10 2017.

[2] N. Afrose Huq and A. Mahmood Dewan, "Launch disasters in bangladesh: A geographical study," *Geografia : Malaysian Journal of Society and Space*, 01 2006.

[3] "BANGLA NEWS 24," http://http://www.banglanews24.com, 2015, [Online; accessed 26-January-2019].

[4] "The Daily Prothom Alo," http://en.prothom-alo.com, 2015, [Online; accessed 02-January-2019].

[5] "The Daily Nayadiganta," http://www.dailynayadiganta.com, 2015, [Online; accessed 02-January-2019].

[6] "The Daily Juganttor," http://www.jugantor.com, 2015, [Online; accessed 02-January-2019].

[7] "The Daily Ittefaq," http://www.ittefaq.com.bd, 2015, [Online; accessed 12-January-2019].

[8] "The Daily Star," http://www.thedailystar.net, 2015, [Online; accessed 02-January-2019].

[9] M. S. Uddin, M. M. Ahmed, J. B. Alam, and M. Islam, "Smart anti-theft vehicle tracking system for bangladesh based on internet of things," in *2017 4th International Conference on Advances in Electrical Engineering (ICAEE)*, Sep. 2017, pp. 624–628.

[10] "Robi Tracker," https://www.robi.com.bd/en/personal/digital-solutions/robi-tracker, 2018, [Online; accessed 06-January-2019].

[11] "Finder-GPS tracking System," http://www.finder-lbs.com, 2015, [Online; accessed 12-January-2019].

[12] E. C.-W. Lau, "Simple bus tracking system," *Journal of Advanced Computer Science and Technology Research*, vol. 3, no. 1, 2013.
[13] S. Arslan, M. Gündüzalp, and E. Türk, "An embedded system vehicle tracking system application," in *Electrical, Electronics and Biomedical Engineering (ELECO), 2016 National Conference on*. IEEE, 2016, pp. 447–451.
[14] M. M. Rahman, J. R. Mou, K. Tara, and M. I. Sarkar, "Real time google map and arduino based vehicle tracking system," in *Electrical, Computer & Telecommunication Engineering (ICECTE), International Conference on*. IEEE, 2016, pp. 1–4.
[15] O. A. Mohamad, R. T. Hameed, and N. Ţăpuş, "Design and implementation of real time tracking system based on arduino intel galileo," in *Electronics, Computers and Artificial Intelligence (ECAI), 2016 8th International Conference on*. IEEE, 2016, pp. 1–6.
[16] "UK aerospace equipment, materials and maintenance sectors brochures," *Aircraft Engineering and Aerospace Technology*, vol. 75, no. 1, feb 2003. [Online]. Available: https://doi.org/10.1108%2Faeat.2003.12775aab.052
[17] " SeeedStudio ," https://www.seeedstudio.com, 2012, [Online; accessed 04-January-2019].
[18] "BeiDou Navigation Satellite System," https://en.wikipedia.org/wiki/BeiDou_Navigation_Satellite_System, 2000, [Online; accessed 04-January-2019].
[19] "Wio Tracker - GPS, BT3.0, GSM, Arduino Compatible," http://wiki.seeed.cc/wio_gps_board, 2017, [Online; accessed 04-January-2019].
[20] "Inertial Measurement Unit," https://en.wikipedia.org/wiki/Inertial_measurement_unit, 2007, [Online; accessed 06-January-2019].
[21] "Contactless Water Level Sensor," https://www.banggood.com/Contactless-Water-level-switch.html, 2014, [Online; accessed 06-January-2019].
[22] "Non-contact Liquid Level Switch," https://www.dfrobot.com/wiki/index.php/Non-contact_Liquid_Level_Switch_SKU:FIT0212, 2016, [Online; accessed 06-January-2019].
[23] "Displacement (ship)," en.wikipedia.org/wiki/Displacement_(ship), 2003, [Online; accessed 08-January-2019].
[24] A. Whitmore, A. Agarwal, and L. Da Xu, "The internet of things—a survey of topics and trends," *Information Systems Frontiers*, vol. 17, no. 2, pp. 261–274, 2015.
[25] S. Li, L. D. Xu, and S. Zhao, "The internet of things: a survey," *Information Systems Frontiers*, vol. 17, no. 2, pp. 243–259, 2015.
[26] GSMA, "Gsma: The impact of the internet of things - the connected home," KRC RESEARCH, Tech. Rep., 2012.

Development of An Android Based Real Time Bus Tracking System

Mohammad Nazmul Hasan and Md. Sharif Hossen

Department of Information and Communication Technology
Comilla University, Comilla, Bangladesh
nazmulhasan5793@gmail.com, sharif5613@gmail.com

Abstract—In this dynamic life, everybody is in an urgency to reach their destinations. In this case, waiting for the buses is not feasible. People who depend on the public transport system should know about the location of the bus for which they are looking for. With the advancement of technology, android smartphones have become universal and affordable for all. Smartphones have vast capabilities to provide rich user experiences with interactive facilities. In this paper, we propose a real-time bus tracking system for the students using Android smartphones which can be used for tracking and positioning of the buses by using Global Positioning System (GPS). This android based real time application enables the students to find out the exact location of the buses so that they will not get late or will not arrive at the bus stop too early. It provides the exact location of the student's respective buses which can be viewed on the google map. In addition, this application also gives the information like bus details, driver details, contact numbers, routes, etc. Therefore, our developed application saves the student's time to wait for the buses as they can know their current location of the buses updated every moment in the form of latitude and longitude on Google maps.

Keywords— GPS, google map, android, real-time, vehicle tracking

I. INTRODUCTION

Nowadays, buses are used rapidly as a transportation system in the different organization, e.g., school, college, university, business firm, banks, etc. But, with the increase of population, vehicles are increasing day by day which is further leading to heavy traffic. There are many reasons for heavy traffic in developing countries like Bangladesh which results in more delay to reach at the destination of the respective buses [1]. These reasons may be harsh weather, traffic conditions, etc. [2]. On the other hand, at the bus station, most of the time people have to wait for the arrival of buses because of the delay of a particular bus at a particular stop [3]. Therefore, an effective transportation system has effective movement of vehicles and people which leads to better quality of life and better economic growth of the society. Hence, the bus tracking system can be useful for passengers which will provide the location of the buses with routes using Google map at the current time [4]. Embedding the equipment in the buses to track those is important which increases the cost of the system. Hence, we use IoT based technologies to reduce the costs of processing and communication equipment [5]. Today smartphones are a very good source for tracking the vehicles (i.e. buses in this case), called mobile tracking system [6], because of giving the real-time location of the buses with the help of Global Positioning System (GPS) [7]. GPS uses satellites to locate the presence of smartphones using the values of longitude and latitude. Using those values of GPS coordinates our developed application will show the location of the smartphone used in each bus carried by the driver [8]. Users (i.e. students, teachers, etc.) can retrieve information relevant to the arrival time of buses with respect to the current location by selecting the bus number through the android application [9]. The main objective of this research is to notify the locations of the buses from anywhere for the students so that they do not need to wait for their buses at the stops for a long time and save their valuable time. To complete this, we develop an android application and demonstrate the location of the bus at the particular time.

The remainder of this paper is arranged as follows: Section II discusses the related works with the proposed system. Section III gives a brief of the technologies used in the proposed system. Section IV discusses the architecture of the proposed system and its specifications. Section V includes the step by step experimental analysis and discussion. Section VI summarizes the conclusion with future works.

II. RELATED WORKS

Bus tracking systems are developed to display the exact location of buses which can be useful for the public transport system. The system requires an internet connection and may or may not be GPS tracker. Real-time bus tracking system using GPS displays the current locations of the bus which consists of a GPS tracker, real-time server, and the user application. It provides relevant bus routes and other information to the user. Bus tracking system uses the built-in GPS service provided by the Android. A smartphone is used on each bus and gets its GPS coordinates. These coordinates are shifted to the server. Users can retrieve information through their Android application. Then, selecting the respective bus number, users receive the arrival time of the bus with respect to the user's current location. Google map is also used to graphically plot the bus location [10] [11].

There are about 17,000 location-based applications on the marketplace. Some apps can make our traveling a little easier, a bit more fun and more memorable. But, most of the applications are hardware dependent. There is a fact of high cost as the devices they use are costly and there is more complexity in the integration of those devices in a single unit [12] [13]. To fix such problem is also costly and time-consuming. In this research endeavor, we develop a real-time android based bus tracking system which is hardware independent, less costly and easy to install.

III. Technologies Used

A. GPS Technology

GPS [10], also known as a space-based radio navigation system, consists of 24 satellites orbiting this earth under the management of the US government. It was primarily developed for the military purpose and later used to help general people for various applications, e.g., navigation, tracking, timing and positioning services [14] [15].

B. Android Platform

Android is a Linux based open source operating system which was founded by Andy Rubin, et al [16]. There are some important features that make android different from other operating systems [17]. These are as follows:

- Open Source
- Easy to customize the Android platform
- Support the multiple inheritance since it uses the concept of Java [18]
- Allow all graphic dimensions, i.e., 3D and 4D
- Support the low-level interface for having a Linux operating system kernel
- Support standard libraries like OpenGL, Web Kit and SQLite, etc.
- Appealing look-and-feel features including live Really Simple Syndication (RSS) feeds, weather information, etc.
- Support many more connectivity and messaging services like Bluetooth, GSM, Wifi, MMS, SMS, CDM, etc.

C. Firebase Server

Firebase is a web and mobile application developing platform that provides tools and services to develop real-time apps. It was developed in 2011 by Firebase Inc. and in 2014 acquired by Google. It provides a backend and real-time database as a service. It is a cloud-hosted database system [19] [20].

D. Google Map

Google map is a free software that renders 3D graphics of earth using satellite images around the world. It is a version of Google earth that shows the maps and can be embedded into web pages through Google maps API [21].

IV. System Design and Specifications

Our proposed and developed android based system provides the students to find out the exact location of the buses from anywhere. The bus routes are shown on the user interface displaying the position of the buses using google map. Accordingly, user can plan and start. The advantages of the developed system are to:

- Provide the exact position on Google map
- Show the details of the buses from anywhere
- Provide details like bus number, drivers' contact number, bus route, stops, etc.

Figure 1 shows the architecture of the developed system. The system consists of a client and a server interface. The developed android application is used on both driver and user module. The user can track the location of buses controlled by a driver on Google map from the server using GPS. Admin maintains and updates all relevant information like bus routes, driver number, number of buses, etc. on the server.

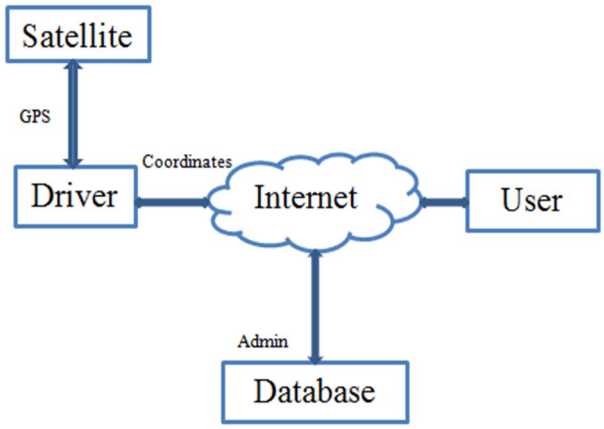


Fig. 1. Block Diagram of System Architecture

The developed system contains the three modules:

a) Driver
b) User
c) Admin

a) Driver Module

This module is designed for the bus driver. The authorized bus drivers can use this module by providing their unique login credentials. They need to start their location services before driving. The current location of the bus will automatically be updated from the driver's mobile to the server every moment in the form of latitude and longitude. Figure 2 shows the flow diagram of the driver module.

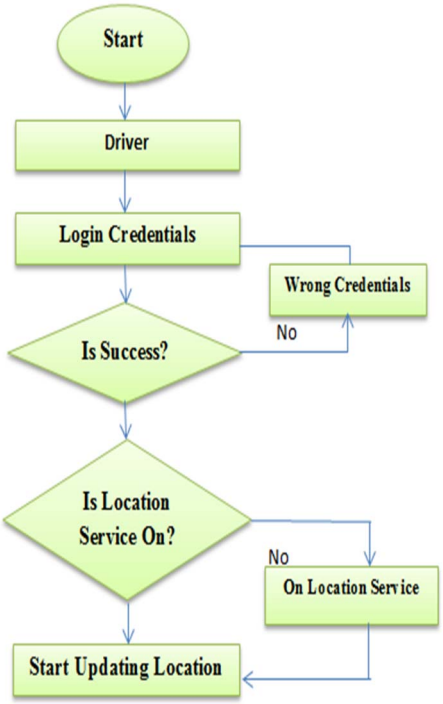


Fig. 2. Driver Module

b) User Module

This is the most important module and the soul of the system. Users of this module need to click on the user button. They can access the details of all the buses through their smartphones. Here, they will get all the buses and driver-related information offline too. Students can track the location of their buses from any location. A student must make sure that their location service is active. Figure 3 represents the flow diagram of user module.

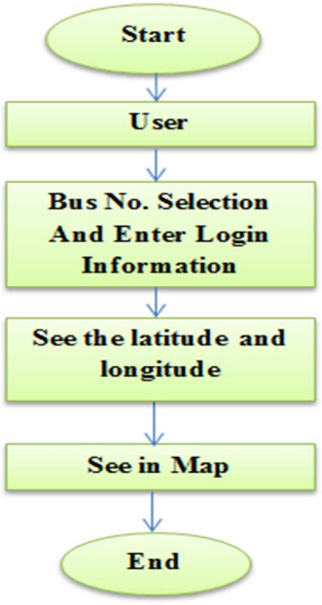


Fig. 3. User Module

c) Admin Module

Figure 4 shows the flow diagram for the admin module. This module is designed for the bus administrator for updating the information. Admin can log in to the admin account after authentication and authorization. He can enter new route details and also he has the options to add or remove a route. The authority has the right to update the driver's name, driver's contact number, routes, stops, etc. If the admin wants to send any information to the driver, then he can send the message to the driver's mobile.

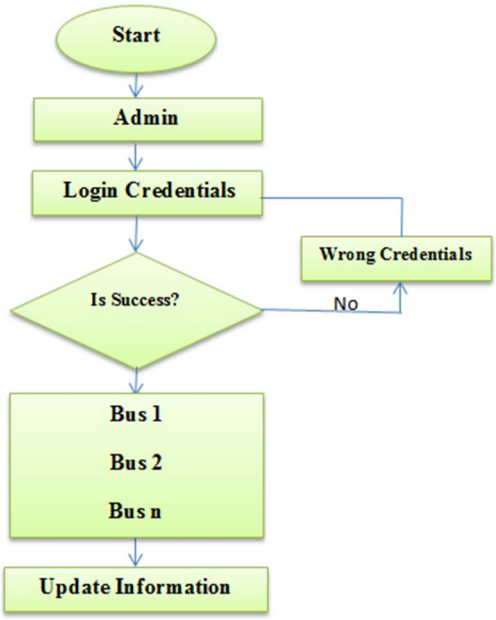


Fig. 4. Admin Module

V. Experimental Result and Discussion

When a driver enters into the driver module, Figure 5 will be appeared for the driver to enter his login credentials. The location of the bus will be shown on the screen as latitude and longitude and it will automatically be updated in every 5 seconds (shown in Figure 6).

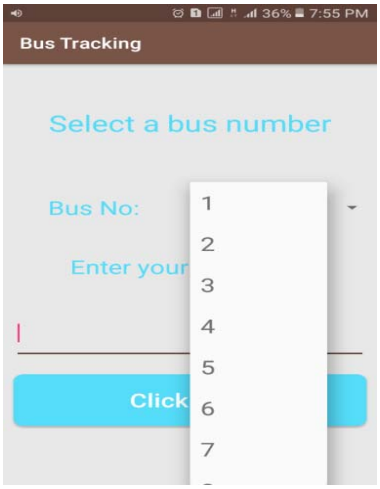


Fig. 5. Driver's Bus Selection



Fig. 6. Bus Location

To search for a bus location a user has to enter the bus number and login credentials. Then, the map will be displayed which will show the current location of the bus. They can get access to the details of all the buses through their mobile phones. Here, they will also get the bus and driver-related information as shown in Figure 7.



Fig. 7. Drivers' Details

When the application is launched, home activity shows the user module. When a user selects this module, another activity will be opened where the user will select the bus number and enter the login credential shown in Figure 8. Then, the map will be displayed which will show the current location of the bus as shown in Figure 9.

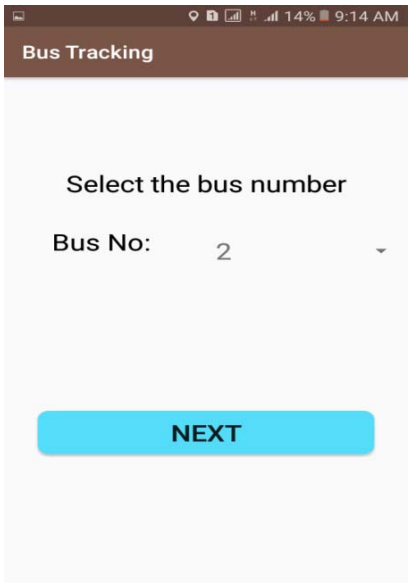


Fig. 8. User's Bus Selection



Fig. 9. Bus Location on Map

VI. Conclusion and Future Works

In this research study, we design and develop a real-time bus tracking system using GPS tracking technology which needs only a smartphone and a real-time server. Our application consists of two fundamental concepts: first it collects the real-time location information of buses via GPS technology and secondly updates the location information in the database server. The bus-side, server-side and client-side modules provide all the expected functions. Since this application does not need any external hardware except a smartphone which is available to anyone in the world, the overall cost is very low or no cost needed for tracking the bus location. It provides nearly accurate data in real time that makes possible for the user to track the buses.

In near future, we would like to enrich our proposed application by adding the following features.

- Designing a reliable server with a user friendly API
- Distance and time calculation between two places
- Anti-theft system for cars and bikes
- As a vehicle management software for transport companies

REFERENCES

[1] P. Chaiprapa, S. Kiattisin, and A. Leelasantitham, "A Real- Time GPS Vehicle Tracking System Displayed on a Google-Map-Based Website," University of the Thai Chamber of Commerce Journal, 2011.

[2] P. Zhou, Y. Zheng, and M. Li, "How Long to Wait? Predicting Bus Arrival Time with Mobile Phone Based Participatory Sensing," Transactions on Mobile Computing, IEEE, vol. 13, no. 6, June 2014.

[3] S. A. E. Yosif, M. M. Abdelwahab, M. A. E. ALagab, and F. Muhammad, "Design of Bus Tracking and Fuel Monitoring System," In Proc. of International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), IEEE, Khartoum, Sudan, 16-18 January 2017.

[4] P. W. Buana, I. Sukarsa, I. B. Purwania, and I. G. Prasetya, "Real Time Trans Bus Tracking and Passenger Information System using Hybrid Application Technology," International Journal of Software Engineering and Its Applications, vol. 10, no. 9, pp. 35-50, 2016.

[5] M. Nieto, L. Unzueta, J. Barandiaran, A. Cortés, O. Otaegui, and Pedro Sánchez, "Vehicle Tracking and Classification in Challenging Scenarios via Slice Sampling," EURASIP Journal on Advances in Signal Processing, Springer, 2011.

[6] M. Abdurohman, A. Herutomo, V. Suryani, A. Elmangoush, and T. Magedanz, "Mobile Tracking System Using OpenMTC Platform Based on Event Driven Method," In Proc. of International Workshop on Machine to Machine Communications Interfaces and Platforms, IEEE, 2013.

[7] T. H. Win, and Z. M. Aung, "Vehicle Navigation System Using GPS Receiver," In Proc. of International Conference on Advances in Electrical and Electronics Engineering, 2011.

[8] O. A.Elrhman, A. M. Ahmed, T. H. Mekki, and G. M. Taha, "Android Based Vehicle Tracking System," In Proc. of Advances in Business Research International Conference, IEEE, Langkawi, Kedah, Malaysia, pp. 17-18 November 2015.

[9] M. Ghareeb, A. Ghamlous, H. Hamdan, A. Bazzi, and S. A. Nabi, "Smart Bus: A Tracking System for School Buses," In Proc. of International Conference on Sensors Networks Smart and Emerging Technologies, IEEE, Beirut, Lebanon, 12-14 September 2017.

[10] M. A. A. Rashed, O. A. Oumar, and D. Singhm, "A Real Time GSM/GPS Based Tracking System Based on GSM Mobile Phone," In Proc. of International Conference on Future Generation Communication Technologies (FGCT 2013), IEEE, London, UK, March 2014.

[11] M. Kahvecı, "Recent Advances in GNSS: Technical and Legal Aspects," In Proc. of International Conference on Recent Advances in Space Technologies (RAST), IEEE, Istanbul, Turkey, June 2017.

[12] A. Mulla, J. Baviskar, A. Baviskar, and A. Bhovad, "GPS Assisted Standard Positioning Service for Navigation and Tracking: Review & Implementation," In Proc. of International Conference on Pervasive Computing (ICPC), IEEE, Pune, India, January 2015.

[13] H.Y. Yu, Y.Y. Wang, and Z.F. Liu, "Research on Android-Based Data Management System for Fitness Equipment," In Proc. of International Conference on Chinese Automation Congress (CAC), IEEE, Jinan, China, January 2018.

[14] A. Puscasiu, A. Fanca, and H. Valean, "Tracking and Localization System using Android Mobile Phones," In Proc. of International Conference on Automation, Quality and Testing, Robotics (AQTR), IEEE, Cluj-Napoca, Romania, May 2016.

[15] K. Tan, and K. Wong, "Low-Cost Campus Bus Tracker Using Wifi Access Points," In Proc. of International Conference on Consumer Electronics-Taiwan (ICCE-TW), IEEE, Nantou, Taiwan, May 2016.

[16] P. A. Shinde, Y. B. Mane, and P. H. Tarange, "Real Time Vehicle Monitoring and Tracking System Based on Embedded Linux Board and Android Application," In Proc. of International Conference on Circuits, Power and Computing Technologies (ICCPCT), IEEE, Nagercoil, India, 2015.

[17] M. Behzada, A. Sanab, M. A. Khanc, Z. Walayatb, U. Qasimd, Z. A. Khane, and N. Javaida, "Design and Development of a Low Cost Ubiquitous Tracking System," In Proc. of International Conference on Future Networks and Communications (FNC), Elsevier, Canada, 2014.

[18] M. S. Hossen, "A Java Based GUI Application for Substitution Encryption Techniques," In Proc. of International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), RU, Bangladesh, January 2017.

[19] S. Lee, G. Tewolde, and J. Kwon, "Design and Implementation of Vehicle Tracking System Using GPS/GSM/GPRS Technology and Smartphone Application," In Proc. of World Forum on Internet of Things (WF-IoT), IEEE, Seoul, South Korea, March 2014.

[20] M. M. Rahman, J. R. Mou, K. Tara, and M. I. Sarkar, "Real Time Google Map and Arduino Based Vehicle Tracking System," In Proc. of International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), IEEE, Rajshahi, Bangladesh, December 2016.

[21] N. Chadil, A. Russameesawang, and P. Keeratiwintakor, "Real-Time Tracking Management System Using GPS, GPRS and Google Earth," In Proc. of International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, IEEE, Krabi, Thailand, August 2008.

An IoT based mHealth Platform for Chronic Kidney Disease Patients

Alaa Noor, Anonnya Banerjee, Mayeesha Fairooz Ahmad, Mohammed Nazim Uddin
School of Science, Engineering & Technology, East Delta University
Chattogram, Bangladesh
Email: alanoor262@gmail.com, anonnyabanerjee95@ymail.com, mayeesha.f.a@gmail.com, nazim@eastdelta.edu.bd

***Abstract*—Medicinal care signifies as a standout amongst the most striking application ranges for the IoT. Hospitals utilize new advances and IoT to screen the situation of patient's therapeutic conditions and other related data. There are vast IoT applications that are being readily used in the medical care nowadays from. It can possibly keep patients safe and sound, as well as it will enhance how doctors provide care. Thermometers, electrocardiograms, ultrasounds, glucose monitors, and more are altogether beginning to end up associated and giving patients a chance to track their wellbeing. This is significant for those circumstances that require line up meetings with specialists. In this study, we presented a preliminary healthcare system for the treatment of Chronic Kidney Disease (CKD) or Chronic Renal Disease patients with the concept of IoT, which includes patient log, the nutrition of the food basically focusing on the level of salt intake, monitoring the sleep pattern, activity level, water intake and so on in order to provide with required improvements needed for the betterment of their health condition. In addition, the data and information obtained from the patient log and from their activities will stored in cloud for future purposes and simultaneously it can be accessed directly by the caregivers for monitoring of a patient's health condition.**

Keywords— ***Internet of things; Chronic Kidney Disease; HealthCare system, CNN; Blood potassium level***

I. INTRODUCTION

Internet of Things (IoT) is a method for interfacing gadgets to the Internet and to one another utilizing remote innovation. This enables engineers to add innovation-based highlights to physical gadgets including the capacity to convey encounters continuously. It is a prominent articulation, which is conjectured as a standout amongst the most encouraging advancement drivers for automation and control in the business. The Internet of Things (IoT) has moved toward becoming the industry-skeptic language to portray how innovation is as of now been inserted in numerous business sectors, including medicinal services, and developing the manner by which business is directed. By 2020, it is projected that 200 billion gadgets will be associated, which measures up to 26 smart gadgets for each individual on Earth. The cumulative market capability of IoT invention will reach at its peak with 38% of the cumulative value originating from gadgets in health care [2]. The IoT has several applications in medical healthcare, from remote intensive care to smart devices or sensors and medicinal device incorporation. Moreover, IoT can properly recognize ideal occasions for restocking supplies for various gadgets for their ceaseless and smooth activity**.** It has the possibility not only retain patients healthy and safe, but to expand how medical doctor provide care additionally. Healthcare Internet of Things can also enhance patient satisfaction and engagement by providing patients with more interactive time with their consulted doctors. The capability to track patient's location and therapeutic equipment by means of sensor-based expertise has shaped a service of its own that is currently known as LaaS (location-as-a-service). Few of the instances consist of tracking of therapeutic devices that is used to examine patient's condition and their location with the help of geo-tracking in order to provide emergency medical aid which do leave a huge impact for the patients in need. Consequently, imaging, diagnostic devices, and various therapeutic sensors can be perceived as smart things establishing a fundamental factor of IoT. Healthcare built using IoT features is projected to reduce expenditures, increase the duration of life and improve the involvement of users. The added benefit of healthcare is bulk amount of statistical data. But the security of those collected statistical data must be ensured.

However, in developing and developed states, chronic diseases are rapidly growing such as cardiac diseases, diabetes and renal or kidney disease. 30 million of adults are projected to be diagnosed with CKD (Chronic Kidney Disease) and furthermost most of them are not aware of it. Risk influences for renal disease consist of high blood pressure, diabetes, obesity, heart disease and family background [1]. The administration and analysis of Chronic Renal Failure or kidney disease are complex. The hazard related with the absence of knowledge, weak management, and impassive way of life can be considered as essential explanations behind growing certainty of chronic kidney disease on the planet. Chronic kidney disease is escalating the damage role of kidneys over interval of years or months. CKD is determined by the occurrence of damaged or decreased kidney, glomerular filtration rate (GFR), where GFR < 60 ml/min/1.73 m2 for three months or so. Such deterioration is difficult to excess and waste fluids development in body and has an impact on its performance, possibly projecting to complications. It remains undiagnosed and overlooked until the state of the patient becomes worse gradually over time. Nevertheless, in order to have a healthy lifestyle, an individual need to have appropriate information about sustenance, a realization of hazardous diseases and so on. Accomplishment of such system needs an active contribution and involvement of users.

In this study, we proposed a healthcare system that will be beneficial for monitoring the health condition of chronic kidney disease patients with the help of smart devices such as wearable's. Section II consists of background study and related works. The design and the construction of Healthcare System for CKD patients have been delineated in Section III. The investigational design has been studied in Section IV. Section V illustrates the outcomes or results and finally the paper draws the conclusion in section VI.

II. Related Work

eHealth (Electronic Health Record) can be treated as a part of biomedical engineering, which is technologically advanced to help therapeutic procedure to expand the value of life by improving public healthcare [3]. In paper [4], the author correspondingly projected that if IoT is applied in the eHealth then will play as the leading role in the upcoming future.

Numerous researchers have endeavored to execute IoT for health care utilizing Raspberry Pi [5] [6]. Some study on Internet of Things (IoT) for healthcare tried implementing the system to reduce the morality rate of mother and an infant. In India, a system was developed known as Mother and Child Tracking System (MCTS), which was used to simplify universal access to such health services. It provides and help the healthcare providers with their facilities. In [7] the developers build a system based on smartphone for remote real time tele-monitoring of patient's physical activity with chronic heart failure (CHF). The tele-monitoring system was designed to measure patient activity minute-by-minute and energy expenditure (EE) estimated from accelerometer. In addition, the system tracked relative user location via global positioning system (GPS) to track outdoors activity and measure walking distance and also administrated self-reported vital and relevant cardiovascular symptoms. However, the drawback of the system was that it consumed a lot of battery life as some execution was done in the phone and the data was being collected from many sensors, in result it had an effect on usability.

Convolutional Neural Network (CNN), which is a special architecture of artificial neural network is now widely used for image classification. For the work in [8], the authors have used Deep Convolutional Neural Network to detect object with the data size of 1650 food images and epochs of 25. They had achieved 60.74% of accuracy.

The authors in [9], designed a mobile health (mHealth) self-control of diabetes that supports real-time clinical communication and response personalized to the individual necessities of the diabetes patient, using current and past patient records. Physical layer nodules are associated to a Network centered application layer over a present telecommunication infrastructure. It connects the several objects to other objects in the physical layer. The application units, that handles all clients associated functionality, is planned to be adaptable with the Model-View-Control (MVC) outline. Depending on the gathered readings, the cell phone delivers the support and required opinion in measuring the necessary insulin bolus. The researchers at [10] proposed a wireless living assistance IoT based system for elder people to monitor and register patient's important information to provide mechanisms for triggering alarms in case of emergency situations. The research suggests a system including wrist band which can be connected to a cloud server for assisting and monitor elderly people. The researchers claimed it as cost efficient with low power consumption system having wireless communication features.

III. Health care System for Chronic Kidney Disease PATIENTS

The framework in fig 1. uses MQTT, which is a frequently used protocol in IoT projects, and it stands for Message Queuing Telemetry Transport, as it is a lightweight messaging protocol which utilizes operation to interchange data between clients and the server. Additionally, less power consumption, optimal data packets, and simplicity of implementation contributes in the betterment of the health care system. Similar to internet protocol, MQTT is based on client and the server, where MQTT acts as the server that is use to handle the client's requests of sending or receiving data between the connected devices. Thus, it is used to establish communication between patients and care providers in order to provide an affluent healthcare facility.

The proposed framework makes it possible to provide superior medicinal services management with a scope of various modules. Our framework lies on end clients who are utilizing a mobile application, a cloud as a service and a web-based interface for health care system. The majority of the clients are associated with the end goal to keep up an astounding correspondence benefit. The framework includes various modules such as: - *(1)* Patients log to record one's data, *(2)* Sleep pattern module to uphold the sleeping cycle of the patient as it plays as a vital role amongst the CKD patients, *(3)* Water intake that will notify the patient to stay hydrated, *(4)* Food classification which will provide patients with the nutriments of the food in order to precisely check on the amount of potassium, sodium and other nutrients that he or she intakes on a regular basis as the intake of salt, *(5)* Patient's physical activities which is used to ensure that a patient is not living an inactive life and this is done by the movement detection with the help of smartphone sensors or a wearable. The majority of the previously mentioned modules will be capable enough to deal with the medicinal services administration of a CKD patient in the near future.

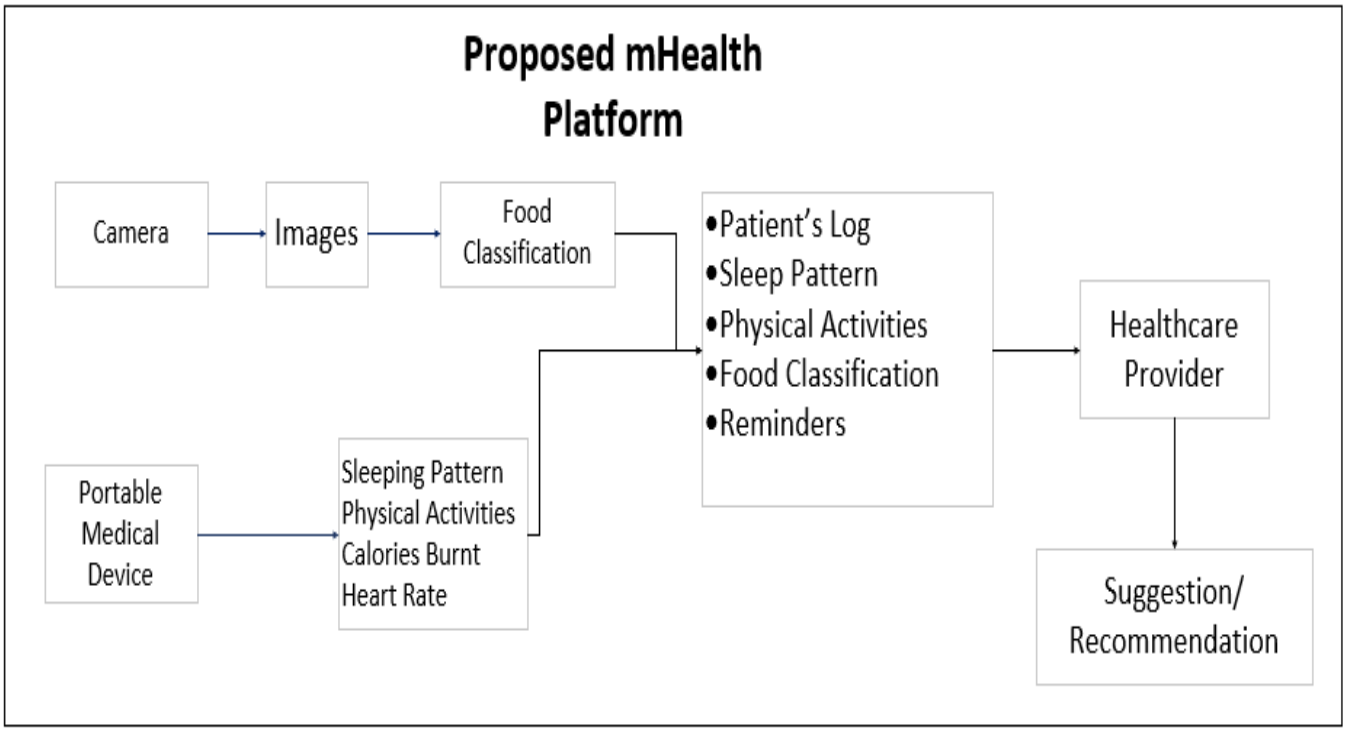


Fig. 1. Framework of the proposed healthcare system

A. *Patient Log*

A computerized health care system to record patients' information and basic statistics of the patient, such as name, daily intake of water, previous records of dialysis or transplantation, current stage of CKD if any, this information is stored in cloud storage for future purposes, which allows programmed monitoring. Patient logs are basically helpful to monitor patient's condition which is a significant issue.

B. *Sleep Pattern*

Sleep issues are a typical affliction for most individuals despite the fact that sleep is a vital factor amongst the most critical factors in keeping up both emotional health and physiological. A proper sleep is very important for general health, in other words, the kidneys. The function of kidney is usually regulated by the sleep pattern. People who sleep less have a faster decline in their kidney function and nighttime patterns affect CKD.

Patients with chronic kidney disease have problems staying asleep and falling asleep. If a patient has a habit of consuming too much caffeine, reducing the intake can help to have a better sleep, as caffeine is a stimulant it changes the sleep cycle. Therefore, to monitor the sleep patterns for CKD patients we will be using a wearable or smart fitness tracker to track the sleeping pattern it monitors the movements. When the device is synced then the subsequent morning, the software decodes those activities and movements into sleep statistics. And the technique of tracking movements by means of a device to measure the cycle of sleep is known as actigraphy.

Actigraphy nowadays have movements (accelerometers) as well as an adequate memory to store data for few weeks. Movements are examined a few times each second and analyzed later on. The remarkable component of actigraphy that separates it from early innovations is the capacity to connect the gadget to the wrist (or lower leg) of a person for a long time and give movement information continuously.

C. *Water Intake*

Keeping an account of water level is important for Chronic Kidney disease patients. Water intake varies with different stages of CKD. Hence, maintaining the water level consumption is a crucial factor for CKD. There is presently a proof on the particular impact of high-water consumption on the kidney. [13] Increased water consumption has been appeared to decrease the threat of kidney stones and lower the chances of forming stones. But at the point when the kidneys fail, individuals don't discharge enough water. For the individuals who are accepting dialysis treatment, water intake should really be extremely limited. Thus, to notify patients to stay hydrated, we are using Webhooks services. This will integrate the services on IFTTT (If this then that), where one can build Applets that will work with any app or device which can send and receive a web request. From the changes, that takes place inside other web facilities, for example Facebook, Gmail, Instagram etc. Applets incorporate sending messages to Messenger from the Google Assistant, getting estimates by means of Weather Underground, along these lines; the conceivable outcomes are relatively inestimable with a touch of creative ability. Here, in this study, the IFTTT bot on Messenger was utilized to notify the users to accomplish something as straightforward as a glass of water every 4 to 5 hours.

D. Food Classification

Pictures using smartphone camera have been used to measure the nutrients of the food. In particular, Patients with CKD cannot remove excess of salt so it is crucial to maintain the intake of salt. Excess of salt can cause swelling of hands, ankles, feet, breathing problems, increased blood pressure, etc. Therefore, balancing salt level is important in CKD [14].

To enlighten users about the nutrients their foods are having we will be using Convolutional Neural Networks that is referred as CNN, which is a kind of Artificial Neural Networks that is recognized to be vastly potent within the field of image classification and identification [15]. CNN performs the four following steps: -

(i) Convolution: - The most utilize of the Convolution operation in case of a CNN is to recognize suitable features from the picture, which is an input for the primary layer. Spatial interrelation of pixels is maintained by convolution. It is done by utilizing tiny squares of picture by achievement of picture features. Formula for convolution is given below:

$$I_{new(x,y)} = \sum_{j=-1}^{1} \sum_{i=-1}^{1} \alpha_{ij} I\_old(x-i, y-j) \quad (1)$$

(ii) Non-linearity: - Rectified Linear Unit (ReLU) that is a non-linear procedure. It acts on a basic level. ReLU is a process that is done per pixel, which then surpasses all the negative numbers of every pixel by zero in feature map. It is mainly a smooth estimate. (1+ex (x) = lnf).

(iii) Sub-sampling or Pooling: - Spatial Pooling that is additionally called down sampling or subsampling helps in decreasing the scale of every feature map however even whereas doing therefore, retains the foremost important info of the feature map. Once pooling is completed, ultimately the three-dimensional feature map will be regenerated to at least 1D feature vector.

(iv) Fully Connected layer (*Classification):* - The outcome of pooling operation and convolution give noticeable feature that are removed from the picture. The features are then used by FC layer for lowering the input picture into distinctive classes centered on training dataset.

E. Physical Activites

Physical inactivity causes a decrease in daily activities of day to day life and nature of life. [16] By improving the quality of their bones, joints, and muscles through working out, individuals with CKD can enhance their coordination and balance. This can additionally avoid the patients from falling also, secure their autonomy as they age. Also, it is broadly realized that CKD patients are at high hazard for unexpected death because of heart disease due to some degree to their inactive behavior. [17]. By expanding the physical activity levels of people with CKD, it is conceivable to reduce their danger of heart disease and enhance their physical working, thus decreasing the risk of unexpected death [18].

In this case, a portable medical device or sensor such as accelerometer will be used to constantly monitor the activities of the patient's body taking place in an accelerometer of 3-axis. The statistics are logged every nanosecond starting from the time the wearable is worn, which allows it to trace if an individual is running fast, walking forward even standing without any motion.

All these statistics are stored for further processing. This processing happens when the information is exchanged to the product related with the wellness tracker on the cell phone or PC for matching up. Since the individual has officially imparted individual points of interest to the product, the information gathered is gone through a customized calculation. This makes it feasible for the product to recognize what the distinctive developments recorded really suggest. It orders the developments into various exercises and afterward creates more data dependent on these points of interest. These are generally put away under various headings in the wellness tracker application. At the point when the individual taps the application a wide assortment of data is accessible from the handled information, prepared for perusing. The application lets the user know the count of steps that is taken, the rate of speed, the individual's pace, and also the number of calories liable to be scorched. The application enables the person to communicate with the data in an easy way.

IV. Results And Experimental Design

A. Data Collection

Data received from the patient's log, sleep pattern, and physical activities are collected. After collecting the user records, the application makes a calculation on each user and takes necessary actions. Incorporating the app of end client and web-based interface together with cloud help in this process.

B. Experimental Design

An android-based application was developed using Java and web portal of the mHealth was built using PHP. Notifications and reminders were triggered combining IFTTT (If This, Then That) and Facebook Messenger. The entire system was hosted using cloud storage, as the data are stored in cloud, henceforth no further trouble of uploading and downloading data were needed. Additionally, when the healthcare providers, which is a doctor or any medical specialist receive the patient's data, she/he further evaluates the information and addition steps are taken and suggestion or recommendation are given based on the severity of the patient's condition.



Fig. 2. The sleep pattern of a patient with the help of wearable in this case Fitbit.

In Fig. 2, the level of sleep and the wake duration are determined by utilizing computer programs, for example, time of total sleep, percent of time spent sleeping, time to wake, and percentage of total time staying awake. It monitors the activity of brain rather than counting number of turns.

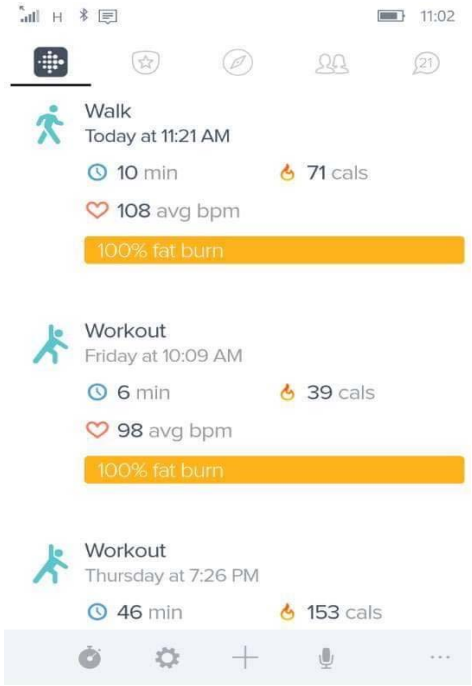


Fig. 3. Walking and workout sessions of a patient on the following days

Daily Physical activities of a patient is recorded using Fitbit, where the wearable sensor such as optical sensor calculate or record the frequency, acceleration, intensity, outlines and duration of a patient's movement. It takes a good bunch of statistic which helps a track whether one is just walking or just using hands to wave someone goodbye. In Fig.3, it is shown that the patient was walking for good 10 minutes and was working out as in was doing some light exercises for 6 minutes and 46 minutes on the following days respectively.

Due to the complexity of the system, efficient tools and technologies was used to simplify the entire infrastructure. An app has been build using Java Alerts and notifications have been implemented with the help of If This, Then That (IFTTT) and Facebook Messenger combining together. A shown in the Fig. 4, the reminder to stay hydrated was geenrated after every 4 hours interval.

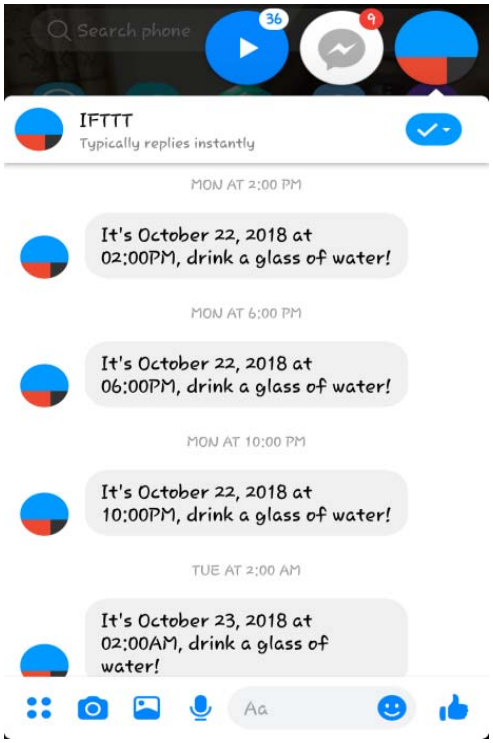


Fig. 4. Daily Reminder to drink a glass of water

C. *Results*

In our model, we used IoT and deep learning technologies. To analyze the nutrients in the food we used Image Recognition. Where the dataset was built with different type of food from existing online repositories. Furthermore, using Convolution Neural Network, the outcomes obtained were as following.

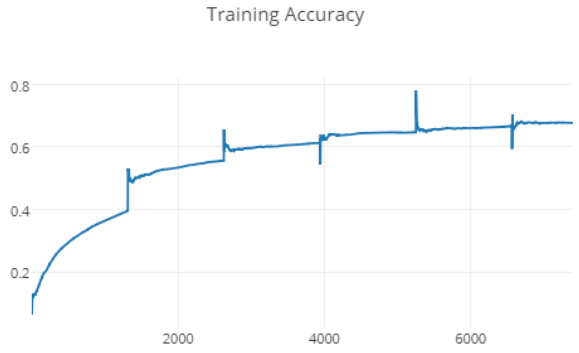


Fig. 5. The training accuracy (67.6%)

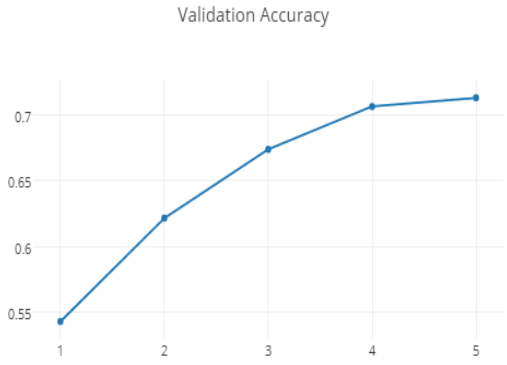


Fig. 6. The validation accuracy (72.8%)

Here, we are using **Deep Learning Studio**, which is also available in cloud solution. It is an individual user solution for deploying and creating Artificial Intelligence (AI). The basic interface causes any user to plan deep learning prototype effortlessly. Pre-learned models and in addition built-in assistive highlights streamline and speed up the model improvement process. We used our food dataset and modified the model with the visual interface. The stage consequently spares each model form as repeat and tune hyper-parameters to enhance execution. The model achieved a training accuracy of 67.6% and validation accuracy of 72.8%. Comparing to other existing system previously which also used CNN, our model could provide better result.

V. Conclusion & Future Work

Internet of things (IoT) has been separating different conventional figuring techniques in all field particularly healthcare. In IoT healthcare industry the idea of remote checking, the condition of patients' is a different break-through. The proposed model is developed to be of low-cost and more reliable particularly for CKD patients with real-time monitoring system, which is online based. The system is designed to monitor a patients' vital information. Various device and sensors are used here for the monitoring purpose. The system consists of a patient log, sleep pattern, level of water and salt intake, nutrition information mainly based on salt. All the data collected from the patients' log are stored on the cloud, which can be later used for, further changes, and generate an overall report for further analysis. The system is therefore not only helpful for patients but also for nephrologists and renal dieticians. Finally, the proposed model was successful to provide visions regarding renal conditions.

For future enhancements, a knowledge-based agent will be designed using intuitionistic fuzzy theories, which will be playing the role of healthcare provider, as a doctor will not be able to respond 24*7. More descriptive work for each module will be covered, such as custom-made mobile portal, wireless body sensor to monitor a patient's blood pressure will be introduced as hypertension can be the progressive cause of chronic renal disease.

References

[1] Facts avaliable :- https://iot.do/home-test-kit-smartphone-kidney-disease-2018-04

[2] Facts:- https://www.forbes.com/sites/andrewweinreich/2017/12/18/the-future-of-the-smart-home-smart-homes-iot-a-century-in-the-making/#39ee31f657ac

[3] S. Soegijoko, "Recent Developments of Telemedicine and E-Health and Its Application Prospects in Indonesia" [Perkembangan Terkini Telemedika dan E-Health Serta Prospek Aplikasinya di Indonesia], *Seminar Nasional Aplikasi Teknologi Informasi 2010 (SNATI 2010), 2010.*

[4] Y. Gunardi, A. Adriansyah, and T. Anindhito. "Small Smart Community: An Application of Internet of Things", *ARPN Journal of Engineering and Applied Sciences*, *Vol. 10, No. 15, August 2015*

[5] B. S. Rani and N. Venkatram, "Implementation of PHM System using IoT and Raspberry Pi", *International Journal of Control Theory and Applications, Vol. 10, No. 35, 2017.*

[6] S. Sagar, D. Niranjan, V. Pandurang, and M. Dighe, "IoT Based HealthCare System Using Raspberry Pi", *International Research Journal of Engineering and Technology (IRJET)*, Vol. 04, Issue 04, *Apr 2017.*

[7] D. Aranki, G. Kurillo, P. Yan, DM. Liebovitz, R. Bajcsy "Real-Time Tele-Monitoring of Patients with Chronic Heart-Failure Using a Smartphone: Lessons Learned", *IEEE Transactions on Affective Computing 7 (3), 206-219.*

[8] XJ. Zhang, YF. Lu, and SH. Zhang, "Multi-task Learning for Food Identification and Analysis with Deep Convolutional Neural Networks," *Journal of Computer Science and Technology 31(3): 489–500, DOI 10.1007/s11390-016-1642-6, May 2016*

[9] M. A. Al-Taee, W. Al-Nuaimy, A. al-Ataby, Z. J. Muhsin, S. N.Abood, "Mobile health platform for diabetes management based on the Internet-of-Things", *IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT),IEEE Conference Publications, pp. 1-5,2015.*

[10] S. Pinto, J. Cabral, T. Gomes, "We-Care: An IoT-based Health Care System for Elderly People", *IEEE International Conference on Industrial Technology (ICIT), 2017.*

[11] K. Kaiya, A. Koyama, "Design and Implementation of Meal Information Collection System Using IoT Wireless Tags", *10th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS), 2016.*

[12] X. Long, P. Fonseca, R. Haakma, R. M. Aarts, "Actigraphy-based sleep/wake detection for insomniacs" *IEEE Journal of Biomedical and Health Informatics, Vol. 18, No. 14, July 2014*

[13] Clark WF, Sontrop JM, Huang S-H, Moist L, Bouby N, Bankir L. Hydration and chronic kidney disease progression: a critical review of the evidence. Am J Nephrol, 2016.

[14] Facts Available :- https://kidney.org.au/cms_uploads/docs/salt-and-your-kidneys-fact-sheet--kidney-health-australia.pdf

[15] J. Rathod, V. Wazhmode, A. Sodha, & P. Bhavathankar, "Diagnosis of skin diseases using Convolutional Neural Networks." *Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2018.*

[16] Hiraki K, Yasuda T, Hotta C, et al. "Decreased physical functioning pre-dialysis patients with chronic kidney disease. Clin Exp Nephrol.", 2013.

[17] Greenwood SA, Koufaki P, Mercer TH, et al. "Effect of exercise training on estimated GFR, vascular health, and cardiorespiratory fitness in patients with CKD," a pilot randomized controlled trial. Am J Kidney Dis. 2015.

[18] Rita Huie Milam, Journal on "Exercise Guidelines for Chronic Kidney Disease Patients"

Design and Implementation of IoT based Ideal Fish Farm in the Context of Bangladesh Aquaculture System

Abdul Kadar Muhammad Masum
Dept. of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh
akmmasum@yahoo.com

Md. Shahin Mia
Dept. of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh
Shahin.iiuc@gmail.com,

Md. Kalim Amzad Chy
Dept. of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh
kalim.amzad.chy@gmail.com

Shahidul Islam Khan
Dept. of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh
a.k.azad.um@gmail.com

Ahmed Shan-A-Alahi
Dept. of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh
ahmed.iiuc@gmail.com

Md. Golam Rabiul Alam
Dept. of Computer Science and Engineering, BRAC University, Bangladesh
gra9710@gmail.com

***Abstract*—With the first growth of population, the protein demand in the world is also increased that can't be overcome by meat. Fish is a great source of protein even low-cost comparing with meat. Traditional fish farming and most of the existing smart system fail to produce the expected level of fish. This research paper demonstrates an IoT based Ideal fishery that will be low in cost and high in fish production level. A microcontroller is engaged in measuring several essential features of the water like pH, temperature, water level, oil layer and the fish behavior to understand their hungry level that directly affects in the first growth of the fish. The user will be notified of the information through our developed mobile app and an interactive web interface. Bluetooth and ESP8266 wifi module are integrated with the system to deliver the data to the mobile app and web interface. In case of any abnormality detection, the system informs the user to take the immediate steps. Besides, forecasting of the several water parameters in the following day to take an earlier preventive action makes the proposed framework more exceptional. In consequence of this feature, the system will assist the fish farmer to produce more fish, and the protein challenge for the large population will be diminished drastically.**

Keywords— fish farm, pH sensor, oil layer detection, Support Vector Regression, Random Forest Regression

I. INTRODUCTION

The present 7.3 billion world population is estimated to be 11.2 billion in 2100. With the first growth rate of the current population, at least 70% more protein will be necessary in year 2100[1]. Oceans, a great source of protein, is taken about 2.5% more fish than what ocean can naturally support[2]. WWF (World Wildlife Fund) reported that since 1950, 90% of very usually captured fish have vanished from the ocean[3]. We are increasing stress on the ocean that's a threat to the natural balance. To meet the demand, an alternate source should be though without threatening nature. According to a report by FAO, aquaculture has experienced a significant annual growth rate of 7.4% with escalating meat requirements[4].

Aquaculture-based fish farming is an excellent solution to this crucial problem. Fish farming is very cheap as only 1.5 pound feed is needed to produce 1 pound fish while in case of meat from cow its 8-9 pound feed and 8000L water[5].In respect to financial gain, a significant difference between IoT based fish farm and non-IoT based fish farm has been found on a study of Zhang et al. [6].The performance of the fish relies on various factors: production factors [7],environmental factors [8],and biotic factors [9].It is possible to monitor and control some of those parameters.

The primary objective of this work is to develop a low-cost smart fish monitoring and controlling system that can measure different important water characteristics like pH, temperature, oil level, water level, and fish hungry level by observing the fish behavior. Several low-cost sensors are equipped with the system to accomplish the assigned task. To achieve an accurate value several data manipulation techniques like mean, mode, data normalization has been followed. The gathered information is stored in the cloud server for further analysis like forecasting the probable pH, oil layer, water level, required fish feed etc. to take advance preventive steps. If the parameters exceed the predefined threshold level, then the system will notify the user through the mobile app. Also, the Bluetooth module is used to update the data in the mobile app for a local user. To develop the system only $50 - $55 is required. This amount is so low cost that a fish farmer can easily adapt it and take part in resolving the world protein challenge problem.

This paper is structured into five sections containing the introduction. Section II presents related work. Section III combines methodology where circuit design, smart fish monitoring, and controlling process is explained. An experimental result is shown in section IV. Section V The conclusion and future work are deliberated in section V.

II. RELATED WORK

A few researchers from the genetic engineering field produced genetically modified fish having more breeding capabilities. But their survival rate and nutrition level are lower than the natural fish[10]. Also supplying artificial fish feed containing more chemical, microbes, seaweeds, etc. causes various disease. Alternate fish feeds are developed having lower chemicals to control the diseases.

Nowadays, Internet of Things (IoT) makes a revolutionary change in diverse fields. In consequence of this revolution, IoT improves the fish farming system that resolves a large portion of the nutrition problem for more growing population[11]. Conventional fish farming faces a lot of difficulties in the change of the climate like heavy rainfall, extreme weather,

limited water in the farm, high temperature, etc. FAO highly recommends that modern technologies should be integrated with traditional farming to overcome the challenges. SK telecom developed an IoT based fish farm system that observes pH, temperature, dissolved oxygen and notify the fish farmer in real time through their smartphones[12].To test the water quality, the conventional method takes several samples from various locations and process in the laboratory. To smart this method, Shareef et al. [13]presented an IoT based decisions support system using WSN (Wireless Sensor Network) technologies that determine various physic-chemical characteristics of water and inform the end user through smart mobile. Daudi et al. also developed a fish farm monitoring system to examine water quality based on WSN technology[14]. Their equipped sensors collect temperature, oxygen level, pH, water level of the fish farm and transfer the data to the mother server via ZigBee protocol. Raspberry Pi based system continually determine the water quality through equipped sensors and store the data in a cloud server for further analysis[15]. If any water parameter exceeds its threshold level than an auto-generated message will be sent to the user. They utilized Lab View software to show the gathered information. But the high cost is the main drawback of this system.

A three-layer architecture-based system is proposed where a remote layer collects the data, server layer transfer and store the data. The obtained information is displayed in the client layer[16]. The software is also integrated with the system to monitor the sensor data. The proposed IoT based system of Cesar Encinas et al. deployed Arduino, Atlas pH prove digital sensor, dissolved oxygen probe, analog temperature sensor, and a ZigBee module to transfer the data into MySql database[10]. They also used 200 mA/h rechargeable batteries lasting for 8h. The system for fish farm monitoring system determines similar factors instead of considering other essential considerations like the density of the oil layer. Besides, most of the work doesn't identify which sensors they have employed or equipped high-cost sensors that make the system difficult to cope up in the fish farm. Moreover, most of the system supply the fish feed after a certain period ignoring feeding when they required. Prediction of the amount of next week fish feed is missing in most of the system.

In the context of Bangladesh aquaculture system, lots of fisheries follow the traditional methods to cultivate the fish. As a result, the cost becomes high, but the output is not at an expected level. To the best of our insight, this is the first attempt to develop a low-cost fish farm monitoring and understanding when fish require feed by analyzing their movement. Also, the forecasting features of the next week fish farm parameters willhelp the fish farmer to take a preventive step in advance.

III. Methodology

Figure 1 demonstrates the whole picture of the proposed IoT based fisheries monitoring and controlling system, how the fisheries information will be gathered and how transferred to the end user of the system. The system estimates several parameters of the water like pH, temperature, oil layer, water level and the movement of the fish that directly effect on the quick growth of a fish. After supplying the power, the system is activated. At the first time launching, to measure the targeted water parameters and fish movement, the system takes a few moments to initialize the equipped sensor. As pH sensor doesn't give a proper reading at a static situation, it needs to shake slightly to measure the pH of the water appropriately. The system uses a servo motor to perform this task that rotates the pH sensor 30degree in one direction and 30degree in reverse direction. For more accuracy and lower the amount of power consumption of the system, it takes an average of five ph value where each value is read in every 3min. By using equation (1) this average value is calculated where P is the value of ph in every 3min and pH is the desired result.

$$pH = (\sum_{i=1}^{5} P_i)/5 \quad (1)$$

Dense oil layer hinders the fishes to meet the sunlight that reduces the grown of the fish even can cause death. By IR optical sensor, the proposed system measures the density of the oil layer and stores the data in memory. Like pH value measuring, the value of the oil layer is calculated using equation (2) where D is the value of the oil layer in every 3min and Oil layer is the ultimate result.

$$Oil\_layer = (\sum_{i=1}^{5} D_i)/5 \quad (2)$$

Usually in a fishery, after a certain period, fish food is supplied to the fishes. But maybe continuously increasing fishes need food earlier than the given time or more food. We know, starve fishes grow very slowly. So, feeding the fish when requires is a mandatory feature of a model fishery. The system can determine either the fishes is Extreme hungry or medium hungry or normal hungry so that enough fish food is supplied. The abnormal movement, i.e. not a frequent movement of the fish can help us to understand the fish hunger level. Three ultrasonic sensors (*a, b, c)* are deployed in this purpose where each ultrasonic sensor takes six value after every 10 sec. For each sensor average value is calculated using equation 3,4 and 5where *a* is a location of a fish measured by sensor *a* and *b* and *c* are like *a*. Then for each sensor RMS value is calculated using equation 6, 7 and 8 $a_rms$is measured RMS value by sensor *a*,$b_rms$and c_rms are like a_rms. From these three RMS value of the sensor, we can determine the hungry level easily. If three RMS value exceeds the threshold value, then the hungry level is extreme, if two value exceed then medium hungry level, if one value exceeds then normal hungry level, if any RMS value doesn't exceed the threshold value then satisfied level.

$a = [a_1, a_2, \dots a_6]$ $b = [b_1, b_2, \dots b_6]$ $c = [c_1, c_2, \dots c_6]$

$$a\_avg = (\sum_{i=1}^{6} a_i)/6 \quad (3)$$

$$b\_avg = (\sum_{i=1}^{6} b_i)/6 \quad (4)$$

$$c\_avg = (\sum_{i=1}^{6} c_i)/6 \quad (5)$$

$$a\_rms = \sqrt{(a_1 - a\_avg)^2 + \dots + (a_6 - a\_avg)^2} \quad (6)$$

$$b\_rms = \sqrt{(b_1 - b\_avg)^2 + \dots + (b_6 - b\_avg)^2} \quad (7)$$

$$c\_rms = \sqrt{(c_1 - c\_avg)^2 + \dots + (c_6 - c\_avg)^2} \quad (8)$$

Too hot and too cold temperature is a great reason for the slow growth of a fish. The system gets the current temperature of fisheries through the temperature sensor. In similar to pH and oil layer value measuring, the system measures the temperature of the fishery using equation (3) where T is the temperature in every 3min and Temperature is the estimated value.

$$Temperature = (\sum_{i=1}^{5} T_i)/5 \quad (9)$$

Similarly, too high and low water level is not suitable for the fishes. A liquid level sensor can measure the water level, but it is costly. To make the system more cost-effective, instead of the liquid level sensor a low-cost ultrasonic sensor is equipped to measure the water level. The system recognizes either the fisheries are in high or low or normal water level. Like previous value measuring water level is determined using equation (4) where W is the water level in every 3min and water level is the outcome of the right-hand side calculations

$$Water\_level = (\sum_{i=1}^{5} W_i)/5 \quad (10)$$

Through two communication media, the microcontroller sends the gathered data to the end user. A mobile application and web interface are integrated with the system to operate it easily. Bluetooth module sends the data in the mobile app so that local fishery supervisor can easily gain the necessary information. ESP8266 WIFI module sends the data to the web interface so that one can monitor the fishery remotely. In the web interface, the data is plotted graphically and stored in the database that will help the supervisor of the fishery to check the previous data when necessary. Also, the ESP8266 WIFI module can interact with the mobile app in case of Bluetooth module failure. Also, if the parameters of the water quality and movement of the fish exceed the preset threshold value then buzzer will be on and fishery supervisor can be notified through mobile app also.

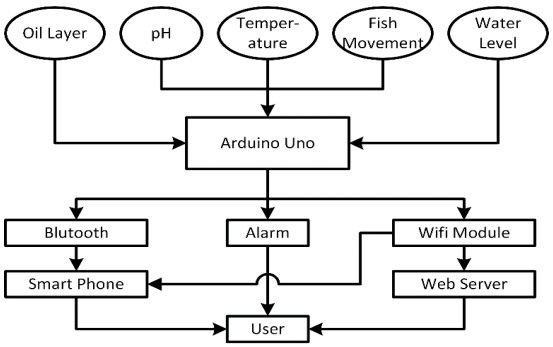


Fig. 1. Block Diagram

The proposed system is classified into the following parts:

A. Circuit Design

The proposed system is developed by following the designed circuit, shown in Figure 2.The ground and VCC pin of every sensor are connected to the ground and 5v pin of Arduino Uno. pH sensor relates to the analog pin of Arduino. From the analog output of the pH sensor, the value of pH is calculated using the formula of pH calculation. Servo motor, buzzer, temperature sensor, IR optical sensor, ultrasonic sensor, Bluetooth module, WiFi module are connected to the Arduino Uno microcontroller board according to the circuit diagram.

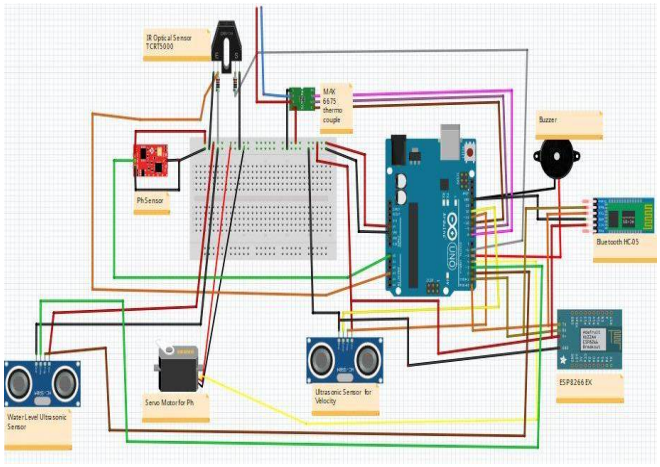


Fig. 2. The circuit design of the proposed system

The proposed model is designed with the following components:

B. Hardware: Following pieces of equipment are utilized to develop the hardware part.

1) Arduino Uno: ATmega328 IC based Arduino Uno has six analog inputs, 14 digital i/o pins, USB connection, 16MHz oscillator, ICSP header, reset button, power jack and 6 PWM output pins[17]. The physical computing platform of the Arduino Uno board is an open source.

*2) pH Sensor:*pH sensor measures the alkalinity or acidity of a solution that's scale range is from 0 to 14 where 7 is considered as neutral, below 7 acidic, above 7 alkalic.TOL-00163 model of pH sensor is equipped to measure the pH level of the fisheries water[18].

3) HC-SR04 Ultrasonic Sensors: Ultrasonic wave based, HC-SR04 ultrasonic sensor's head emits an ultrasonic wave and receives the reflected wave from the target. Measuring the taken time between the emission and reception it measures the distance to the target. In this system, HCSR04 ultrasonic sensors are used to measure water level and fish movement for detecting hungry level. At 40 KHz frequency, it can identify an object at range from 2cm to 400 cm[19].

4) IR Optical Sensor: The TCRT5000 IR optical sensor is a reflective sensor which incorporates an infrared emitter and phototransistor in a leaded package that prevents visible light[20].

5) Bluetooth Module: Among various types of Bluetooth module HC-05 Bluetooth module is selected for the system that works on 3.6V to 6V. 9600 is its default baud rate [21].

6) Wifi Module: TCP/IP protocol based ESP8266 Wi-Fi Module has a powerful enough processing capability that allows it to be integrated with the sensors with

minimal loading during runtime. In this system, it is used to transfer real-time data at the server side[22].

C. Softwares:

Arduino IDE and language is used to write the code and burnt into the system that is easy to understand and reliable.

D. Algorithm:

The algorithm is described by a flow chart. The whole flowchart of an IoT based ideal fishery system appears in figure 3.At the first level, all sensor values are initialized. To measure the pH value, servo motor rotates 30 degrees left, and 30 degrees right as pH sensor needs to shake for proper value. By using pH formula, pH is determined 5times in every 3min. Average of the pH value is calculated for a 15min duration. As optimal pH value for speedy growth of fishes is 7 – 8.5, the system turns on the alarm if it exceeds this range[23]. IR optical sensor engages to measure the oil layer of the water. The output of the IR optical sensor gives a wholevalue, heat and noise value. By removing the heat and noise value form the entire value, the system gets the actual value.

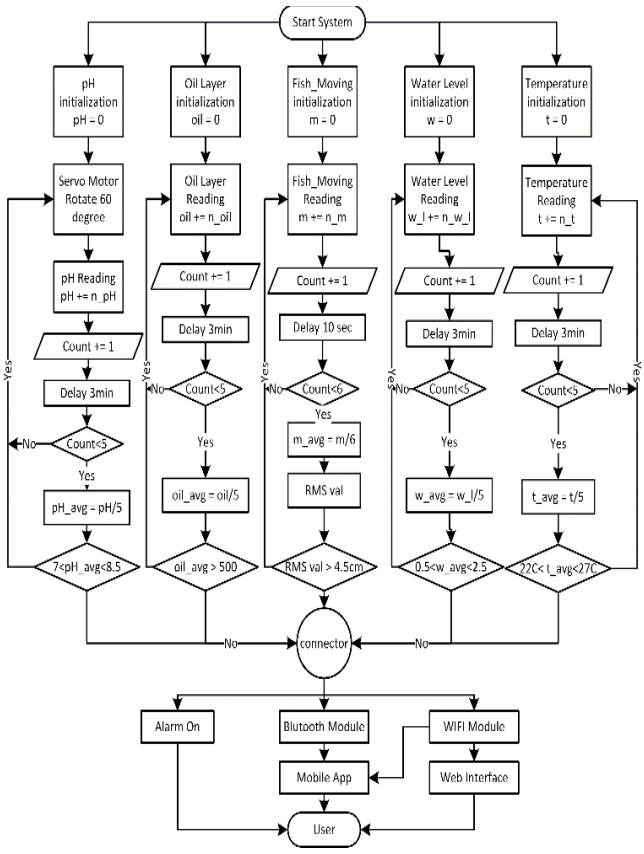


Fig. 3. Working process of the proposed system

Like the pH sensor, the average value of this sensor is measured in the same fashion. If the oil layer value less than 500 than it notifies the user and turns on the alarm. To estimate the fish moving rate, every 10 sec is taken in 1min duration. 4.5cm Root Mean Square (RMS) value of the sensor act as a threshold value to determine the hungry level of the fishes. Average value of the water level and temperature within 15min is calculated like the pH sensor.[24]says that 22C to 27C is the optimum temperature for the rapid growth of the fish. So, in case of exceeding this range, the system informs the owner of the fishery through mobile app and web interface. The flow of the data from the device to the mobile app and web is performed through the Bluetooth and WIFI module.

E. Predicting the following day power consumption:

The system can forecast the various essential parameters and notify the user. For simplicity here we have considered just forecasting of the probably required fish feed in next day. Like this process, the system can predict other parameters. To select the best regression model for our forecasting feature, different kinds of machine learning regression algorithm have experimented. From a local fishery, their daily fish feed supplied amount is recorded for the purpose of making a data set. We have recorded about 200-day data of that fisheries and used the data to train our model. Dataset is split into 70% and 30% where 70% data is for training and 30% data is reserved for testing the model. Through different kinds of machine learning regression algorithms such as linear regression, decision tree regression, polynomial regression, support vector regression (SVR) and random forest regression, we fit the training data to prepare the model to forecast the amount of the fish feed in the following day. After finishing the model training part with training data, the model is examined by testing data. Then Mean absolute error and accuracy level is calculated to evaluate the model. From these results, we have found Random Forest Regression gives better outcomes than other regression models.

IV. Experimental Results and Discussion

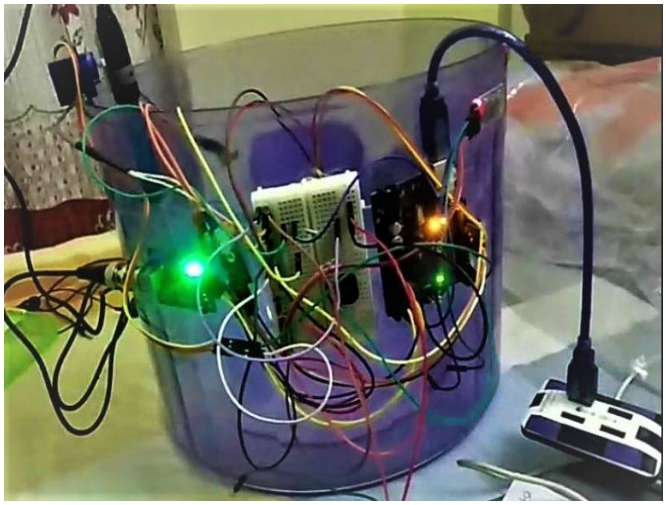


Fig. 4. Photograph of the prototype

A prototype of the proposed system appears in figure 4. pH, temperature, oil layer, water level, the hungry level of the fish

are the outcomes of the system. The essential parameters of the water are sent to the mobile app through Bluetooth module for local user and data sent to the web server through the ESP8266 WIFI module for the remote user. Also, data is saved to the cloud storage for further analytics. In case of any parameters exceeding the predefined threshold value, the system notifies the local user as well as the remote user.

In figure 5 we have seen some sample outcomes of the system taken at different time and different situation. In figure 5(a) we have seen all parameter is ok while at the time of medium hungry level the system notifies it successfully (figure 5(b)) and inform the user. In case of pH level decreasing the system notify the user, shown in figure 5(c) while in the oil layer exceeding the threshold value case the system successfully determine the situation.

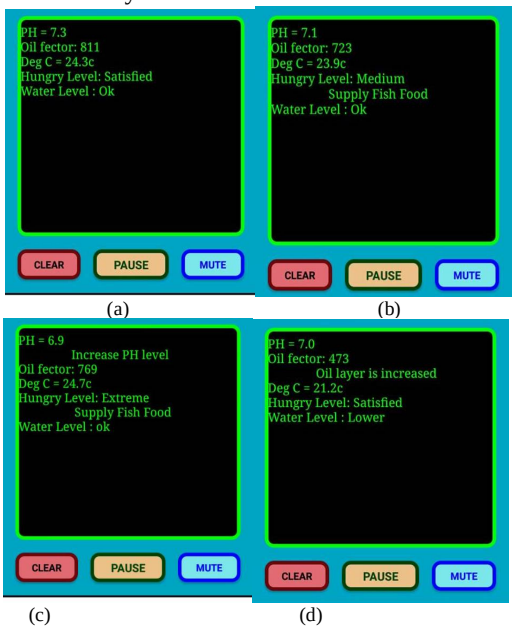


Fig. 5. (a) A sample of everything ok, (b) A sample of fish hungry notification (c)A sample of Ph and hungry level notification, (d) A sample of oil layer notification

In the web interface, the system sent the measured data and visualized it graphically to make it more interactive with the user. In figure 6, the system visualizes the pH data over time and store it into the server for further analytics. Like the pH data visualization, the system also visualizes the temperature data (figure 7), oil layer data (figure 8) and water level data (figure 9) graphically into the web interface for the remote user.

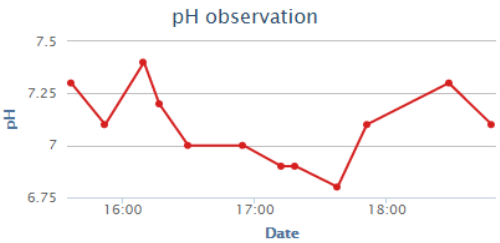


Fig. 6. pH observation

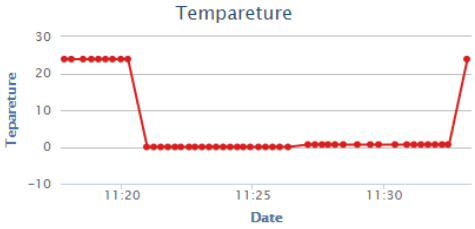


Fig. 7. Temperature observation

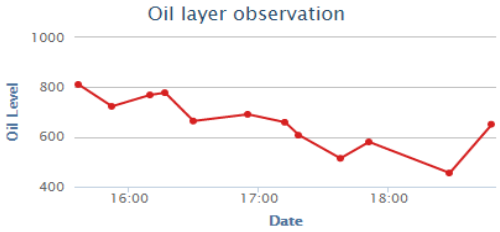


Fig. 8. Oil Layer observation

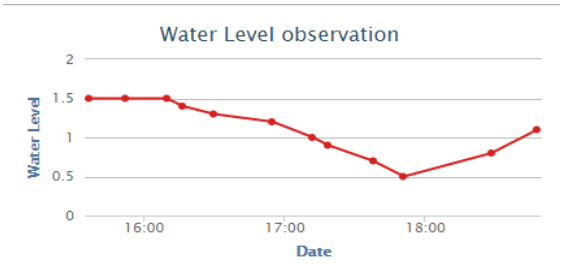


Fig. 9. Water Level observation

TABLE I. TESTING DATA

Date Test	Fish Feed (g)	Linear Regression	Polynomial Regression	Support Vector Regression (SVR)	Decision Tree Regression	Random Forest Regression
113	10096.59	10101.15	10101.31	10294.05	10602.16	10524.65
218	10173.3	10176.01	10178.49	10294.79	10693.29	10418.38
303	10235.84	10236.61	10237.68	10294.44	10732.53	10593.94
515	10394.56	10387.76	10383.78	10295.23	10561.82	10582.49
717	10572.98	10531.77	10539.52	10295.58	10732.19	10689.75

To forecast the following day power consumption, our trained model is examined by testing data. Some forecasted data are displayed in table 1 where the 1st column “Date Test” refers to the forecast of the power consumption of that date. “MonthDate" format is used to represent a date where the last two digits refer to data and remainder value to month. For example, "218" implies 2nd-month 18th day, “717” implies 7th-month 17th day and so forth.

Accuracy level and mean absolute error is considered to judge the model that is shown in table 2 and table 3

respectively. Because of varying the parameters, the outcomes diverge from one regression analysis to another regression analysis. For example, shifting the degree of polynomial regression analysis provides a good result than linear regression. Random forest regression gives a better result than decision tree regression and other models. So, it can be decided that the random forest regression-based model is more suitable to forecast the water parameters.

TABLE II. ACCURACY COMPARISON

Accuracy	Linear Regression	Polynomial Regression	Support Vector Regression (SVR)	Decision Tree Regression	Random Forest Regression
Training Set	85.45	85.49	87.49	92.09	95.75
Testing Set	84.76	84.85	86.05	90.94	94.24

TABLE III. MODEL EVALUATION

Evolution of the Linear Model		Linear Regression	Polynomial Regression	Support Vector Regression (SVR)	Decision Tree Regression	Random Forest Regression
Mean Absolute Error	Training Set	44.2489	44.2063	29.2189	26.3687	23.0747
	Testing Set	43.6042	43.2575	35.5301	30.2692	23.8518

V. CONCLUSION

In demand for the increased level of protein for the fast-growing population of the world, the fish production needs to be raised. Traditional fish cultivation method is failed to meet the demand because of the adverse environment and ignoring the current behavior of the fish. This study proposes an IoT based smart fish monitoring and controlling system where various essential water parameters and fish behavior is continuously measured and notify the local user as well as the remote user in case of any abnormal situation arisen like increased pH or dense oil layer etc. To make the system more user-friendly, a web interface and mobile app are created so that the user can examine all earlier data from everywhere in the world. Alongside these characteristics, the system can forecast the water parameters like pH, oil level, water level, temperature, the fish feed of the following day by machine learning regression algorithm that will assist them to take an advanced step in the fish farm. The proposed framework is more cost-efficient and smarter than existing tradition system as well as other smart systems. Reducing the cost of the system and integrating other important feature like determining the poison in the water is a future scope of the present system. The fish farmer should integrate our efficient and effective system to increase their fish production and defeat the challenge of the protein problem.

VI. REFERENCES

[1] I. Ullah and D. Kim, "An optimization scheme for water pump Control in smart fish farm with efficient energy consumption," *Processes*, vol. 6, p. 65, 2018.

[2] F. a. A. O. o. t. U. N. Rome. (2014, 28 January). *The State of World Fisheries and Aquaculture*. Available: http://www.fao.org/3/a-i3720e.pdf

[3] O. Heffernan, "No more fish in the sea," *Nature*, vol. 460, p. 956, 08/19/online 2009.

[4] FAO, *FAO Yearbook: Fishery and Aquaculture Statistics*: Fao, 2008.

[5] W. O. Review(WOR). (2013, 28 January). *WORLD OCEAN REVIEW 2, THE FUTURE OF FISH – THE FISHERIES OF THE FUTURE*. Available: http://hdl.handle.net/20.500.11822/8484

[6] Y. Zhang, J. Hua, and Y. B. Wang, "Application effect of aquaculture IOT system," in *Applied Mechanics and Materials*, 2013, pp. 1395-1401.

[7] M. A. Adewolu, C. A. Adeniji, and A. B. Adejobi, "Feed utilization, growth and survival of Clarias gariepinus (Burchell 1822) fingerlings cultured under different photoperiods," *Aquaculture*, vol. 283, pp. 64-67, 2008.

[8] V. Rubio, F. Sánchez-Vázquez, and J. Madrid, "Effects of salinity on food intake and macronutrient selection in European sea bass," *Physiology & behavior*, vol. 85, pp. 333-339, 2005.

[9] P. Kestemont, S. Jourdan, M. Houbart, C. Mélard, M. Paspatis, P. Fontaine, *et al.*, "Size heterogeneity, cannibalism and competition in cultured predatory fish larvae: biotic and abiotic influences," *Aquaculture*, vol. 227, pp. 333-356, 2003.

[10] J. A. Beardmore and J. Porte, "Genetically modified organisms in aquaculture," 2003.

[11] T. S. F. A. E. T. I. Vision. (2014, 28 January). *Towards Smart Farming—Agriculture Embracing The IoT Vision*. Available: http://www.beechamresearch.com/files/BRL%20Smart%20Farming%20Executive%20Summary.pdf

[12] (2014). *SK Telecom Starts Pilot Operation of IoT-based Fish Farm Management System*. Available: http://en.c114.com.cn/576/a855896.html

[13] Z. Shareef, S. Reddy, and D. IGDTUW, "Design and development of IOT Based Framework for Aquaculture," in *National Conference on Product Design (NCPD 2016)*, 2016.

[14] D. S. Simbeye and S. F. Yang, "Water quality monitoring and control for aquaculture based on wireless sensor networks," *Journal of networks*, vol. 9, p. 840, 2014.

[15] K. R. S. R. Raju and G. H. K. Varma, "Knowledge based real time monitoring system for aquaculture using IoT," in *Advance Computing Conference (IACC), 2017 IEEE 7th International*, 2017, pp. 318-321.

[16] M. Hua, D.-a. Zhao, W. Xia, Y. Zhu, and X. Liu, "The design of intelligent monitor and control system of aquaculture based on wireless sensor networks," in *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*, 2010, pp. 9-12.

[17] Arduino. (2019). *ARDUINO UNO REV3*. Available: https://store.arduino.cc/usa/arduino-uno-rev3

[18] T. Bangladesh. (2019). *PH Sensor with Module*. Available: https://www.techshopbd.com/product-categories/meters/2576/ph-sensor-with-module-techshop-bangladesh

[19] ElectronicWings. (2018, 23 September). *Ultrasonic Module HC-SR04*. Available: http://www.electronicwings.com/sensorsmodules/ultrasonic-module-hc-sr04

[20] T. Bangladesh. (2019). *Reflective Optical Sensor - TCRT5000*. Available: https://www.techshopbd.com/product-categories/sensors-98778/977/reflective-optical-sensor-tcrt5000-techshop-bangladesh

[21] Amazon. (2019). *HiLetgo HC-05 6 Pin Wireless Bluetooth RF Transceiver Module Serial BT Module for Arduino*. Available: https://www.amazon.com/HiLetgo-Wireless-Bluetooth-Transceiver-Arduino/dp/B071YJG8DR?ref_=fsclp_pl_dp_2

[22] sparkfun. (2019). *WiFi Module - ESP8266*. Available: https://www.sparkfun.com/products/13678

[23] Y. E. Corporation. (2016). *pH in Fish Farming*. Available: https://www.yokogawa.com/library/resources/application-notes/ph-in-fish-farming/

[24] T. F. Hobbyist. (2019). *Temperature Control*. Available: http://www.tfhmagazine.com/aquarium-basics/temperature-control.htm

Digital Forensic Investigation Framework for Internet of Things (IoT): A Comprehensive Approach

Md. Jahidul Islam*, Md. Mahin†, Ayesha Khatun‡, Biplab Chandra Debnath§, Sumaiya Kabir¶
Dept. of Computer Science and Engineering
Green University of Bangladesh
Dhaka, Bangladesh
{jahidul.jnucse*, mdmahin3†, ayeshankhatun‡, biplab.cse.pstu§}@gmail.com,sumaiya@cse.green.edu.bd¶

***Abstract*—**
Internet of Things (IoT) is receiving more and more attention in recent years. Even though a great number of researches have been going on focusing the field IoT, research on Digital Forensics Investigation (DFI) techniques are still little focused. Till now IoT platforms have not fully matured to adapt with the existing Digital Forensic(DF) tools, methods, and procedures. The prime reason is the nature of the cloud, network and IoT infrastructures (e.g. heterogeneity, distributed, jurisdiction, data duplication, multi-tenancy etc.). As a result, it is a very challenging task of locating, identifying, examining, analyzing, and presenting the potential IoT-based forensic evidence for digital forensic from the device, Network and, cloud environment. Consequently, there is no fully adapted international standard for an ideal framework to conduct DFIs. Because of this, we are proposing a comprehensive DFI process framework for the IoT environment that is able to enable more efficient and effective for DFI. The aim of this paper is to present a more comprehensible DFI framework for digital forensic professionals and experts. Also, it would be a proper guideline to guide the investigators and would also be useful in training of the fledgling investigators. Moreover, it reduces the dependence on the Cloud Service Provider (CSP) at the time investigation is going on.

***Keywords*—IoT, Digital Forensics, DFI, Forensics, Cloud Computing, CSP, ISP, Cloud forensics, Network Forensics etc.**

I. Introduction

IoT [1] is an infrastructure where smart nodes are connected to each other in a progressive and worldwide network infrastructure. Every day thousands of devices are being connected with the internet and sharing the local information. IoT can be applied in different areas like the smart city, smart home, transportation, farming, health care, etc. According to the information by Gartner, there will be an estimated 20.4 billion IoT devices be connected by 2020 (things). In accordance with the vast growth of IoT devices- that faces some new challenges in terms of security, it is becoming a new battlefield for cybercrime. Since billions of IoT devices are connected to each other and sharing their information from the device to device through various distributed network and cloud environment. Sensitive data communication from a complex system may encounter security threats as the data travel through various communication routes and devices. Various researchers have shown that threats, attacks, and different security issues on the IoT platforms that have been identified include: Virus & Worms, Spyware & Malware, Distributed Denial of Service (DDoS) etc [2]. These types of attacks may lead to power systems failure, failed user devices and, even sometimes infrastructure failures. As a result, it is always at risk of intruders, who may find out the weakness of crucial IoT infrastructure related to health, safety and communication systems and introduce malicious data to the systems. Investigation of these kinds of threats needs sophisticated DFI methods into the IoT domain. We call this new paradigm *IoT Forensics*.

In present years, numerous attacks on various social networks like Google, Yahoo, Facebook, Twitter etc are observed. Most of the social networks are affected by the Distributed Denial of Service (DDoS) attack. Moreover, capturing confidential information of users by phishing attacks, like doing illegal transactions by obtaining personal security information.

Furthermore, cloud computing is another part of networking that has limitless storage and processing capacity. With the sporadic growth of global cloud adoption in private, public and hybrid sectors, cloud computing environments face some new challenges every now and then in terms of security through increased crime-ware. In cloud computing, sensitive data communication from such complex virtual systems may often become vulnerable. Sometimes attackers find out the vulnerability or weak-point to trigger malicious data. Investigation of these kinds of attacks requires DFI methods into the cloud computing domain. This is also known as the *Cloud Forensics*.

Contributions: The contribution of this paper are as follows:

- We are proposing a comprehensive IoT-Based forensics framework, to reduce the dependence on the CSP and the network logs at the time investigation is going on.
- We are systematically analyzing and explaining the opportunities for IoT-Based forensics. We are designing this model to help researchers while focusing on very specific research sub-problems within the Cloud, Network and, Device forensics problem domain d for being a proper guideline to guide investigators and to be useful in training the fledgling investigators.

Organization: The organization of this article is as follows: We are describing related works in Section 2. Background knowledge on digital forensics and digital forensics models regarding the digital forensic investigation process models are discussed in Section 3. After that, we are proposing a comprehensive a DFI Framework for IoT and we are discussing how the proposed framework can work in section IV. Finally, we are concluding the paper in Section V.

II. Related Works

Significant research efforts have been observed in recent years on cloud and IoT forensics. In this section, we are presenting some related works: A recent study [3] has proposed a DFI Framework for IoT (DFIF-IoT). This framework is focused to support forensics investigative capabilities related to IoT in the future. The special features related to the framework is, it complies with the ISO/IEC 27043 standard: 2015.

Another paper [4] has suggested that application-aware forensic methods are needed to be implemented for specific IoT applications to ensure that data evidence remains within the context of the application. The study has considered three IoT applications (Wearable, Smart home & Smart city) and has proposed a forensic process which is application specific.

In another study [5] the researchers have tried to reduce the dependence on the Cloud Service Provider (CSP). The authors of the study have suggested a model for the collection of evidence by remaining outside of the cloud environment. The study has tried to mitigate the problem of over-dependence on CSP by suggesting a pre-implemented forensic server for evidence collection.

A survey work [6] has presented a study from the opinions of experts and practitioners related to digital forensics. This study has explored the important issues that are vital for cloud forensics like the definition, challenges, and opportunities.

Another similar study [7] has explored the critical challenges, recent proposals and technical solutions from the literature, related to cloud forensics.

A study [8] has tried to explores the issues related to cloud computing such as the role of CSP and it has also proposed a model named Open Cloud Forensics (OCF) model. The researchers also tried to validate the model using a civil lawsuit based case study.

Another similar work [9] has also explored the challenges and their probable solutions related to cloud forensics. This paper has also proposed a framework for cloud forensics and has also proposed a method for the collection of evidence from the Eucalyptus Cloud.

III. Background

In this part, the following topics will be covered: Digital Forensics, Digital Forensic (DF) models.

A. Digital Forensics

Various types of DFI techniques are required to breach the computing system for seizing the digital evidence to find out who has performed the digital attack, what malicious attacks have occurred, possible degree of risk and the potential outcomes after security assessing. The National Institute of Standards and Technology (NIST) defines digital forensics as an applied science to identify an incident, collection, examination, and analysis of evidence data. Based on the premise, digital forensics includes four processes:

- **Identification:** Identification of an incident and identification of the potential evidence.
- **Collection:** After identification of incident and evidence, the investigator tries to extract the IoT crime-related evidence from numerous physical & logical devices, cloud environment, network devices, and etc.
- **Organization:** Organization are closely related to two stages: Examination and analysis In the first stage, after extracting the IoT related information from the crime scenes investigator examines the evidence and observes its characteristics. In the second stage, all investigators analysis the evidence and try to establish a co-relation the receivable evidence for judgment.
- **Presentation:** In this process, evidence should be documented to store the chain of custody. An investigator generates a detail report with an explanation of presenting into the court for proof.

B. Digital Forensic Models

Since the first DF Research workshop [10], digital forensics models are being proposed on a regular interval but none of them are adopted fully. One of the models is- Digital Forensic Investigation Process (DFIP) model -presented by Valjarevic & Venter [11]. This model made an effort to incorporates the existing models. Other mentionable models include: Event-Based Digital Forensic Investigation Forensic Framework [12], Hierarchical, objective-based framework for digital investigation processes [13], Integrated Digital Forensic Process (IDFP) model [14], Generic Digital Forensic Investigation Framework for the Internet of Things(IoT) [3], Top-Down Forensic Approach Methodology [15] and Application-Specific Digital Forensics Investigative Model in an Internet of Things (IoT) [4]. Still, there are no models or standards that are accepted globally to collaborate with IoT.

IV. Proposed Framework for IoT-Based Forensics

In this section, Fig. 1 is introducing a DFI Framework for IoT. The proposed framework includes Readiness process, IoT forensics, Initialization process, Acquisition process, Investigation process, and Concurrent process.

A. Readiness Processes

The readiness process is also called iterative process. Readiness Processes including the following sub-processes:

- **IoT Scenario Definition:** In this step, all possible IoT scenarios should be examined to gather potential digital evidence with a proper risk assessment.
- **Identification of Potential IoT Evidence Sources:** In this step, all potential evidence sources like smart home,

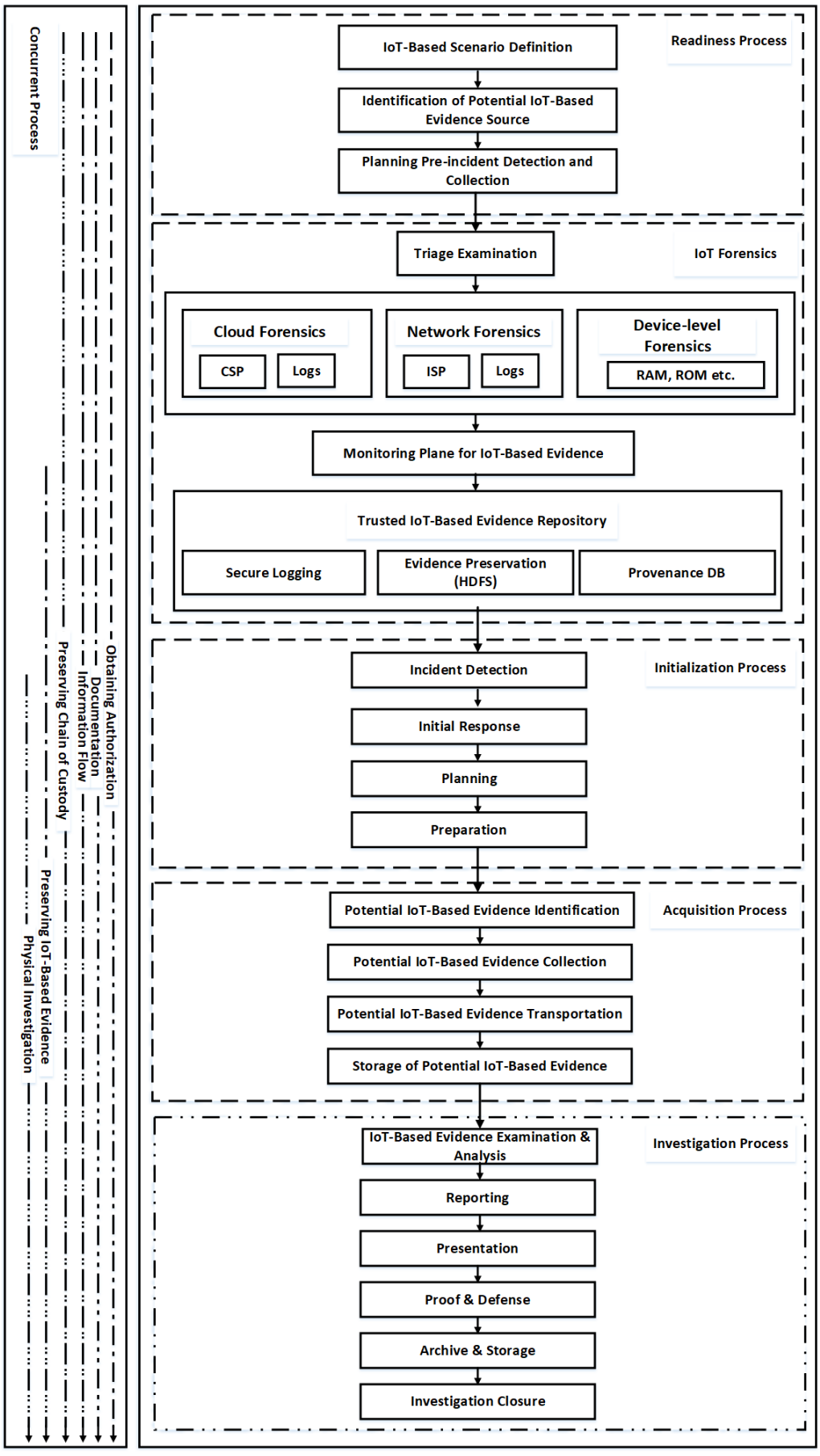


Fig. 1. Proposed Framework for IoT-Based Forensics

smart cities, health devices, cloud services, smart grid etc. are identified.

- **Planning Pre-incident Detection and Collection:** In this step, all activities for IoT-based pre-incident detection, storage and handling of IoT data should be defined.

B. IoT Forensics

IoT forensic is a special branch for DFI, where the identification, acquisition, organization, investigation, and presentation establishes an attacked incident with detail. There present numerous types of IoT forensic schemes, e.g. cloud forensic, Network forensic, device forensic etc.

1) Cloud Forensics: Data extraction and analysis is very difficult from the cloud environment because of lacking in data integrity, the huge amount of data, data diversity, multi-tenancy, jurisdiction, data duplication, and the high degree of virtualization. Previously, There is various type of solutions have been implemented and proposed for cloud forensics. But no one has considered about the CloudIOT forensic. However, For collecting the IoT forensics evidence, investigators fully dependent on the CSP or Cloud logs [5]. Cloud Logs are collected and stored the forensic evidence for investigators. So, most of the investigators and users are dependent on the CSPs server logs or network logs. To overcome this limitation, we are considering a Forensic Monitoring Plane (FMP) for data collection issues shown in Fig. 1. Data from the cloud will be stored in the proposed IoT-Based evidence repository through the FMP [5] [8]. The investigators are needed to analyze the evidence from the data repository. For that reason different types of cloud computing forensic tools will be needed such as Forensic ToolKit (FTK), E-detection, Encase, Recover, Recuva, Blade, etc. [16] running on the repository outside the cloud computing environment. After that, forensic evidence is forensically imaged bit by bitstream encryption and stored the data in the forensic server. Because we need to create the image of the original forensic evidence without any damage. These tools automatically capture the evidence from the current state of the cloud and stored the forensic evidence on the cloud server or trusted repository.

2) Network Forensics: This represents the IoT environments which incorporate different networks such as industrial network, home network, LANs, MANs, and WANs. Attack logs can be gathered from all these networks to be used with the DFI process.

But, data extraction and analysis are very difficult from the distributed network environment because of lacking in data integrity, the large volume of data, data diversity, data duplication, high-speed data transmission, data extraction location, data storage on the network devices, data privacy, access to IP addresses and, etc. So, most of the investigators and users are dependent on the ISPs server logs or network logs. To overcome this limitation, we are considering data collection issues presenting a Forensic Monitoring Plane (FMP) outside from the network environment infrastructure shown in Figure 1. Data from the network environment will be stored in the proposed IoT-Based evidence repository through the FMP [5] [8]. The investigators are needed to analyze the evidence from the data repository. For that reason, different types of network forensic tools will be needed such as network analysis tools, vulnerability assessment tools, network sniffing and packet analyzing tools, network scanning tools, and Network monitoring tools [17] running on the repository outside the environment.

3) Device-level Forensics: IoT-based evidence can be collected from physical devices like memory (RAM, ROM, or various secondary device), graphics, audio, video, image, files, and other IoT devices. But, data extraction and analysis are very difficult from the device environment because of lacking in data integrity, data diversity, data duplication, data extraction location, data privacy and, etc.

When a crime has occurred in a cloud, network, and device-level scenarios, an investigator can directly access (with secure logging) in the trusted evidence repository with user credentials and collect the pieces of evidence. Various types of Forensic tools are conducting in the cloud, network and, device-level when an attack occurs in the environment and finally, it collects the forensic images from the environment and stores in the evidence repository. In this paper, We are considering a secure evidence preservation module for evidence storage, as shown in Fig. 1. While evidence is stored on the module, it separates the evidence based on the IoT-based devices and its owner. In that way, we can separate the multiple user's data. This module stores the evidence by using public-private key based encryption (RSA, AES) [18] so that only investigator can access this module with secure logging and can examine or analyze the evidence. But it is difficult to manage very large data-sets. Hence, we are proposing a Hadoop Distributed File System (HDFS) [8] as the trusted IoT-based evidence repository, as shown in Fig. 1. This is Google's proprietary Map-Reduce framework, which is designed to solve problems involving big amounts of data sets computationally efficient. For big datasets, HDFS provides better performance than any other conventional repository system. Moreover, any hardware failures in the system will be automatically handled by this framework. We are using another module named Secure Provenance Module (SPM), which is used to store the access history of digital evidence. Finally, we are using the Application Programming Interface (API). An investigator can access (with secure logging) the trusted repository by using these APIs.

C. Initialization Process

Initialization Process(IP) including incident detection, first response, planning and preparation processes [11].

- **Incident Detection:** This is the first procedure of the initialization process. The procedures can define where the incident might occur.
- **Initial Response:** The initial response procedure should be included when the initial response to be detected IoT-based incident. Based on the incident types, that will disconnect the materials from the network environment and will detect the attacked equipment's.

TABLE I
COMPARISON OF PRESENT MODELS WITH PROPOSED MODEL

The Proposed model	Kebande et al. [3]	Perumal at el. [15]	Zawoad at el. [19]	Zia at el. [4]	Alex at el. [5]	Ragib at el. [8]	Valjarevic at el. [11]	Harbawi at el. [20]
Readiness Process								
Scenario definition	Yes	No	No	No	No	No	Yes	No
Evidence source identification	Yes	No	No	No	No	No	Yes	No
Planning pre-incident detection & collection	Yes	No	No	No	No	No	Yes	No
IoT Forensics								
Cloud forensics	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Network forensics	Yes	No	Yes	Yes	No	No	No	Yes
Device-Level forensics	Yes	No	Yes	Yes	No	No	No	Yes
Monitoring Plane	No	No	No	No	Yes	No	No	No
Evidence Repository	No	No	Yes	No	Yes	No	No	No
Initialization Process								
Incident Detection	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Initial Response	Yes	No	No	No	No	No	Yes	No
Planning	Yes	No	No	No	Yes	Yes	Yes	Yes
Preparation	Yes	No	No	No	Yes	Yes	Yes	Yes
Acquisition Process								
Evidence Identification	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Evidence Collection	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Evidence Transportation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Evidence Storage	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Investigation Process								
Evidence Examine & analysis	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reporting	Yes	No	No	Yes	Yes	No	Yes	No
Presentation	No	No	No	No	Yes	No	Yes	No
Proof & Defense	No	Yes	No	No	Yes	No	No	No
Archive & Storage	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Investigation Closure	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Concurrent Process								
Obtaining Authorization	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Documentation	Yes	No	No	No	No	No	Yes	No
Information flow	No	No	No	No	No	No	No	No
Chain of Custody	No	No	Yes	No	No	Yes	Yes	Yes
Preserving Evidence	No	No	No	No	Yes	No	Yes	Yes
Physical Investigation	Yes	No	No	No	No	No	Yes	No

- **Planning:** In this procedure, the investigator plans about the IoT-based Forensics Investigation. Planning should be included about the development of the relevant procedure such as procedure methodology, tools, technology human resource, etc.
- **Preparation:** Preparation activities should be included about the IoT-based forensics investigation. This should be included, but there are no limitations of preparation. IP Preparation including hardware, software, documents, training etc.

D. Acquisitive Process

The acquisitive process is closely related to the physical investigation [11]. It is purposed to uncover Evidence (DE) that are related to IoT. This process insists on the following procedures:

- **Potential IoT Evidence Identification:** This is the first procedure of the acquisitive process, the aim of this procedure is to identify the potential IoT-based evidence. But It is a very difficult and challenging task to identify the IoT pieces of evidence from the incident place in the cloud environment because the cloud environment deals with an exceptionally large amount of data.
- **Potential IoT Evidence Collection:** After identifying the IoT-based evidence from the incident scene, this evidence will be used for analysis in the later procedure. So, anyhow evidence should be collected from the crime scene for analyzing purposes.
- **Potential IoT Evidence Transportation:** During this procedure, after collection of the IoT-based evidence, it should be transferred from one place to another secure place for storing purposes that we can easily analyze next.
- **Storage of Potential IoT Evidence:** This is the last procedure of evidence acquisitive process. For examining and analyzing purposes, the potential evidence is needed to be stored. In this time extra care is needed to avoid the damage of digital devices. IoT evidence can be lost and damaged because of shocked by electricity, dead heat, cold, dirtiness, sudden power loss, function failure, and etc.

E. Investigation process

Investigation process consists of six procedures e.g. IoT Evidence Examination and Analysis, Reporting, Presentation, Proof & Defense, Archive & Storage and Investigation Closure [11]. In the beginning, evidence from the IoT devices is

examined and analyzed. After examining and analyzing the seized evidence, a report is needed to be created for presenting to all stakeholders. The stakeholder includes the judge, jury, accused, lawyers, prosecutors etc. The process is needed to be proofed based on all IoT-based evidence that is collected from an incident scene. Investigation process also checks the validity of the evidence for the investigation process. Deciding on the court, evidence could be accepted or rejected. Investigation closure concludes the investigation. After concluding the evidence documents it is needed to be archived & stored in a secure physical disk.

F. Concurrent Process

The concurrent process has been introduced by Valjarevic & Venter [11]. This process can improve the acceptability of IoT environments without any difficulty. Authorization is involved with the whole process of DFI framework which contains authorities, users, owners information. Moreover, the chain of custody is used to enhance the proficiency which stores the information related to the DFI. It also used to identify the information flow and secured the DFI in IoT-based environment. Furthermore, preserving the evidence of forensics is an important part of DFI. After collecting the evidence you should preserve and check the integrity of original evidence with a secure lock and must ensure the evidence will be unchanged. These process must confirm that the original IoT-based evidence is not altered. Sometimes, evidence needs to be transferred to the third parties. For these reasons, evidence should be stored in a physical device that also ensures the protection of evidence and easily handed to the trusted person. The whole concurrent process is used to conduct the DFI process properly with secure integration with IoT environment.

V. CONCLUSION

Even though IoT environments are immature; the standards of security and service are still evolving. Day by day, various research challenges have been arising on IoT forensic and some researchers are trying to solve the difficulties. However, The difficulties and problems that this paper has addressed previously are stated as follows: there is no global or international DFI method or framework which can facilitate the IoT environment. Based on the above analysis, we are introducing a comprehensive DFI process framework for the IoT environment that is able to enable more efficient and effective for DFI. The primary aim of our research is to conduct useful DFI for investigators and experts. In this paper, we are systematically analyzing and explaining the opportunities and try to solve some of the problems for IoT forensics but solving all the challenges are still very difficult. So, there are many aspects that need improvement. Moreover, the proposed framework reduces the dependency on CSP or cloud logs for acquiring evidence from the cloud. Moreover also reduces the dependency on ISP or network logs for acquiring evidence from the network. In the future, the proposed framework should also be included in more procedures and techniques for DFI.

REFERENCES

[1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (iot): A vision, architectural elements, and future directions," *Future generation computer systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[2] J. Deogirikar and A. Vidhate, "Security attacks in iot: a survey," in *I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2017 International Conference on*. IEEE, 2017, pp. 32–37.

[3] V. R. Kebande and I. Ray, "A generic digital forensic investigation framework for internet of things (iot)," in *Future Internet of Things and Cloud (FiCloud), 2016 IEEE 4th International Conference on*. IEEE, 2016, pp. 356–362.

[4] T. Zia, P. Liu, and W. Han, "Application-specific digital forensics investigative model in internet of things (iot)," in *Proceedings of the 12th International Conference on Availability, Reliability and Security*. ACM, 2017, p. 55.

[5] M. E. Alex and R. Kishore, "Forensics framework for cloud computing," *Computers & Electrical Engineering*, vol. 60, pp. 193–205, 2017.

[6] K. Ruan, J. Carthy, T. Kechadi, and I. Baggili, "Cloud forensics definitions and critical criteria for cloud forensic capability: An overview of survey results," *Digital Investigation*, vol. 10, no. 1, pp. 34–43, 2013.

[7] S. Alqahtany, N. Clarke, S. Furnell, and C. Reich, "Cloud forensics: a review of challenges, solutions and open problems," in *Cloud Computing (ICCC), 2015 International Conference on*. IEEE, 2015, pp. 1–9.

[8] S. Zawoad, R. Hasan, and A. Skjellum, "Ocf: an open cloud forensics model for reliable digital forensics," in *Cloud Computing (CLOUD), 2015 IEEE 8th International Conference on*. IEEE, 2015, pp. 437–444.

[9] D. R. Rani, S. N. Sultana, and P. L. Sravani, "Challenges of digital forensics in cloud computing environment," *Indian Journal of Science and Technology*, vol. 9, no. 17, 2016.

[10] P. Gary, "A road map for digital forensic research," in *Digital Forensics Research Workshop*, 2001.

[11] A. Valjarevic and H. S. Venter, "A comprehensive and harmonized digital forensic investigation process model," *Journal of forensic sciences*, vol. 60, no. 6, pp. 1467–1483, 2015.

[12] B. Carrier and E. H. Spafford, "An event-based digital forensic investigation framework," in *Digital forensic research workshop*, 2004, pp. 11–13.

[13] N. L. Beebe and J. G. Clark, "A hierarchical, objectives-based framework for the digital investigations process," *Digital Investigation*, vol. 2, no. 2, pp. 147–167, 2005.

[14] M. D. Kohn, M. M. Eloff, and J. H. Eloff, "Integrated digital forensic process model," *Computers & Security*, vol. 38, pp. 103–115, 2013.

[15] S. Perumal, N. M. Norwawi, and V. Raman, "Internet of things (iot) digital forensic investigation model: Top-down forensic approach methodology," in *Digital Information Processing and Communications (ICDIPC), 2015 Fifth International Conference on*. IEEE, 2015, pp. 19–23.

[16] M. P. Mohite and S. Ardhapurkar, "Design and implementation of a cloud based computer forensic tool," in *Communication Systems and Network Technologies (CSNT), 2015 Fifth International Conference on*. IEEE, 2015, pp. 1005–1009.

[17] R. Joshi and E. S. Pilli, "Network forensic tools," in *Fundamentals of Network Forensics*. Springer, 2016, pp. 71–93.

[18] C. Stergiou, K. E. Psannis, B.-G. Kim, and B. Gupta, "Secure integration of iot and cloud computing," *Future Generation Computer Systems*, vol. 78, pp. 964–975, 2018.

[19] S. Zawoad and R. Hasan, "Faiot: Towards building a forensics aware eco system for the internet of things," in *2015 IEEE International Conference on Services Computing (SCC)*. IEEE, 2015, pp. 279–284.

[20] M. Harbawi and A. Varol, "An improved digital evidence acquisition model for the internet of things forensic i: A theoretical framework," in *2017 5th International Symposium on Digital Forensic and Security (ISDFS)*. IEEE, 2017, pp. 1–6.

Path and Partner Selection in Unstructured Resource-Constrained Environments Based on Node Psychology in the Context of IoT

Tanusree Debi
Department of Computer Science and Engineering
Chittagong University of Engineering and Technology
Chittagong, Bangladesh
tanusreedebi11@gmail.com

Assaduzaman
Department of Computer Science and Engineering
Chittagong University of Engineering and Technology
Chittagong, Bangladesh
asad@cuet.ac.bd

***Abstract* – Internet of Things (IoT) is the network of the huge number of resource-personated nodes whose main feature is the reliable communication and fast data exchange among devices in any network topology. In the upcoming world of internet, proper selection of desire or appropriate destination node and finding the optimized path from the huge number of the activated nodes is a very important task. Here many nodes are available in one or more hop distance and each node has various network parameters in the IoT world, so the selection of best path and nodes is quite complex. For this reason, we need to develop an algorithm to find out the best path towards the destination node considering significant network parameters. This paper proposes a general algorithm considering any network topology for discovering the path towards the destination node from the source node and give a solution to the contradiction which may arise if there is more than one consonant path. The proposed algorithm considers the node property of all node for calculating the behavior value, path influence behavior value and finally discover the optimized path from source node to destination node. The proposed algorithm is also evaluated by simulation to determine the impact of parameters in path determination.**

Key words- Node, Path and Partner selection, Path Influence Behavior, Node Parameter, IoT.

I. INTRODUCTION

In a cooperative service scenario like IoT, where many nodes are present in the IoT field, the most appropriate node selection can be a complex task for transmission of data [1]. Nowadays, IoT is the rising technology in scenarios such as smart city, smart home, eHealth, service, etc. Information is generally carried over by lossy networks and low power consists of smart nodes, many of those have limited power capacity and processing capability, and over very dynamic IP-based connections [1]. IoT operation depends on the identification, automatic management and use of a huge number of diverse physical and virtual objects (i.e. both physical and virtual representation), connected to the Internet [2], [3]. Additionally, the identification of a particular object and its addressing scheme ability to dynamically configure their addresses is also an open issue [4].

Because of many available nodes in network topology, proper selection of nodes and path is an important decision for data exchange between nodes to nodes. Therefore, the behavior of nodes may be changed after a certain time period. Also, the influence behavior of nodes may differ depends on specific communication service requirements. In IoT research field, the behavior of path and node selection is still in the initial stage. If a node wants to start communication, there may be a destination node is defined or not. Another case is that destination partner can be far away one or more than one hop distance. Some research work has been done on node discovery, node interpretation, and established the connections among available nodes in the world of IoT [5], [6].

A representation for the discovery and selection of a node in large-scale, resource-constrained environments in the context of service and application requirements was proposed and assessed [5], [6]. This representation relies on a fundamental mathematical approach which describes and predicts the positive outcome of human interactions and communications in terms of long-term relationships and recognizes various key variables in resource-constrained environments [5], [6]. The methodology is used to develop this model adopted from [7]. The mathematical model of the network parameters, such as Device Class, Influence function and Influence behavior are shown in [5], [6]. Three network architectures are proposed for

data storage and reliable communication in the context of IoT [5]. An algorithm named "choosing a partner" is proposed in [6] and extended in [5]. An algorithm named "the behavior of a whole path" is proposed in [5], but there are some limitations. Also, a new priority of node called Emergency Priority and a new parameter called Reliability & service type properties proposed and performed simulations on behavior parameters and convergence of the behavior value [5].

The following paper proposes an algorithm for node selection and path evaluation based on path influence behavior, behavior value, node parameters, which is applicable for any kind of network topology either destination node is one hop or more hop distance and help us to find the optimized path. There was no such algorithm for path and partner selection in [5], [6].

The rest of the paper organized as follows: Section II addresses the related works. Proposed Methodology is represented in Section III. Section IV describes our experimental evaluation, Section V provides the conclusion of the proposed paper and Section VI suggests some future work over the algorithm.

II. BACKGROUND

A.Node Parameters:

Here we consider some IoT node parameters which help us to evaluate the fastest and best path to communicate among nodes [5] [6]. The Node parameters are shortly described here:

Device Class: Device Class is a number which indicates the potentiality of the node. It is composed of resource parameters (device speed, processing power, data storage, battery life, digression from the power efficiency defined by the throughput, manufacturer). The range of Device Class is [4; 20] where 4 means low power device and 20 means high power device.

Free Resources: Free resources is another node parameter which discloses the actual free resources present in that node. The range of free resources is [0; 0.9] where 0 means 100% of the resources are occupied and 0.9 means 100% of the resources are free.

Reliability Score: Before any communication, the source node needs to know how reliable the candidate nodes. The *Reliability* is another IoT node parameter which gives information that how reliable a node is. For IoT node, we use reliability instead of security. The range of Reliability parameter is between [1; 10] where 1 means least reliable and 10 means most reliable.

Influence Function: Another most important node parameter is Influence function. It is a function of Influence Behavior which indicates the influence of one node has on its partner. It is also used in the evaluation process of communication is described in [5], [6] with its score and description.

B. Modelling Node Interaction:

In [5] [6], the authors introduced a mathematical model which is used to define the *Influenced Behavior* scores. It is a score showing the strength of a node depending on its available resources and the demands it has to meet. The relationship equation of this mathematical model is given below:

$$N_{1(t+1)} = I_{N_2 N_1}(N_2)t + r_1 N_{1\,t} + a \quad (1)$$

$$N_{2(t+1)} = I_{N_1 N_2}(N_1)t + r_2 N_{2\,t} + b \quad (2)$$

Where N_1 is the Influence behavior of Node 1 and N_2 is the behavior score of Node 2. The details of the above equations are described in [5], [6]. Behavior has an only positive value and this value is used in connection establishment among nodes.

C. Algorithm for choosing a partner:

An algorithm for choosing a new partner in IoT topology was proposed in [6] which was extended in [5] including Reliability parameter. When a node wants to communicate with another node away from one hop distance and there is no fixed destination node, this algorithm is used to find out the appropriate partner node from the available nodes. At first, the source node calculates its uninfluenced behavior score and then sends the application requirements to its all neighbor nodes in one hop interval. The neighbor nodes receive the requirements and send back their own Device Class and influence behavior score to the source node. Then the source nodes choose the best partner node considering the Behavior value, Device Class and Reliability score.

D. The Behavior of a whole path:

A procedure for establishing an entire path and calculate behavior score named "Behavior of a Whole Path" was described in [5]. By this mechanism, we can calculate the Behavior value of the node to node communication and used it in the further calculation. For a specific proposed scenario, source node broadcast its requirements towards all its neighbors. Neighbor nodes of the source node modified the requirements and broadcast to their neighbors and so on. When the destination node receives the requirements it calculates its own behavior value and sends this value towards the node of the previous path. Then, a node of the downstream path calculates the path influence behavior.

The weak points of this algorithm are:

- This mechanism is only applicable to the proposed scenario shown in [5]
- There is no exact algorithm without only a scenario model.
- If there is more than one candidate path, this algorithm can't give any solution.

III. PROPOSED METHODOLOGY

A. SYSTEM ARCHITECTURE:

The proposed methodology is applicable for any kind of network topology. Here, we consider the following network topology, for example,

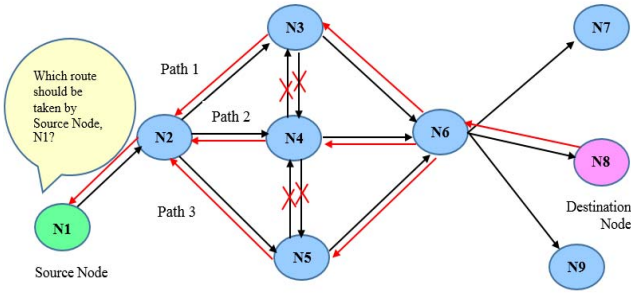


Figure: Sample Network Topology

Fig. 1: A network topology contains 9 nodes

In the scenario, Source node N_1 wants to find a reliable path towards Destination node N_8.

There are three possible paths -

1. $N_1 \rightarrow N_2 \rightarrow N_3 \rightarrow N_6 \rightarrow N_8$
2. $N_1 \rightarrow N_2 \rightarrow N_4 \rightarrow N_6 \rightarrow N_8$
3. $N_1 \rightarrow N_2 \rightarrow N_5 \rightarrow N_6 \rightarrow N_8$

We propose the following algorithm to choose an optimal path by the Source node to communicate with the Destination node for exchanging the data.

B. PROPOSED ALGORITHM:

The proposed algorithm Path and Partner Selection in Unstructured Resource-Constrained Environments Based on Node Psychology in the Context of IoT is applicable for any kind of network topology. This algorithm is applicable for one hop or more hop distance and also can give an exact solution if there is more than one candidate path.

The steps of the algorithm are as follows:

1. Appended the network parameters with the nodes present in Network Topology.

2. Define the Source node (N_s) and the Destination node (N_D) among all the nodes.

3. Setup Quality of service (QoS) requirements for the Source node. Quality of service means application requirements which are essential for any application. Suppose, for a certain application, source node needs Delay 500ms, Throughput 1.5Gbps.

4. Source Node (N_s) wants to find a path to the Destination node (N_D) (n hop distance).

5. N_S sends a package including its application requirements, the name of destination node as header to its neighbors.

6. Neighbor nodes modify requirements and forward packets to their neighbors and continue.

7. N_D receives requirements and identifies itself as a final destination, calculates its influenced behavior based on the modified requirements and replies the packets along with the previous path.

8. Every node in the path calculates the path influence behavior, an average of its influence behavior and the influence behavior of the previous neighbor node.

9. N_D receives feedback.

10. It checks whether there is one path having the highest Influenced Behavior or not:
a. If Yes, a path is found;
The source node can communicate with the destination node through this path.
b. If No, continue to Step 11;

11. The node analyzes whether the nodes with an equal highest value of the Influenced Behaviors are from the same Device Class or not:
a. If Yes, continue to Step 13;
b. If No, continue to Step 12;

12. The capable partners with the highest scores for the Influenced Behaviors and with the highest value for the Device Class among all potential partners are saved. Continue to Step 10;

13. Choose the potential partner(s) with the highest value of the Reliability score
14. The node checks whether there is only one potential partner left or not:
a. If Yes, a partner is found;
b. If No, a partner is picked randomly from the remaining pool of partners.

C. REQUIRED EQUATIONS:

Influence Behavior Score:

$$N_{1(t+1)} = I_{N2\to N1}(N_{2t}) + r_1 N_{1t} + a \quad (3)$$

$$N_{2(t+1)} = I_{N1\to N2}(N_{1t}) + r_2 N_{2t} + b \quad (4)$$

Uninfluenced Behavior:

$$B_{UB} = \frac{Device\ Class}{1 - Free\ Resources} \quad (5)$$

Influence Function:

$$I = \alpha I_L + \beta I_D + \gamma I_{PE} + \delta I_S + \varepsilon I_T + \zeta I_R + \eta I_{PDR} + \theta I_{BS} + \iota I_J + \kappa I_{CP} + \lambda I_{EP} + \mu I_{IC} + \nu I_{DC} \quad (6)$$

Path Influence behaviour:

$$PIB = \frac{I_{Current\ Node} + I_{Adjacent\ Node}}{2} \quad (7)$$

IV. Experimental Evaluation

A. Sample Input:

The network topology contains 8 nodes. N_1 is the source node and N_8 is the destination node. The node parameter (Device Class, Free resources, Reliability score and Influence behavior) are given for all 8 nodes. The Quality of Service (Delay time, Latency, Throughput, packet delivery ratio, etc.) are attached to the source node. It is not obvious that all quality of services is required for any communication. Only the interested parameters are set up for distinct communication. The requirements for any service is delivered from source node towards all the nodes in the network topology. The proposed algorithm is applied to the network topology [Fig. 1] and we will get the best path for communication.

B. Numerical Example:

Influence Function,

$$= \alpha I_L + \beta I_D + \gamma I_{PE} + \delta I_S + \varepsilon I_T + \zeta I_R + \eta I_{PDR} + \theta I_{BS} + \iota I_J + \kappa I_C + \mu I_{IC} + \nu I_{DC}$$

```
Node 8 IB : 60
Node 6 PIB : 40
Node 3 PIB : 65
Node 2 PIB : 44.5
Path THREE Selected : Final Influence Factor : 44.5
```

C. Sample Output:

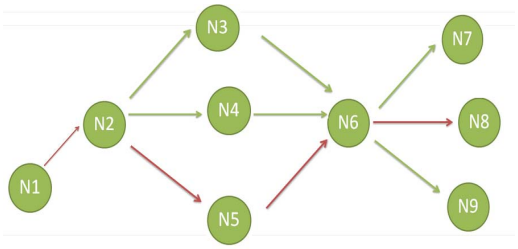


Fig. 2: Simulation output of sample network topology

D. Accuracy Measure

The accuracy of the system can be observed from the theoretical and experimental result. We have tested seven samples for all the network parameters range. The performances of the proposed system are listed here-

Subscript a, b, c, d, e, f for node number n= 4, 5, 6, 7, 8, 9 respectively.

Average accuracy, A_{averge} =∑Ai

$$= \frac{A_{a1} + A_{a2} + A_{b1} + A_{b2} + A_{c1} + A_{c2} + A_{d1} + A_{d2} + A_{e1} + A_{e2} + A_{e3} + A_{e4} + A_{f2}}{14}$$

$$= \frac{100 + 99.99 + 99.99 + 100 + 99.99 + 100 + 99.99 + 100 + 100 + 100 + 99.99 + 100 + 100 + 99.98}{14}$$

$$= 99.995\%$$

We can see that, the accuracy of the performance doesn't depend on the number of nodes. But if the number of node in a path increase the value of path influence behavior will also increase.

E. Simulation Result

a. Relationship between Device Class and Path Influence Behavior

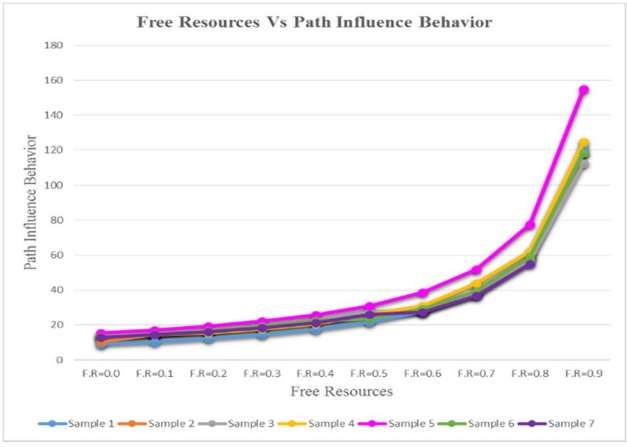


Fig.3: Device Class vs. Path Influence Behavior

In Fig. 3, X-axis shows the Device Class [4; 20], where 4 indicates the less powerful device and 20 indicates the most powerful device and Y-axis shows the Path Influence Behavior.

Observations:

- From Fig. 3, we can see that increasing the value of Device Class will increase the value of path influence behavior which means that powerful device has the highest influence value and prefer for communication among device.
- The lower the value of Device Class, the lower the value of path influence behavior.
- The Path Influence Behavior value doesn't depend on the number of nodes, but depends on the value of Device Class.
- The lower the Device Class, the free resources play a significant role in the value of path influence behaviour.

b. Relationship between Free Resources and Path Influence Behavior:

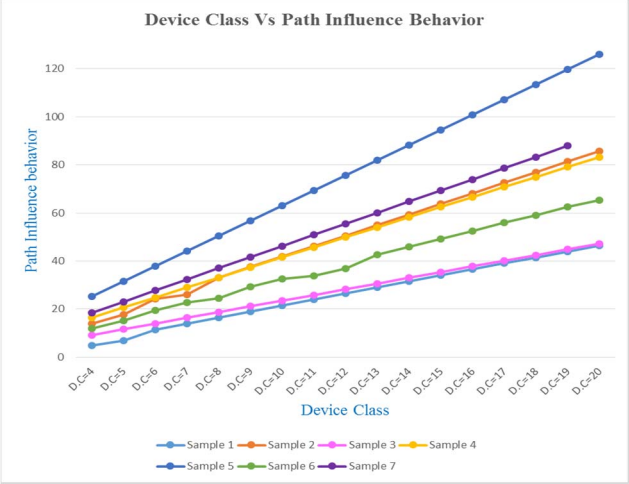


Fig. 4: Free Resources vs. Path Influence Behavior

In Fig. 4, X-axis shows the Free Resources [0.0; 0.9], where 0.0 means 100% resources are occupied and 0.9 means 100% resources are free and Y-axis shows the Path Influence Behavior.

Observation:

- From the figure, we can see that increasing the value of Free Resources will increase the value of path influence behavior that means device contains free resources to get chance first for communication.
- The lower the value of free resources, the lower the value of path influence behavior.
- The Path Influence Behavior value doesn't depend on the number of nodes but depends on the value of free resources.
- Here, we can see an exponential relationship between the Path influence behaviors and free resources.

From the range [0.6; 0.7] of free resources, the value of path influence behavior increases extremely fast.

Based on the figures and simulation result, the following conclusion can be drawn:

- The greater the value of Device Class the greater the possibility for the value of the Behavior to be with higher value. That means that more powerful device can perform well comparing low powerful device. For example, if any device has high processing power, RAM speed, battery life will show the high behavior value and get a chance than a low power device for communication.
- The greater the value of Free Resources the greater the possibility for the value of the Behavior to be with higher value. That means that the device which contains more free resources easily occupied for data communication.
- There is a linear relationship between Device Class and Path Influence Behavior which means that increasing the device class will increase the path influence value and may be capable of communication.
- There is an exponential relationship between Free Resources and Path Influence behavior which means that for any communication we need to increase the free resource value.

c. Comparison with Existing Algorithm

"Behavior of whole path" algorithm which is proposed in [1] is applicable for one topology contains 5 nodes. The proposed algorithm has removed this limitation. The proposed system is built to analyze any network topology for finding out an optimized path in order to node-to-node data communication.

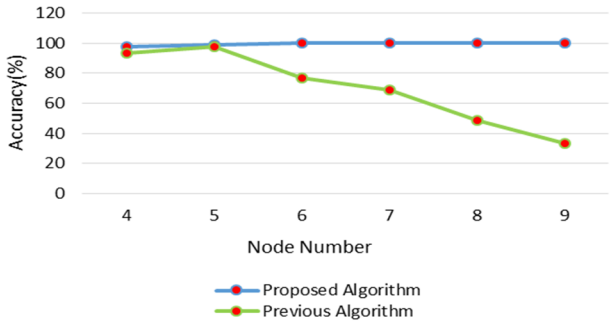


Fig. 5: Comparison of previous and proposed algorithm in respect of Node number

We have tested both previous [1] and proposed algorithm for 7 network topology samples contains several nodes. In Fig 5, X-axis shows the node number and Y-axis shows the accuracy of successful path and partner selection based on Path Influence Behavior. We can see that the accuracy of the previous algorithm is decreasing when we increase the number of nodes, but the accuracy of the proposed algorithm is near to constant that case.

V. CONCLUSION

In IoT network topology, path and partner selection for data communication is a very important and essential part. The proposed Path and Partner selection algorithm in Unstructured Resource Constrained Environment solved the path and partner selection problem when the destination node is more than one hop away from the source node in the hybrid flexible architecture. This algorithm can enable to discover the most suitable route for the need for a complete source-to-destination connection establishment. The proposed algorithm can also solve the contradiction when the Path influence behavior of two or more route is the same. Here, we examined seven samples including a variable number of node and evaluate the system by using a large set of data range. The performance is also measured by using summation and average calculation. In the end, we analyzed the relationship between Network parameters such as Device Class and Path Influence Behaviour from simulation result to validate the proposed algorithm.

VI. FUTURE WORK

The following enhancements can be further looked into:

- The Device Class and Free Resources parameter should be further researched and should increase the range.
- Should reduce the exponential time for decision making into logarithmic or linear time when a number of node are huge.
- Can introduce more network parameter, security mechanism and different priorities such as Emergency Priority which can classify the service type.

REFERENCES

[1] J. Ko, A. Terzis, S. Dawson-Haggerty , D. E. Culler , J. W. Hui , P. Levis "Connecting low-power and lossy networks to the internet" in *IEEE Communications Magazine*, 49(4), 96–101, April 2011.

[2] O. Vermesan and P. Friess, "Internet of Things – Global Technological and Societal Trends," *River Publishers*, 2011.

[3] Dave Evans, "The Internet of Things: How the Next Evolution of the Internet Is Changing Everything", April 2011

[4] T. Savolainen, et al, "IPv6 Addressing Strategies for IoT," *in Proc. of IEEE Sensors Journal*, Vol. 13, No. 10, October 2013.

[5] S. Kasabova, M. Gechev, V. Vasilev, A. Mihovska, V. Poulkov and R. Prasad "On Modeling the Psychology of Wireless Node Interactions in the Context of Internet of Things" *in Wireless Personal Communications* 85(1):101-136, June 2015.

[6] M. Gechev, S. Kasabova, A. Mihovska , V. Poulkov and R. Prasad ''Node discovery and interpretation in unstructured resource-constrained environments" *in Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronics Systems (VITAE),* 2014 4th International Conference on GWS2014, 11-14 May 2014, Aalborg.

[7] J. Gottman, J. Murray, C. Swanson, R. Tyson, Kristin R. Swanson, "The Mathematics of Marriage: Dynamic Nonlinear Models", A Bradford Book, 2005.

[8] Z. Shelby, "Internet Engineering Task Force (IETF) RFC 6690", August 2012.

[9] J. Hollar, V. Tsiatsis, C. Mulligan, S. Karnouskos, S. Avesand and D. Boyle "Machine-to-Machine to the Internet of Things: Introduction to a New Age of Intelligence" Elsevier. ISBN 978-0-12-407684-6.

[10] Monnier, Olivier "A smarter grid with the Internet of Things" in *Texas Instruments*, May 2014.

[11] Y. Erlich "A vision for ubiquitous sequencing" Genome Research *published by Cold Spring Harbor Laboratory Press*, May 2015.

[12] T. A. Butt, et al., "Adaptive and Context-aware Service Discovery for the Internet of Things," in *Proc. of 6th conference on Internet of Things and Smart Spaces (ruSMART 2013), St. Petersburg, Russia,* pp. 36-47, 2013.

[13] M. Nitti, R. Girau & L. Atzori "Trustworthiness management in the social internet of things*" IEEE Transactions on Knowledge and Data Engineering*, 26(5), 1253–1266, 2014.

Bridge Strength Preservation by Automatic Traffic Density Control: an IoT Application

Aisha B. Rahman
Department of Electrical & Electronic Engineering
University of Chittagong
Chittagong, Bangladesh
aiisha331.ar@gmail.com

Mohammed Shahriar Kamal
Department of Electrical & Electronic Engineering
University of Chittagong
Chittagong, Bangladesh
srksmaran786@gmail.com

Ariful Islam
Department of Electrical & Electronic Engineering
University of Chittagong
Chittagong, Bangladesh
ariful@cu.ac.bd

***Abstract*— Bridges are expensive and important constituents in the transportation network and any failure associated with it results in a huge loss of money, time and sometimes, lives. Bridges are important structures considering its use and purpose-serving. It may fail due to aging, loading conditions and natural disasters. Every bridge construction specifies a limit of weight of the vehicles that can ride it. However, still there is a very large record of bridge failures due to overloading. The proposed system is a wireless sensor network which monitors the weight of the vehicles on the bridge and provides immediate protection when the failure of the structure is nearly estimated due to excessive load. This design uses embedded sensors (piezoelectric sensors and accelerometers) over the bridge deck and its girder for the continuous governance of the bridge conditions. Data collected from the sensors are sent to a database server connected to the internet (forming an advanced wireless technology called Internet of things (IoT)) to activate traffic control when the limit crosses in order to eliminate high losses which would've occurred if the bridge have collapsed.**

Keywords—accelerometer, piezoelectric sensor, wireless sensor network, internet of things.

I. INTRODUCTION

Strength and Stability of a bridge are indications of the prosperity of a society and the safety of lives and properties of its dwellers. This makes the proper monitoring of bridges to preserve its strength an important issue. Strength to Weight ratio is a measure of a structures efficiency. It is the ratio of the structures weight to the weight it carries. A ratio>1 indicates stability of the structure. The proposed system ensures this ratio to be above 1 to prevent the bridge from collapsing. It employs a wireless sensor network (WSN) over the bridge that reports the state of the bridge to a remote data acquisition center. Due to all the inconveniences associated with the inspection of the health of such huge structure of a bridge, the development of effective and automated damage diagnosis techniques has been made possible through Wireless Sensor Networks. Based on the data collected simultaneously, the engineer will be able to estimate the state of fatigue or other limit states and provide forecasts of the probable stake of the bridges health and with modern technology, real time control of structures is accomplished. A WSN system consists of 3 units, a sensor unit (also known as sensor nodes), a remote data processing unit and an evaluation system. There are different methods for building up a WSN, which are discussed in the following literature review. This paper designs an IoT platform for monitoring the bridge state (both deck and girder) with automatic control action using arms barriers and spike barriers in case of any emergency. So the authority doesn't have to be at the bridge itself to ensure safety. Moreover, considering the cost of the system due to its great size, the cost can be minimized by cutting off the cost of power supplies i.e self-powering the sensor nodes. The system can thus be divided into 5 subsystems- energy harvesting, data collection, data transmission, data monitoring and control.

II. LITRTATURE REVIEW

Monitoring the strength of a bridge falls within the fields of Structural Health Monitoring of bridges. SHM aims to give, at every moment during the life of a structure, a diagnosis of the "state" of the constituent materials, of the different parts, and of the full assembly of these parts constituting the structure as a whole [1]. The SHM process involves the observation of a system over time using periodically sampled dynamic response measurements from an array of sensors. The basis of different structural health monitoring systems may be damage detection, loading condition, vibration based etc. Some commonly used sensors in SHM systems are mentioned briefly below,

A. Accelerometers

Accelerometers used for SHM are either piezoelectric or spring-mass accelerometers. Piezoelectric accelerometers are light, small, and operate over wide acceleration and frequency ranges [2]. On the contrary spring-mass accelerometers are heavier and requires large area and also operates over a limited range of frequencies.

B. Strain Sensors

Strain sensors used for SHM can be classified as piezo resistive or embedment strain gauge based. Cement-based strain sensors are typically piezo resistive and are capable of measuring strain. These sensors generate signals with an incredibly low frequency (< 1 Hz). Embedment strain gauges can be used for measuring strains inside concrete structures [3].

C. Corroision sensors

Strain Corrosion sensors typically measure the change in resistance introduced by the erosion of the sensor corrosion wafer. These sensors typically measure a structure's overall corrosion [11, 12].

D. Optical Fiber Sensors

The most common type of optical fiber sensor is the fiber Bragg grating sensor [13, 14]. These sensors can be used to measure parameters such as strain, temperature, pressure, and other quantities by modifying a fiber so that the quantity to be measured modulates the intensity, phase, polarization, and wavelength or transit time of light in the fiber. [15]

This project is based only on the general matrices i.e vibration, load, tilt or slope etc. There are different approaches

of implementing SHM of bridges starting from centralized approaches to distributed ones. The concept of distributed approach began with the development of wireless sensor networks. Besides data sensing, tracking and transmitting data is a challenge in SHM systems. With the advancement in technology, Internet of Things (IoT) has made it possible to integrate SHM with the Internet to track data anytime anywhere. In order to connect the sensors to the gateway wirelessly there are many protocols available such as Bluetooth, Zigbee, SIGFOX, Snow, Weightless, LoRaWAN etc. From comparative studies of these protocols it is clear that Bluetooth and Zigbee are compatible with small range IoT applications while SIGFOX, Snow and LoRaWAN are for large scale IoT applications requiring coverage area of few kilometers such as which we are working on. So far, the existing SHM systems are only monitoring which gives a record of the structures behavior for further studies or decisions such as one given in ref. [4]. However, none of them has an immediate protection. Adding an immediate protection feature would double its usefulness. The ability to prevent the breakdown of the bridge and loss of lives will give a clearer image of the importance of an SHM system. Also repairing the bridge before it collapses would decrease the maintenance cost to a greater extent. This implies the need of a bidirectional communication. However SIGFOX and Weightless are unidirectional. Snow is a network architecture that makes use of TV white spaces (allocated but unused TV channels) to operate. This makes a long communication range possible because VHF and UHF have excellent propagation characteristics over long distances and also a nice obstacle penetration, making communications in urban scenarios more reliable. This technology offers the possibility of deploying thousands of sensors and connecting them to the base station without the need to use either a multi-hop network or gateways[9]. Snow can cover large distances over 1.5km with relatively low transmission power, enabling longer life of batteries used in the sensor nodes. However, Snow is still under development and modules are not available yet. The advantages offered by Snow is also offered by LoRaWAN. The LoRa end device communicates with gateways that employ LoRa with LoRaWAN. LoRa gateways pass raw LoRaWAN packets from the end devices to a LoRa network server with a high throughput based on the backhaul interface, which is typically 3G or Ethernet. Accordingly, LoRa gateways act as a bidirectional communication or protocol adapter with the LoRa network server [10].

III. Internet of Things Platform for Bridge Strength Monitoring

At present the most efficient way to implement an SHM system is using IoT in order to monitor the structure from any location. The internet of things is a system of interrelated computing devices, mechanical and digital machines or objects that are provided with unique identifiers and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction [5]. It is a rising trend where enormous devices (things) are connected to the internet. To implement an SHM system using IoT, the required hardware components fall within 4 categories; sensors, microcontroller, transceiver and gateway. This design uses Accelerometers and Piezoelectric sensors at the sensor nodes to collect information about the state of the bridge (both deck and girders) and send it to a database server. The data is fetched and displayed on the web application for monitoring. In this design, we propose to ensure safety as data reaches the threshold by closing the entrance of the bridge immediately using automatic barriers or spike barriers which is discussed in section IV.

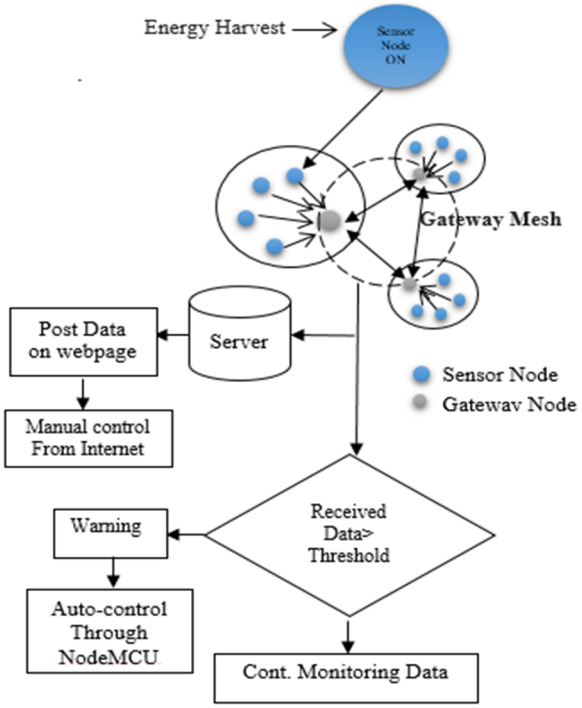


Fig. 1. Overall architecture

A. Sensors

1) Piezoelectric Vibration Sensor

Piezoelectric transducers respond to deflection, strain, vibration, flex, mechanical stress etc. by generating electrical charge proportional to the applied forces. Piezoelectric Vibration Sensors operate on the piezoelectric effect. A voltage is generated as the sensor is vibrated which is applied to the data pin of the microcontroller to be processed prior to transmission over the LoRaWAN network.

2) PiezoelectricPressure Sensor

Piezoelectric effect occurs not only due to vibration of the piezo material but also occurs when we pressurize the material. Besides piezoelectric vibration sensors we have used piezoelectric pressure sensors integrated in series which together, will measure the force on the bridge per unit area (pressure) by the vehicles to assure the pressure to be below the threshold value. The voltage that the sensor gives is proportional to the applied force, and by numerous calibration the resulting relationship between pressure and voltage for 2mm piezoelectric plate was found to be,

$$Y=0.02011X \quad (1)$$

Where, Y is the generated voltage (V) and X is the applied force (N). [7] The pressure created by the vehicles passing through the piezo sensor sheet (a sheet prepared by connecting a large number of piezo sensors in series to cover up a wider space to increase accuracy in measurement of the force applied) is measured by the given formula.

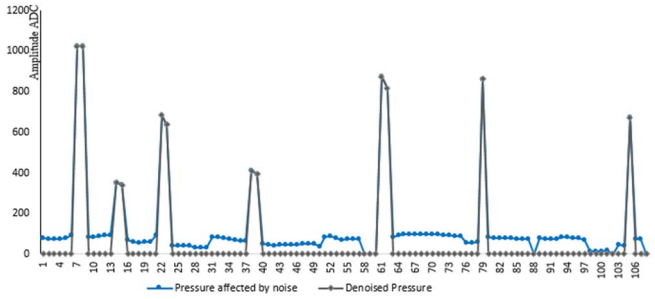


Fig. 2. Vibrational noise compensation using microcontroller

The vibration of the bridge will also affect these piezo sheets. However the voltage generated due to vibration has very small amplitudes compared to those generated due to mechanical force. The compensation to this low amplitude noise is performed by the microcontroller used. The blue lines in Fig. 2 shows noise due to bridge vibration. These low amplitudes are clamped down to 0 level by the microcontroller. The resulting undisturbed response of the piezoelectric pressure sensor is represented by the grey lines. Upon filtering out the noise, the voltage is converted to pressure and transmitted over the LoRaWAN network and received in the gateway to pass it to the database server. Readings from all the sheets within the unit area are integrated and compared simultaneously with a threshold value.

3) ADXL335 Accelerometer

On a bridge, not only the deck gets damaged but also the girder, on which the bridge stands, may get damaged especially in cases where the bridge is built over a river. The angles at which the girders are built is fixed and any deterioration suspects a collapse of the bridge**.** ADXL335 has been chosen to detect this in this design. It is a small, thin, low power, complete 3 axis accelerometer with signal conditioned voltage outputs. It can measure the static acceleration of gravity in tilt sensing applications, as well as dynamic acceleration resulting from motion, shock or vibration [6]. Acceleration is measured by ADXL335 in g unit. Max voltage for 0g is 1.65V. From the X_{out}, Y_{out}, Z_{out} output of the module, we calculate the angle of inclination.

$$A_{xout} = \frac{\frac{X_{out} \times V_{ref}}{1024} - 1.65V}{.330V} \quad (2)$$

Similarly we calculate for A_{yout} and A_{zout}.

The angle of Inclinations are calculated from the formula,

$$\theta = \operatorname{atan}\left(\frac{A_{xout}}{\sqrt{A_{yout}{}^2 + A_{zout}{}^2}}\right) \quad (3)$$

Similarly we calculate for angles ϕ, ψ. As an inclination is detected on any girder, immediately the entrance of the bridge will be closed.

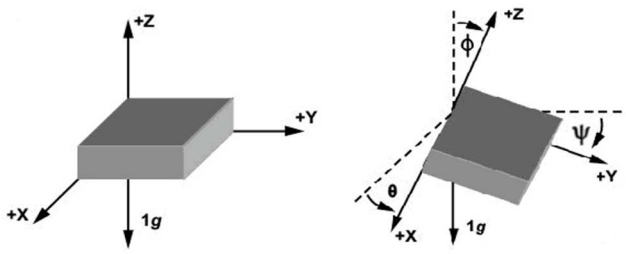


Fig. 3. Illustration of angle of inclination [17]

B. Microcontroller

Microcontrollers processes the data before transmitting it, For IoT applications, to exchange data between web server and web applications, the data collected is converted to JSON format by the microcontroller and fed to the web application from the server. There are 3 different typical microcontrollers which are used in WSN application. They are Microchips, AVR series and MSP430 series. More of which can be known from the reference [8]. For prototyping, one of the most successful and most popular microcontrollers is the Arduino platform. Arduino is a development board that houses a microcontroller on it. To make the nodes as small as possible, Arduino Nano can be used to implement the sensor nodes

C. Transiever

Each node transmits and receives data over the WSN via a transceiver. The right transceiver to be employed is chosen depending on the size of the network. LoRa (Long Range wireless data telemetry) technology is the most exciting technology for longer range communication. LoRa technology offers a 20dB link budget advantage compared to existing solutions, which significantly extends the range of the application while delivering the lowest current consumption to maximize battery life [Semtech], fulfilling all the requirements of the transceiver an SHM system for bridges might require.

D. Gateway

The topology of the nodes and the gateways is an important issue considering the huge size of the structure. ESP8266 is a low cost Wi-Fi module chip that can be configured to connect to the internet for IoT applications. However it has been designed without any holes for pin headers, it is meant to be mounted on a PCB. ESP8266 NodeMCU is preferable for easy bread-boarding and can directly be connected to a computer to start writing programs and connect to the Wi-Fi network.

In case of designing a large scale IoT system such as that one for monitoring a bridge, the availability of the gateway, without crashing or any security breaches is important. A peer-to-peer connectivity between a large numbers of gateways can be employed. This constitutes a number of gateways connected to one another via a mesh network and clusters of sensor nodes to each gateway as shown in Fig. 1. If one goes down, the redundancy of that network ensures reliability and the continuity of the cloud communication for commands and storage of data. During a failure, IoT gateways can transfer their applications and device connections to one another.

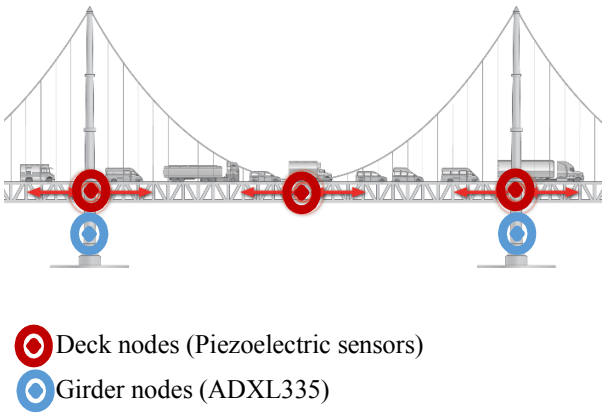


Fig. 4. Distribution of sensor nodes

E. Self Powered Sensor Nodes

IoT applications faces a puzzle to solve the issue of limited energy resources. A solution to this is adding energy harvesters. IoT and energy harvesting is a perfect match for monitoring applications. As energy harvester Solar panel is the best choice as solar power is certainly the cleanest form of renewable energy available. The solar panel must provide sufficient current to power the Arduino and also charge the battery pack during the day. So only connecting it to a charging circuit won't do any good as eventually the charge won't be sufficient for the entire day if the sensors consumes very high power. In most cases of SHM using WSN, the Arduino read the sensors periodically and transmit their values then it usually enters sleep mode until it reaches the next sampling time. During this idle time, instead of putting the Arduino into sleep, which won't save much power as other peripherals (sensors and communication modules) keep working, we can employ a timer circuit, which turns the power on/off periodically. When the Arduino is ON the sensors will read and transceiver will transmit the data. After that the timer circuit turn the Arduino OFF for few minutes and then restart the entire process again.

A 555 timer circuit operating in the astable mode can be used for this operation. In astable mode, the 555 timers output voltage alternates between Vcc (+5V) (high state) and GND (0V) (low state). The frequency of the wave can be adjusted by changing the values of 2 resistors and a capacitor connected to the chip. The timer circuit is shown in the Fig. 5. It consists of a 555 timer, resistors, capacitors and a diode. By changing values for R1, R2 and C1, the frequency can be altered (the length of ON time and OFF time. Increasing C1 or R2 will increase both the ON time and OFF time which Increasing R1 will increase the ON time without affecting the OFF time.

Mathematically, this is represented as,

$$Ton = 0.69 \times C1 \times (R1+R2) \quad (4)$$

$$Toff = 0.69 \times C1 \times R2 \quad (5)$$

Where,

Ton is the time (seconds) during which o/p pulse is high

Toff is the time (seconds) during which o/p pulse is low

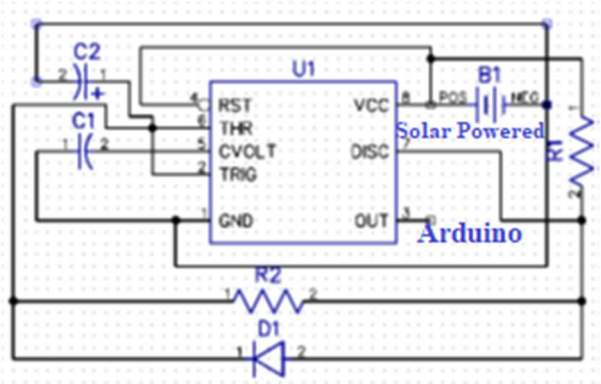


Fig. 5. Timer Circuit Schematics

Minimum duty cycle for this kind of circuit is 50% (Half time ON and half time OFF). To increase the OFF time beyond the ON time, a simple logical inverter method at the output of the timer is employed. This way the values of R1, R2 and C1 is chosen (2 MΩ, 1 MΩ and 47 µF respectively) so that the duty cycle is close to 75% (before the logical inverter). After the inverter output it is ON only for 25% of the time (32.57 seconds ON and 97.71 seconds OFF, resulting duty cycle is 25%). The inversion is performed by using a diode parallel to R2.This extends the battery charge to power the nodes up to few consecutive days which is an advantage really needed during non-sunny seasons.

IV. Experimental Results

For initial testing, ThingSpeak was used as our database. ThingSpeak is an open source Internet of Things (IoT) application and API to store and retrieve data from things using the HTTP protocol over the internet or via a Local Area Network [16]. Data is fetched from ThingSpeak and enrolled on the logger section of our designed webpage in which the bridge state is monitored and controlled. The control strategy is divided into 2 categories. One is control via spike barriers and the other is through arm barrier. The entire bridge is divided into regions of 0.5km, each region separated by a spike barrier (inactivated). The data collected are compared against a threshold value in the NodeMCU gateway serving the corresponding region and if it reaches the threshold point, the NodeMCU automatically sends a command via the LoRaWAN network to the control node to activate a spike barrier to stop vehicles approaching the overloaded region. A warning is given to the drivers on board prior to activating the spike barriers. Tilt of the girder is also monitored and when it reaches the threshold the entrance of the bridge is closed by an arm barrier in the same way. The authority can also close the entrance through the internet only by a click on the webpage provided suspicious behavior in the bridge. As this is done through the internet, the authority can take immediate step from any location. An image of our logger for deck monitoring is provided in fig. 6. This displays the data transmitted by the sensor node 100.

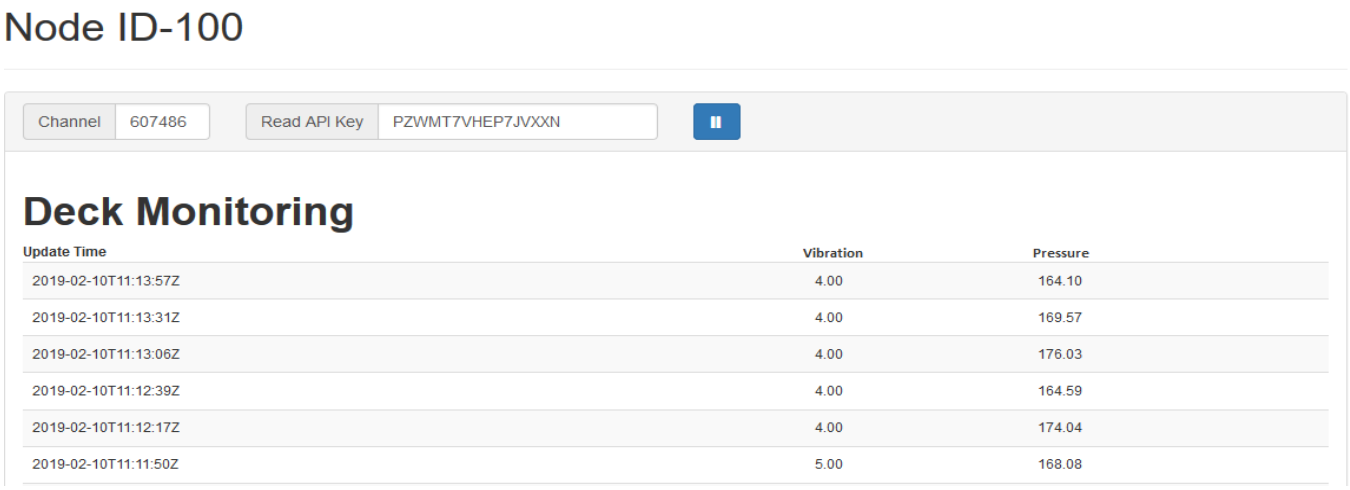


Fig. 6. Deck Monitoring logger for NODE 100

V. Conclusion

Every year breakdown of bridges cause a huge loss of lives and properties. Thus preservation of the strength of a bridge has become a paramount mission. Our proposed system not only decreases the stake of lives but also decreases the cost of maintenance of the Bridge by taking preventive moves. Undoubtedly, the maintenance of a bridge before it collapses is a lot lesser than that of a bridge that has already collapsed .

References

[1] Balageas, B., & Fritzen, P. (2006). C., A. Gemes. Structural Health Monitoring, Wiley-IST

[2] R. De Reus and P. R. Scheeper, "Fabrication and characterization of a piezoelectric accelerometer," Journal of Micromechanics and Microengineering, vol. 9, no. 2, p. 123, 1999.

[3] K.-t. Lau, C.-c. Chan, L.-m. Zhou, and W. Jin, "Strain monitoring in composite-strengthened concrete structures using optical fibre sensors," Composites part B: engineering, vol. 32, no. 1, pp. 33-45, 2001

[4] .5. Liyan Zhang, Guanchen Zhou, Yang Han; Honglei Lin, Yuying Wu; Application of Internet of Things Technology and Convolutional Neural Network Model in Bridge Crack Detection, IEEE Access

[5] https://internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT

[6] https://www.analog.com/en/products/sensors-mems/accelerometers/adxl335.html#product-overview

[7] Yang Guo, Seong Eyl Lee & James B. Mann, Journal of Manufacturing Science and Engineering, Volume 139

[8] https://www.ukessays.com/essays/computer-science/microcontrollers-used-in-wireless-sensor-networks-computer-science-essay.php

[9] A. Saifullah, M. Rahman, D. Ismail, C. Lu, R. Chandra, J Liu, Snow: sensor network over white spaces

Proceedings of the 14th ACM conference on Embeded Network Sensor Systems CD-ROM, SenSys'16,ACM,New York, USA(2016)

[10] L. Alliance, LoRaWAN communication protocol, (https//:www.lora-alliance.org/).

[11] Z. Liu and Y. Kleiner, "State-of-the-art review of technologies for pipe structural health monitoring," IEEE Sensors Journal, vol. 12, no. 6, pp. 1987-1992, 2012.

[12] S. D. Glaser, H. Li, M. L. Wang, J. Ou, and J. Lynch, "Sensor technology innovation for the advancement of structural health monitoring: a strategic program of US-China research for the next decade," Smart Structures and Systems, vol. 3, no. 2, pp. 221-244, 2007.

[13] A. N. Chester, S. Martellucci, and A. V. Scheggi, Optical fiber sensors. Springer Science & Business Media, 2012.

[14] J. Dakin and B. Culshaw, "Optical fiber sensors: Principles and components. Volume 1," Boston, MA, Artech House, 1988, 343 p. For individual items see A90-36770 to A90-36775., vol. 1, 1988.

[15] https://www.sciencedirect.com/topics/engineering/optical-fibers

[16] http://en.m.wikipedia.org/wiki/ThingSpeak

[17] http://www.electronicwings.com/sensors-modules/adxl335-accelerometer-module

IoT Based Automated Solar Irrigation System Using MQTT Protocol In Charandeep Chakaria

Mohammad Minhajul Islam
Dept. of Electrical and Electronic Engineering
University of Science and Technology Chittagong
Foy's Lake, Chittagong, Bangladesh
minhajahmedustc318@gmail.com

Md. Shahadat Hossain
Dept. of Electrical and Electronic Engineering
University of Science and Technology Chittagong
Foy's Lake, Chittagong, Bangladesh
Shahadat_iut@yahoo.com

Rezawl Karim Reza
Dept. of Electrical and Electronic Engineering
University of Science and Technology Chittagong
Foy's Lake, Chittagong, Bangladesh
rezawlkarimreza49@gmail.com

Anik Nath
Dept. of Electrical and Electronic Engineering
University of Science and Technology Chittagong
Foy's Lake, Chittagong, Bangladesh
anik.nathnu@gmail.com

***Abstract*— Agriculture represents a significant index in the evaluation of rural areas of Bangladesh. In order to increase the productivity of crops, a smart irrigation system is needed. Using GSM and IoT technologies, the proposed system makes the agricultural cultivation process more automated. The IoT technology makes the irrigation system capable of monitoring the field and taking the decision to irrigate automatically. The power generated from the solar PV system is the main power source of the proposed system. The sensor is installed on the agricultural field which is frequently sensing the water level. Without visiting the agricultural field, farmers can get information about the water level. When low moisture is detected, then motor is turned on and when moisture is normal then motor is turned off. After that, it sends data to GSM module sim900A which gets information about the motor in the farmer's cellular phone. To transmit and receive the sensor information, Message Queuing Telemetry Transport (MQTT) system is used in the project. By detecting the status of Soil Moisture Content, GSM module controls the water pump status and displays the condition of Soil Moisture Content through a server or mobile application.**

Keywords— Smart irrigation, GSM module SIM 900A, The Internet of Things, MQTT, A Solar photovoltaic system, The Moisture sensor

I. INTRODUCTION

Agriculture is the largest employment sector in Bangladesh. Most of the Bangladeshi people earn their living from agriculture [1]. Farming stands the basic factor of Bangladesh economy, contributing 19.6% to the national GDP and giving employment for 60% of the total population [2].Farmers rely on the monsoon because of water needed, which is insufficient. So the proposed system is concerned, which automatically irrigate the plants depends on the soil moisture condition and atmosphere type [3]. Reducing the waste of water is an important concern. The IoT based automated irrigation system increases the efficiency of cultivation. The moisture sensor is used which observes the plants condition & level of water from distance. To reduce the waste of water and monitor the irrigation system, IoT gets the best results using innovative techniques. MQTT protocol is used in the system. So the proposed GSM and IOT based irrigation design is energy saving and also cost effective [4]. In this paper the system is designed for a remote area like Charandeep, Chakaria, Cox's Bazar. Most of the people in this area are dependent on agriculture, but this area has no grid connection. Petrol, diesel, kerosene etc are used as the fuel source of plant irrigation. This type of sources is not environment-friendly, because of the carbon dioxide emission. The proposed MQTT based smart system presents an idea to use the sunlight as the power source which can fulfill the demands of irrigation in this area. The source is environment friendly and also cost effective.

II. LITERATURE REVIEW

Reduction of waste of water and consumption of nonrenewable energy sources are confirmed by using multiple craftsmanship and technology [5]. If its' energy sources are low cost, then the system is useful to the area (Charandeep) and using GSM Technology for observing and controlling the cultivation method except farmer visiting [6]. Using renewable energy sources like sunlight, the wind has been utilized power consumption [7]. Agricultural sector is improved rapidly using smart technologies [8]. Different types of IoT equipment are available in the market which makes smart irrigation process more automated [9, 15]. The rural area like Charandeep, 80% of the farmers still improvised. So the high cost of equipment buying is not economically viable [10]. Automatic irrigation techniques irrigate using a wireless sensor network, i.e. Zigbee and internet technology [11].To improve the productivity of the crop, there is an immediate change of the manual method from automatic method [12]. The accurate observing of the field data and controlling the cultivation process is being automated through IoT and GSM technologies by taking real time data from the field [13, 17]. The literature review represents that difference between the previous, present and proposed system. The previous state represents that manual based ordinary pump is used where the power sources are diesel, petrol and kerosene. So the system is not suitable for the environment and also costly. The present state has used the power of a source of conventional and renewable energy which energy is cost effective and energy efficient. The proposed system is based on solar driven irrigation system to accumulate GSM technology, MQTT and solar power. So this system is not only energy efficient by using solar power

but also serving the fastest communication on agriculture field. The proposed system also developed an agriculture farm which has different types of crops such as tomato, chilies, brinjal and paddy etc. So, different types of crops need different water level. The analogue pin of Arduino and the moisture sensor probes detect different types of crop water level, which turned on/off the motor. Using sensors in a field area with four sites of corners which supply the water in the field is balanced.

III. PROPOSED SYSTEM

In Fig: 1 represents the block diagram of proposed system. Photovoltaic (PV) cell converts sunlight into electricity directly. Solar power is the main source to drive the whole system. This power will be stored in a dc battery which is used as the backup when solar irradiance is not appropriate for the system. The moisture sensor, Arduino, GSM sim900A and cell phone interface are driven by dc power. The moisture sensor senses the level of water. When sensor senses the low/high moisture, the motor turns on/off automatically. This condition of the field will notify farmer's cellular phone through GSM sim900A. A Jason format links the motor status (On/Off) automatically to the server via MQTT broker. The server can view the motor condition of the field. The MQTT protocol system is observed by the computer system using internet. However DC pump is driven by the dc power through the motor driver circuit. The user can see the motor turn on/off condition and field moisture dry/wet condition in smooth TFT display.

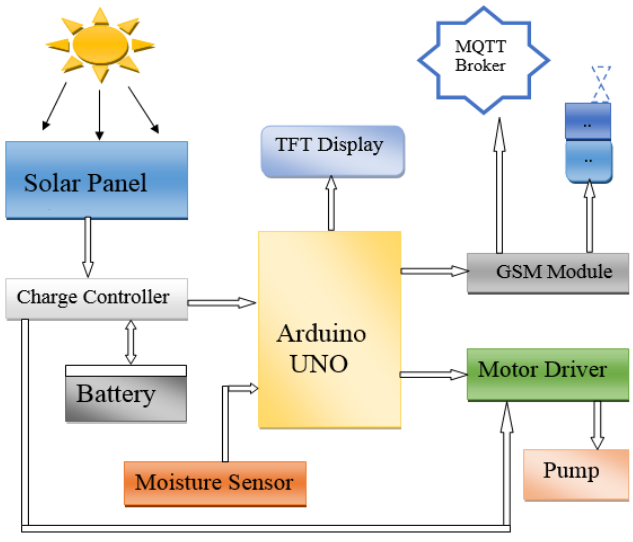


Fig.1. Block Diagram of Proposed Model.

The solar irradiance and clearness index play a vital rule for solar energy conversion. The proposed system designed for Charandeep, Chakaria, Cox's Bazar is encouraged by following NASA solar irradiance and clearness index Fig: 2 and fig. 3.

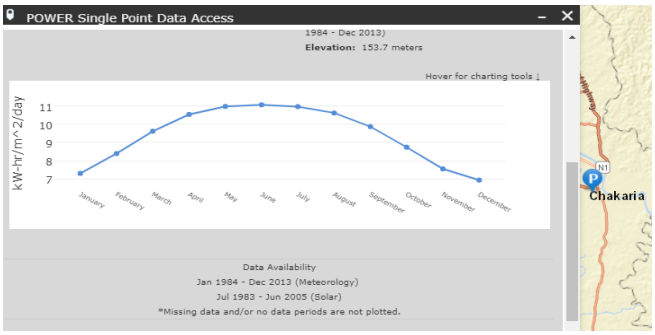


Fig.2. NASA solar irradiance for Chakaria zone [14]

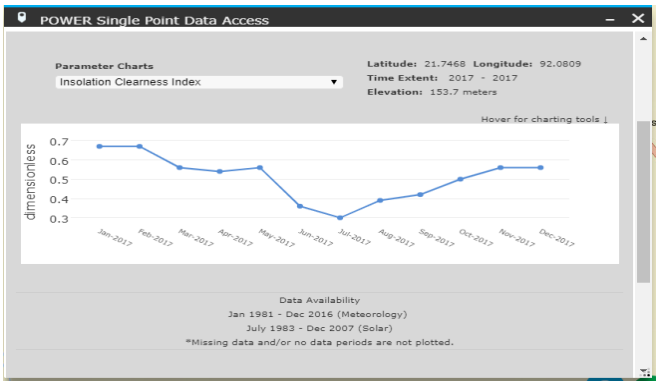


Fig.3. NASA clearness index for Chakaria zone [14].

So this nonconventional source of energy is appropriate for this rural area.

IV. HARDWARE COMPONENTS

Solar Panel:

The solar energy comes from resources which are continuously replenished such as sunlight. Solar cells operate by converting sunlight directly into electricity. This energy will improve the quality of our environment. It's the main source of power in the proposed system.

TABLE I. Specification of solar panel.

Name	Specification
Power	20W
Voltage	17.46V
Current	1.15A
V_{oc}	21.16V
I_{sc}	1.28
Operating Temperature	-40^0C to $+85^0$C

GSM MODULE SIM900A :

GSM uses variation of time division multiple access (TDMA).GSM module sim900A controls the overall system. This module need minimum (1.5A =2A) power source.

TABLE II. Specification of GSM Module SIM 900A.

Name:	Specification
GSM Module	SIM 900A
Frequency	900MHz/1800MHz
VCC5	5V
VCC4	3.3V - 4.5V
Rated Current	1.5A < 2A
Size	(5.6 cm x 4.9 cm x 2.8 cm)
Weight	23g

MQTT PROTOCOL

It is a simple messaging protocol and a perfect solution for internet of things application. It allows to sends command to monitor and controls the output and communicates among multiple devices.

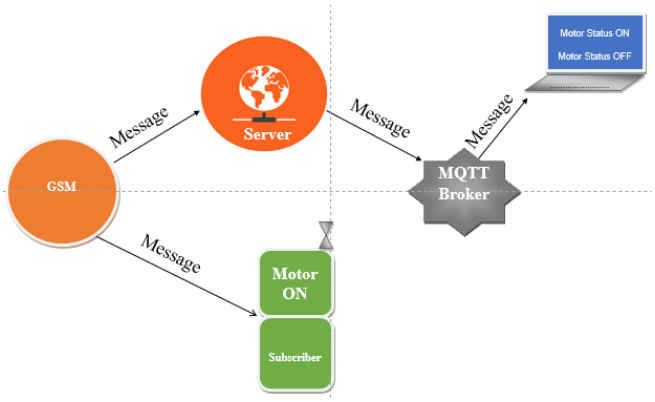


Fig .4. MQTT protocol system.

Message Queuing Telemetry Transport (MQTT) VS Hyper Text Transfer Protocol (HTTP).

- MQTT is data centric whereas HTTP is document centric.
- HTTP is request-response protocol for client-server computing and not always optimized for mobile devices whereas MQTT provides publish/subscribe model clients with independent existence from one another and enhance the reliability of the whole system.
- According to measurements in 3G networks, throughput of MQTT is 93 times faster than HTTP's.
- MQTT uses low bandwidth whereas HTTP uses comparatively high bandwidth.

So MQTT is better than HTTP for IoT development [16].
MOISTURE SENSOR:

The moisture sensors identify the water level of agricultural field. When the water level of farming field is become dry/wet then the moisture sensor's digital output is low/high.

TABLE III. Specification of soil moisture sensor.

Name	Specification
Power indicator	Red color Led
Digital Switch indicator	Green color led
Operating voltage	3.3-5V
Digital output interface (0 and 1)	DO
Analog output interface	AO

V. DESCRIPTION OF PROPOSSED MODEL

A. Flow Diagram of proposed model :

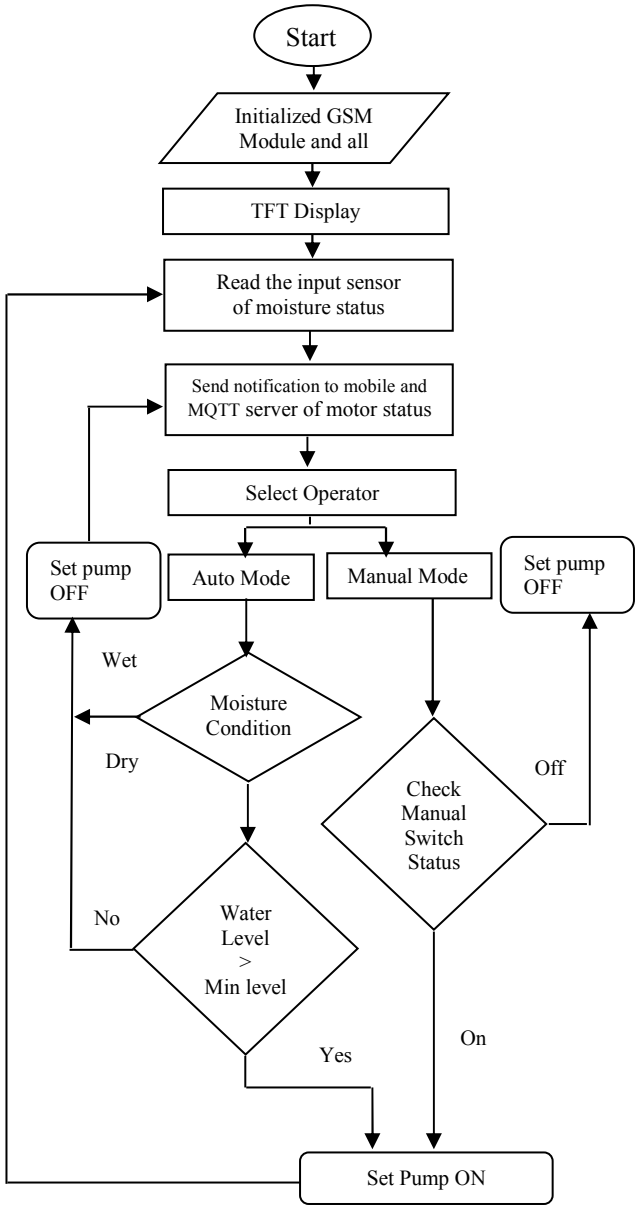


Fig.5. Flow Chart of Proposed Model.

Description of flow chart proposed system:

The following steps involved are working on a system are mentioned below:

Step 1: Power is supplied to the proposed system which initializes GSM and all.
Step 2: After initializing the system, the TFT display shows "MOISTURE BASED PROJECT".
Step 3: The input module moisture sensor senses the moisture condition of farming field and transmits a digital data to the Arduino.
Step 4: Arduino checks the mode operation. If the mode is manual then it cannot operate automatically. Then it depends on farmer's feedback.
Step 5: When the mode is auto then it will operate from step 6 to step 12.
Step 6: The moisture sensor senses two types of moisture condition. I,e (wet/ dry).
Step 7: If the moisture is wet, the moisture sensor sends a digital data to the arduino.
Step 8: Arduino receives this digital data and checks the logic. If the logic is low moisture then the motor is "ON" through the motor driver circuit and also transmits a data to the GSM SIM 900A.
Step 9: The GSM SIM 900A receives this data and sends a notification to the farmer's cell phone " Low soil moisture detected. Motor turned ON" and at the same time sends a notification to the server through MQTT protocol.
Step 10: If the moisture condition is dry then the module transmit a digital data to the Arduino and check this data.
Step 11: If the condition "Water level > Min level" is true, then the pump is turned off through motor driver circuit and also sends a signal to the GSM module SIM 900A.
Step 12: The GSM sends a notification to the farmer's cell phone "Soil moisture is normal. Motor turned OFF" and at a time sends a notification to the server through MQTT protocol.

B.CIRCUIT EXPLANATION

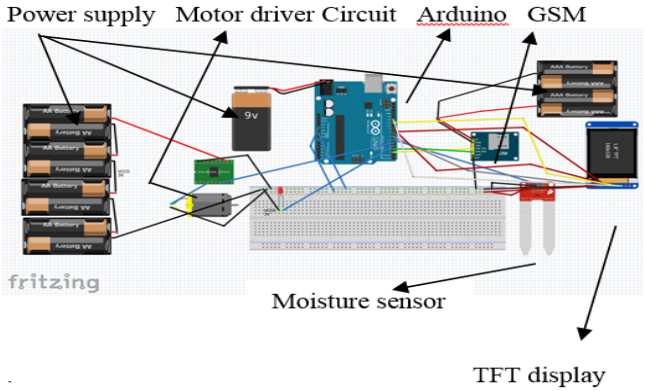


Fig.6. Circuit Diagram of Proposed Model

Power supply:

The power is supplied from the solar power. The solar power voltage is not a uniform voltage. Solar charge controller is used for supplying uniform voltage to the load and control system. To store the power, a 12V DC battery is used. The battery supplies the power when solar power does not work at adverse condition.
Charge Controller Connection:

Solar voltage output is connected to the charge controller input terminal. Battery is connected to the storage terminal and motor system is connected to the load number 1 or load number 2.

Motor connection:

The output of the Charge controller is 12V which is connected to the motor driver input 12V. The motor driver output 12V is connected to the pump and ground terminal is connected to the ground. The pump is controlled by the Arduino digital output through the motor driver circuit.

Moisture sensor connection:

- The moisture sensor Vcc pin connects to the 5V power supply.
- GND pin is connected to the ground.
- Digital input, output pin (D_0) is connected to the Arduino digital pin D_6.

GSM Connection:

- GSM module 900A is power supply takes from 5V, 2A circuit for a dedicated power supply.
- GSM SIM module 900A TX pin is connected to the Arduino digital pin D_2.
- RX pin is connected to the Arduino digital pin D_3.
- GND pin is connected to the ground.

Display connection:

- The 1.8 TFT display module RST pin is connected to the Arduino digital pin D_7 through the 1K resistor.
- CS pin is connected to the Arduino digital pin D_9 through the 1K resistor.
- D/C pin is connected to the Arduino digital pin D_8 through the 1k resistor.
- DIN pin is connected to the Arduino digital pin D_{11} through the 1k resistor.
- CLK pin is connected to the Arduino digital pin D_{13} through the 1k resistor.
- Vcc/BL is connected to the 5V supply.
- GND pin is connected to the ground.

EXPERIMETAL SETUP OF PROPOSED SYSTEM

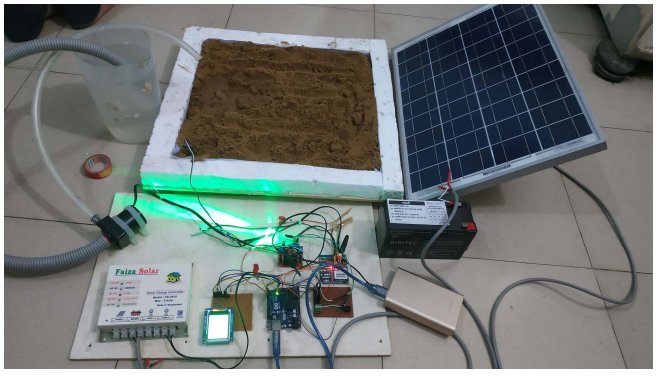


Fig.7. Experimental setup of proposed system

VI. RESULT PART

The screenshot of the moisture condition in the TFT display as bellows:

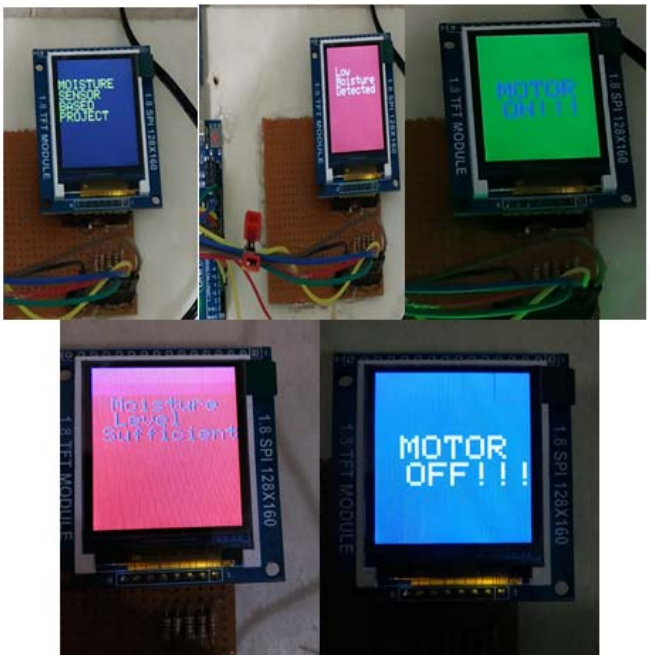


Fig.8. Pump status in TFT display.

The following fig: 8 and fig: 9 of pump status is sent to the farmer's mobile phone via GSM technology. When the moisture sensors detect the low moisture, then it sends a notification to the farmer's mobile phone as "MOTOR ON". Whenever the moisture sensor detects the moisture is normal, then send the notification "MOTOR OFF".

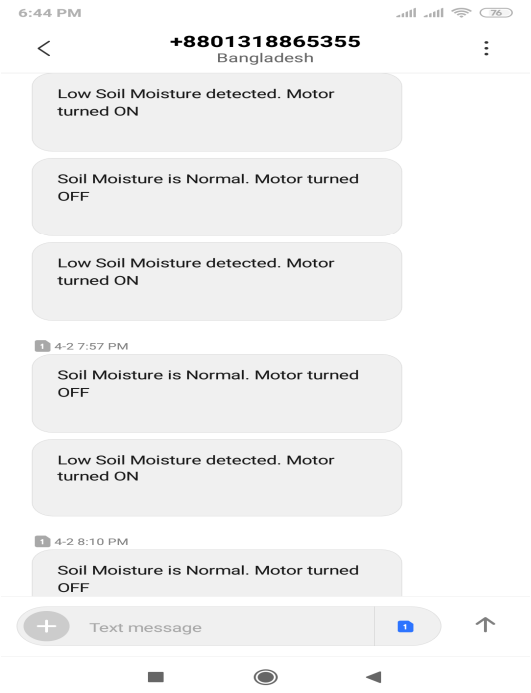


Fig.9. Pump status in farmer mobile phone

MQTT broker system carries the JASON format link to the subscriber's web server of Motor status to monitor and control the system as Following fig: 10 and fig: 11.

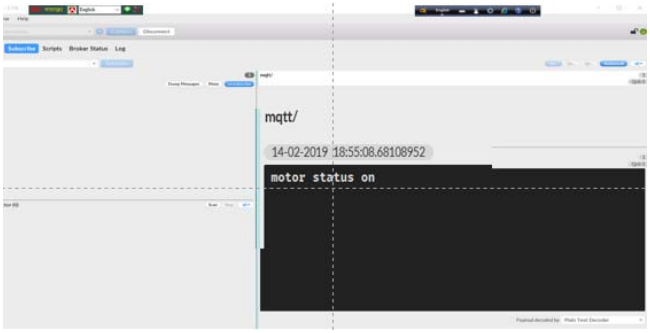


Fig.10. Pump "ON" status monitoring in the server.

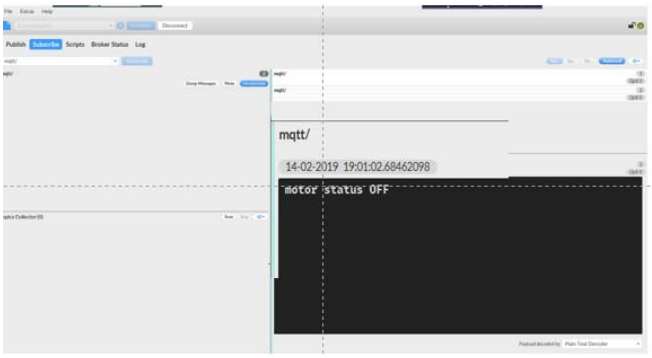


Fig.11. Pump OFF status monitoring in the server.s

VII. CONCLUSION

To increase the productivity of crops, farmers need automatic and IoT technology in the irrigation system which provides more comfort to farmers. To require less human intervention and minimum maintenance for agricultural fields, the system is feasible to use conveniently for all farmers. Moreover time is saved and water is used accurately without wasting. MQTT is used which communicates with multiple devices. It's lower bandwidth and low power consumption makes the proposed system cost effective. This project can be a suitable alternative to overcome the necessity of uninterrupted supply of electricity in irrigation system of our farmers. The system can be implemented in rural areas (Charandeep, Chakaria) with a lower investment.

VIII. REFERENCES

[1] https://en.m.wikipedia.org/wiki/Agriculture.[Accessed 11-Nov-2018].

[2] https://www.nationsencyclopedia,com/Agriculture_in_ Bangladesh [Accessed 11-Nov-2018]

[3] J. guiterrez,J, francisco,v.Medina & M. A. P. Gandara,"Automated irrigation system using wireless sensor network & GPRS module", IEEE 2013

[4] Renu singh, Ahmed sharique anees,Devang tyagi,"Solar irrigation system using brush less DC Motor",IJIREEICE,VOL.3,ISSUE 6,JUNE 2015.

[5] Dr, M, Mahalakshmi,Dr. S.p. Rajaram,s. priyanka, R. rajapriya,"Distant monitoring & controlling of solar driven irrigation

system through IoT",2018 National power Engineering conference(NPEC).

[6] Archana P, Priya R,"Design & implementation of automatic plant watering system",IJAEGT,VOL 04,ISSUE 01,JANUARY 2016

[7] N. A. Fawzi, S. A. jalal,"Design & implementation of smart irrigation using wireless sensor network based on IoT", International journal of scientific & engineering research,VOL 8, ISSUE 4,APRIL 2017.

[8] Pushkar singh, sanghamitra saikia," Arduino based smart irrigation using water flow sensor,soil moisture sensor, tempreture sensor & EsP8266 wifi module",2016 IEEE Conference.

[9] K. Lekjaroen, R. Ponganantayotin, A. Charoenrat, S. Funilkul,U. Supasitthimethee and T. Triyason, "IoT Planting: Watering system using mobile application for the elderly," 2016International Computer Science and Engineering Conference(ICSEC), Chiang Mai, Thailand, 2016, pp. 1-6.

[10] Ravi Kishore Kodali and Borade Samar Sarjerao,"A low cost smart irrigation system using mqtt protocol",2017 IEEE REGION 10Symposium (TENSYMP).

[11] MS. Sweta s. patil,Prof. MS. A.V.malvijay,"Review for ARM based agriculture field monitoring system",IJSRP ,FEB 2,2014.

[12] Nikesh Gondchawr,Dr. R.S. kawitkar,"Iot based smart agriculture",IJARCCE,VOL 5,ISSUE 6,JUNE 2016

[13] R. Nageswara Rao,B. Sridhar," IoT based smart crop-field monitoring and automation irrigation system", 2018 2nd International Conference on Inventive Systems and Control (ICISC)

[14] https://power.larc.nasa.gov/data-access-viewer/ [Accessed:12-Dec-2018].

[15] Tanu Sahu, Ashok Verma, "Automated Smart Irrigation System Using Rashberry Pi" international journal of computer applications(0975-8887),volume 172-No.6,Agust 2017.

[16] https://medium.com/mqtt-buddy/mqtt-vs-http--which-one-is-the-best-for-iot-c868169b3105.[Accessed:20-Jan-2019].

[17] Abhilas shrivastava, M. Rajesh, " Automatic Irrigation system with data log Creation", 2018 Second international conference on inventive computational technologies(ICICCT),27 september 2018

Towards IoT and ML Driven Cardiac Status Prediction System

Md. Ittahad Uz Zaman
Dept. of CSE
United International University
Dhaka-1212, Bangladesh
mzaman152134@bscse.uiu.ac.bd

Shaira Tabassum
Dept. of CSE
United International University
Dhaka-1212, Bangladesh
stabassum152129@bscse.uiu.ac.bd

Md. Sabbir Ullah
Dept. of CSE
United International University
Dhaka-1212, Bangladesh
mullah152144@bscse.uiu.ac.bd

Ashikur Rahaman
Dept. of CSE
United International University
Dhaka-1212, Bangladesh
arahaman152135@bscse.uiu.ac.bd

Samsun Nahar
Dept. of CSE
United International University
Dhaka-1212, Bangladesh
snahar152117@bscse.uiu.ac.bd

A.K.M. Muzahidul Islam
Dept. of CSE
United International University
Dhaka-1212, Bangladesh
muzahid@cse.uiu.ac.bd

***Abstract*—The amount of death rate in Bangladesh because of different types of non-communicable cardiac diseases are rapidly increasing every year. These non-communicable diseases are long-term and slowly progressive and become so severe that the condition of a patient is too critical to handle. As a result, most of the people getting a sudden heart attack or learn about their disease when it is too late. Thus, in this paper, we propose an IoT and Machine Learning derived cardiac status prediction method that can collect necessary data from human body using IoT devices (sensors) and transfer those data to the cloud where the data are stored with user authentication. The data which are received from the human body are then to be normalized before machine learning algorithms are applied to them to calculate and predict the overall condition of a patients heart. To-date we could collect heart rate, ECG signal, and cholesterol through IoT devices from the human body. The results we have received of heart rate and cholesterol are quite satisfactory. There are still some issues with the interpretation of the ECG signal which we are working on.**

***Index Terms*—IoT, Machine Learning (ML), Cardiovascular Disease (CVD), ECG, Cholesterol, Heart Rate.**

I. Introduction

Cardiovascular Disease (CVD) and Coronary Artery Disease are two increasing cause of mortality nowadays in Bangladesh. Over the last few decades, the main reason for death has dramatically shifted from communicable disease e.g.Cholera, Smallpox to non-communicable e.g. Cardiovascular disease, Cancer, Blood pressure, Diabetes disease. Previously from a statistics of 1986, the morbidity and mortality for the communicable disease were 52% and for the non-communicable disease, it was 8% [1]. But according to some studies done in 2014, it has been found that the main cause of death has dramatically shifted to non-communicable disease from communicable disease, which is quite severe in the sense because non-communicable disease like Cardiovascular disease is the most fatal reason of death bearing the statistics of 17% alone [1]. The effect is higher in urban areas although the available treatment for the cardiac disease is quite costly. So looking through different aspects of CVDs in our country it is safe to say that, this side needs some attention and some measures have to be taken to monitor the cardiac condition as well as to get useful suggestions or alerts for upcoming anomalies of a person's cardiac condition.

Thus, in this paper, we describe a prototype which will automate the whole task of data collection and transferring the data with the help of IoT devices to an application running a Machine Learning algorithm on the cloud, which will then calculate the result based on the inputs collected from different IoT devices and sensors connected to the patient or user-end in real time. The ML algorithm is provided with 12 parameters to get a prediction about the possible cardiac anomalies.

Cardiac diseases are highly prevalent in Bangladesh. The death rate is increasing rapidly due to the rising amount of heart attacks, Coronary Artery Disease, Cardiovascular Disease (CVD). According to the Global Health Observatory data of the World Health Organization in 2017 Bangladesh is ranked 104 in the world with 14.31% of total deaths.

According to the recent Non-communicable Diseases Country Profiles of Bangladesh provided by WHO, the number of people who are taking drug therapy to prevent heart attacks and strokes is negligible. People are barely conscious of non-communicable diseases but these long-term and slowly progressive diseases are so severe in future that condition of the patient becomes too critical to handle. According to Bangladesh criteria, people hardly follow a regular checkup or feel the necessity to counsel a doctor even once a month. As a result, most of the people got a sudden heart attack or get to know about their disease when it is too late. So all of these situations motivated us to come up with IoT devices based on ML driven heart condition monitoring system. A device that will take all the necessary sensitive and private data from human body such as Cholesterol, Heart rate, ECG, Blood sugar and Blood pressure and so forth with the help of IoT devices and our Machine Learning algorithm will calculate the result

and provide the overall heart condition of the patient.

Blood pressure monitor, glucose meter devices are popular and useful among people as these devices measure blood pressure, blood sugar level which let people know their current health status within a minute. This service is not available yet for non-communicable disease patients. Hence, a device that can let people learn about their current heart condition will be helpful to take the necessary preventive steps at the right time.

The rest of the article is organized as follows. Section II provides a review on existed works related to IoT based health monitoring systems. Section III presents our proposed methodology on cardiac diseases detection and Section IV demonstrates our experimental analysis and results. Finally, Section V is the conclusion including our future progress and improvement.

II. Related Work

In this section, we have reviewed the existing works related to cardiac issues that already have been done or people still are working on.

A. Blood Pressure and Heart Rate

There is a paper by Snehal Sanjay Kale et al. in [2] where they introduced BSN technology, which is consisted of low-power and lightweight wireless sensor nodes, such as temperature sensor, heart rate sensor to collect blood pressure and heart rate of a patient. These devices are placed on the human body to monitor the health condition of a patient at home and all of these collected health-related parameters will be sent to their physician's server using a long-range wireless technology.

We-Care [3] is an IoT based health care system which introduced We-watch wristband using a SensorTag from Texas Instruments. This wristband is for elderly people which collects blood pressure, humidity from the human body. It can also detect a sudden fall of the patient. The Fall Detection system is implemented by reading the accelerometer data which is used to detect sudden movements, like falls, and also to track any movement activity performed by the elderly person.

A paper by P.Gupta [4] presented an IoT based system which provides emergency medical service by collecting data information through health status monitors including heart rate, blood pressure, and ECG. These pieces of information are sent to the patient's doctor as an emergency alert with full medication information.

In paper [5], Johan Bhurney et al. used the photoplethysmography method via the PPG sensor to detect the variations in blood volume and blood flow in the body.

B. Electrocardiography

In paper [6] the authors have proposed a health monitoring and measurement system by inexpensive wearable sensors. The sensor will collect some physiological parameters and will transmit the data to a gateway server via Bluetooth. The physiological biomarkers that are taken into consideration are ECG, Respiratory rate, Skin temperature, and EMG muscle movement.

Paper [7] mainly focuses on electrocardiogram of a patient through IoT enabled sensors. Despite conventional clinical ECG equipment has 12-lead system, this proposed system implements 3-lead placement in a triangular fashion around the heart. The data is processed through an MCU and sent to the cloud with wireless access points.

In work [8] the researchers mainly worked on developing a software infrastructure that supports ECG signal analysis for feature extraction and classification for diagnosing the heart condition. The features space extracted for the ECG signal led to an accuracy of 98.7%. They also made it clear that ECG analysis and classification can be done in real time.

Paper [9] describes the way of measuring physiological ECG data using a single board computer with the help of e-health sensor platform. The ECG data is processed to get the Heart Rate(HR) and Heart Rate Variability (HRV) which later on will be used to diagnose major diseases.

C. Glucose and Cholesterol

In the work by Meghana Chandrashekar et al. a Near-infrared sensor is used to measure the glucose level of a person which can measure the glucose level without the blood samples from the human body. Another used sensor is BMI sensor which can measure the cholesterol level of a person. In this process, a signal is passed through the human body and the sensor waits for the receiving time of the signal on the other side and then the cholesterol level is calculated based on that time. There are two parts - software and hardware. The software application will monitor a persons health by measuring the blood glucose level and cholesterol level of the human body. The hardware kit is made of Infrared sensor and BMI sensor which helps to measure glucose and cholesterol level to monitor a persons health.

In the work of Mr. Li et al. [10] consider some parameters as the vital element of heart condition prediction e.g. (Blood pressure, ECG, SpO2, Heart rate, Pulse Rate, Blood glucose, Blood fat). For intermediate communication Bluetooth and long distance communication Cellular network, Wireless network is used.

In the paper of Mr. Gomez et al. [11] described the statistics of the mortality of patients suffering from different chronic disease. It appears to be that, most of the patients dying from heart failure have been suffering from Overweight 2.6M, Elevated Cholesterol 4.4M, Blood Pressure 7.1M. Therefore, detecting these vital of the patients on a regular basis and sending them for monitoring to a specialist was their main objective. But they described that due to technological limitations these devices are not smaller quite yet. Although some devices e.g. Glucometer, Wrist blood pressure cuff, HRM including electrocardiogram is available today.

III. Methodology

In this section, we have described our proposed method including the list of parameters required in our work. An overview of the entire prototype design and our implementation so far on 3 parameters (heart rate, ECG, cholesterol) are also explained.

A. Parameters

Our main concern is the cardiac issues of patients so, we would need to collect data-set of patients related to cardiac disease. From the previous study done on the field of cardiac disease, there are actually a total of 75 numbers of parameters. Among them, researchers have worked on 14 parameters with Cleveland data-set till now and the result for the 14 parameters have proven to be accurate enough to continue the research on 14 parameters only. Furthermore, we have seen 12 parameters also provides a similar result. So, these 12 parameters mentioned below would be sufficient enough to provide us with an accurate prediction of a patient's cardiac condition. The parameters below given are associated with their respected available sensors.

Parameters through manual input:

1) **Age:** According to this paper, at age 40 the lifetime risk of developing heart disease is one out of two for men and for women it is one out of three. If we consider age as 70 the statistics suggest that it is one out of three for men and for women it is one out of four [12].
2) **Gender:** Heart disease is more seen among the middle-aged people especially it is two to five times more common in men than it is in women [13].
3) **Chest pain type:** Chest pains are the most common symptoms indicating different Cardiovascular disorders. So that's why it is chosen as one of the parameters.

Parameters through IoT devices:

4) **Resting blood pressure:** From the paper of J.Gomez [11] and C. Li et al. [10], we get the numbers from which it appears that most of the patients dying from heart failure have been suffering from Overweight 2.6M 4.4M, Blood Pressure 7.1M. Which leads us to choose RBP (Resting Blood Pressure) as one of the parameters.
5) **Cholesterol:** Elevated cholesterol is one of the major cause of death among people having heart condition said by J. Gomez et al. in [11].
6) **Fasting blood sugar:** Fasting leads to loss of weight and reduction of blood sugar whereas it is considered a level of less than 100 mg/dL is normal. Which is why it was chosen as a parameter for generating the cardiac prediction of the patient.
7) **Resting ECG:** ECG stands for Electrocardiograms. ECG is used for detecting cardiac arrhythmia [8]. Which is essential because we need to observe the condition of the heart of a patient and detect different anomalies. ECG helps us to do that. Besides we can calculate the hear BPM with ECG.
8) **Max heart rate:** Maximum heart rate is almost equivalent to heart rate although it would help us to observe the heart condition throughout the day.

Future design and implementation using IoT devices:

9) **Exercise included angina:** Angina is a type of chest pain that causes due to the lack of oxygen flow to the heart. It is a symptom of Coronary heart disease.
10) **Old peak:** ST depression induced by exercise relative to rest which is measured by the vertical distance between the iso-electric line and the patient's trace.
11) **Slope:** The slope of the peak exercise ST segment which is a predictive value of coronary disease.
12) **CA:** A number of major vessels colored by fluoroscope which is a predictive value within 0 to 3.

According to the study, these 12 parameters are adequate enough to calculate and predict the overall heart condition of a patient but due to some limitations, we are working on the first 8 parameters and leaving Exercise included angina, Old peak, Slope and CA for future work.

B. Overview of the Design

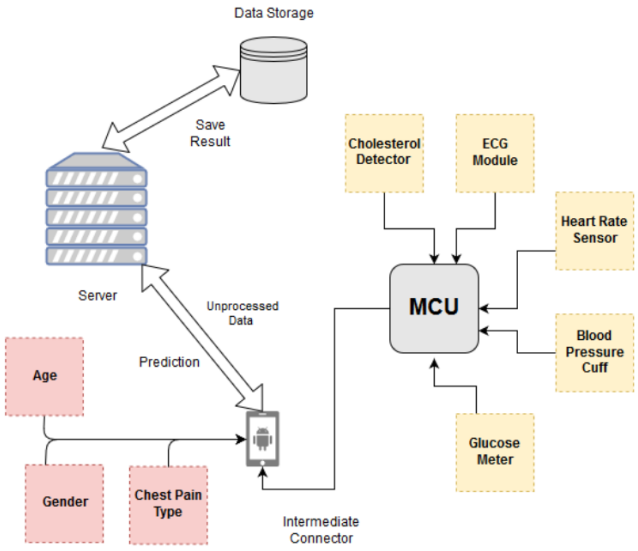


Fig. 1: Block diagram of combined prototype.

This is a simplified block diagram (Fig. 1) of the proposed prototype that we want to implement. We have proposed to use a Micro Controller Unit (MCU). The MCU will be connected to the required hardware sensors and modules. After acquiring all the data from the hardware components it will transmit the data to the smartphone via Bluetooth. The smartphone will also be used to collect the rest 3 parameters. Then all the 8 parameters will be combined and sent to the cloud running the Machine Learning algorithm which will provide the output to the user.

C. Prototype

1) Heart Rate Sensor Module

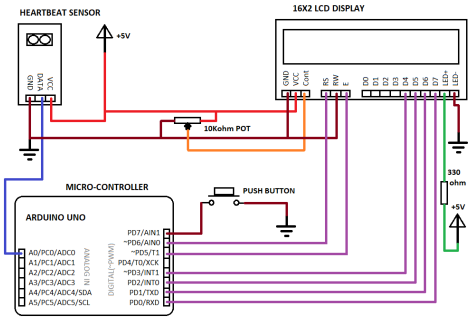


Fig. 2: Circuit diagram of heart rate sensor module.

We have implemented the heartbeat sensor to get the maximum heart rate of a patient. The circuit implementation (Fig. 2) is given which provides us the data of heartbeat of the subject. Heart Rate Module (LM358), Arduino UNO and LCD display are some of the components which have been used in this implementation.

2) ECG Module

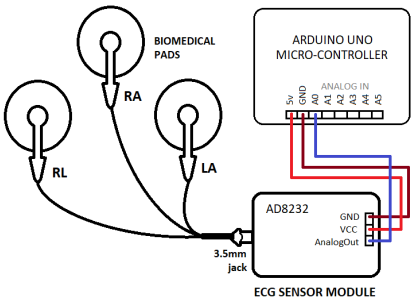


Fig. 3: Circuit diagram of ECG module (AD8232).

We have worked on patient Resting ECG where we used ECG sensor module - AD8232 to get ECG graph and corresponding data. We implemented the module with the existing Arduino UNO micro-controller as it has available peripherals which are required to implement that module (Fig. 3). AD8232 is a single lead ECG module. The sensitive pads are attached to the left hand, right hand and right leg. The output pin of the module is connected to an analog pin of the micro-controller. The output is converted by Arduino then the value is observable in the serial monitor as well as the LCD display.

3) Cholesterol

The non-invasive cholesterol detector (Fig. 4) works by transmitting infrared along the skin of the subject and measuring the received signal reflected back. This circuit contains a photodiode which responses to the infrared light source and converts the optical power into electrical current. The electrical current is then converted into a voltage by using a load resistor which varies between 0-5 V [14].

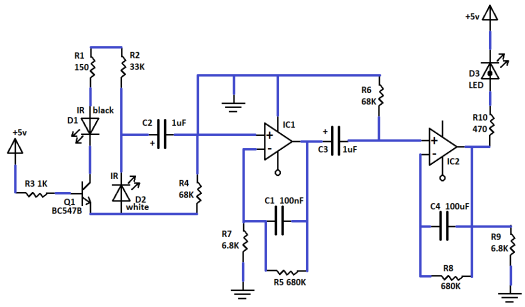


Fig. 4: Block diagram of Near-infrared sensor.

IV. Experimental Results

In this section, we have demonstrated our implemented circuits with the explanation of its working process and visualized the results we have received from different IoT devices.

A. Heart Rate Sensor Module

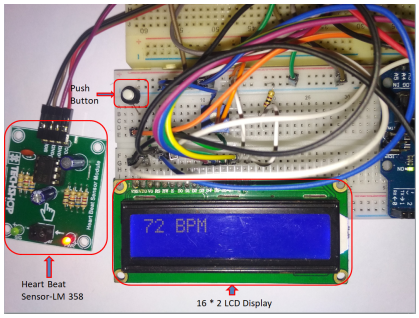


Fig. 5: Implementation of heart rate module sensor (LM358).

Here is the actual implementation of Heart Rate Sensor module (Fig. 5) which is connected to the MCU. The module measures the input for 10 seconds and provides the analog output to the Arduino which is then multiplied with 6 to get the Beats Per Minute (BPM) of the patient.

TABLE I: Sample data of heart rate sensor module

Subject No.	Beats Per Min	Physical Condition
1	66	Normal
2	72	Normal
3	62	Normal
4	90	Increased Respiratory Rate
5	110	Increased Respiratory Rate

Frequent changes in heart rate or regularity of pulse can signify the heart condition of a patient. So in our implemented module, the calculated Beats Per Minute (BPM) of a patient determines whether the physical condition of the patient is normal or the respiratory rate is high. See the table (Table I) providing some sample data of our module for more details.

B. ECG Module



Fig. 6: Output of ECG Module (AD8232).

A single lead ECG module is used which uses 3 electrodes to measure the electrocardiogram. At first, we connected the output pin to an oscilloscope where we found a signal. We have compared our received signal with real ECG signal by studying a paper [15] where we have found some similarities. Then we also visualized the output in the serial plotter where the output waveform was very promising. The main complex part of an ECG signal is the QRS complex (peak value), Q segment and S segment which we have found in our signal output. Still, there are a lot of issues with the ECG signal output on which we are working to improve and find out a much better result.

C. Cholesterol

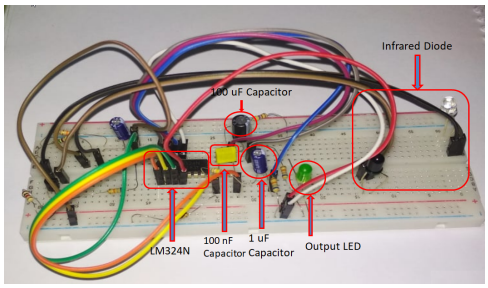


Fig. 7: Circuit diagram of near-infrared sensor.

The near-infrared sensor (Fig. 7) sends a signal through the human body and measures cholesterol without taking blood samples [14]. In the basis of each second, we are continuing the detection process for 30 seconds with 30 different samples taken on each of the seconds. Afterward, we are taking the average value of the voltage drop along the output terminal. We are taking 5.2 mm/L as the standard value of cholesterol in a healthy person.

$$\text{Correction Factor (X)} = 5.2 \,/\, \text{Voltage Drop}$$

(For standard value the Correction Factor is = 9.11)

Finally, we tuned all the output with respect to the value of (X). Which provided us the values similar to expectation. Further study and research are needed in this process.

TABLE II: Sample data of cholesterol in near-infrared sensor

Subject No.	Result	Reference	High
1	4.46		
2	4.56		
3	6.65	4-5 mm/L	>6 mm/L
4	5.10		
5	4.74		

According to the National Cholesterol Education Program (NCEP), a total cholesterol level of 4-5.17 mmol/L is considered normal and greater than 6 mmol/L is considered high. After testing our implemented module on 5 different subjects we found some data. The table (Table II) shows our findings compared to the biological reference.

D. Combined Prototype: Heart Rate and ECG

The H/W implementation (Fig. 8) we did so far is the combination of Heart Rate Sensor module and AD8232 ECG module where both modules work with a single Arduino UNO microcontroller using 9V DC battery. The required components are given below:

- Arduino UNO x 1
- 16 x 2 LCD Display x 1
- 10K Potentiometer
- 330 Resistor (Optional for LCD back light)
- Push Button
- Heartbeat Sensor Module with Probe (finger based)
- Mini Breadboard
- AD8232 ECG module
- 9V DC battery

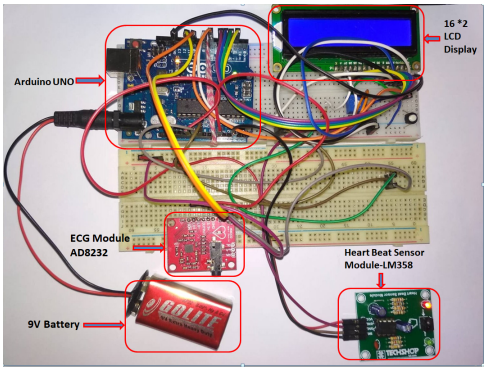


Fig. 8: Combined prototype of Heartbeat sensor module and ECG module.

E. Mobile Application

We have an application (Fig. 9) that is designed to control the complete hardware module. The application is designed for an android platform supporting from android 4.2 to upward versions.

At first, Bluetooth communication has to be made between the hardware module and application. From there we have

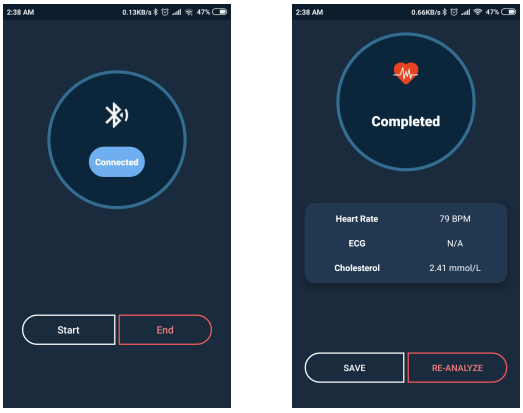


Fig. 9: Sample screen-shots from mobile application

to start the measurement procedure. When the start button is pressed, initially the heartbeat module will be activated and it will calculate the heart rate of a patient for 10 seconds. The final result will be saved for further usage. Secondly, the cholesterol module will start working. Finally, the data generated from the cholesterol module and heart rate module will be combined together and sent to the mobile application and the user will be able to see the final output. Then the user can choose to send the overall output to a cloud-based server where it will be saved in a database for further observation in a sequential way. The ECG module is yet to be connected to be operated from the mobile application.

V. Conclusion and Future Work

In this work, we have proposed a method to collect 8 parameters from the human body that are related to the heart using IoT devices to predict and prevent different severe cardiac diseases. The output of heart rate and cholesterol are quite satisfactory, however, there is still a lot of room to improve the results.

During the implementation phase, we have faced some difficulties. It was challenging to find the actual IoT devices for the implementation of our required parameters because of the lack of supplies in our country. While implementing the near-infrared sensor to collect cholesterol from the human body, receiving the input through a photodiode was tough. After putting much effort, we come up with an electrical current received by the photodiode and converted it into voltage. We tried our device on some subjects and after comparing with the real cholesterol values, we managed to find out the final results of the cholesterol of a patient. Getting the ECG signal from the human body was another challenge. The outputs that we were getting at first were not compatible with our needs. As a result, the ECG output signal was very different from the actual one. Finally, we used oscilloscope which provided us the output signal which was very similar to the actual signal. Also changing the module provided us with a satisfying result. Besides placing the electrodes correctly was quite challenging because it was not obvious which spot would provide the most accurate signal.

In future, we plan to combine all of our modules and circuits to build a complete prototype where after collecting the data, the data will be transferred to a server that runs a machine learning algorithm to predict the current cardiac status. The final results are to be visualized through a mobile application.

References

[1] A. M. Islam, A. Mohibullah, and T. Paul, Cardiovascular disease in bangladesh: A review,Bangladesh Heart Journal, vol. 31, no. 2, pp. 8099, 2016.

[2] S. S. Kale and D. Bhagwat, Highly secured iot based health care system for elderlypeople using body sensor network,International Journal of Innovative Researchin Science Engineering and Technology, vol. 5, no. 10, 2016.

[3] N. Hussain, Ieee 802.15.6 standard in wireless body area networks from a health-care point of view, 01 2016.

[4] P. Gupta, D. Agrawal, J. Chhabra, and P. Dhir, Iot based smart healthcare kit,03 2016, pp. 237242.

[5] J. B. Bathilde, Y. L. Then, R. Chameera, F. S. Tay, and D. N. A. Zaidel, Contin-uous heart rate monitoring system as an iot edge device, in2018 IEEE SensorsApplications Symposium (SAS), March 2018, pp. 16.

[6] M. Hassanalieragh, A. Page, T. Soyata, G. Sharma, M. Aktas, G. Mateos,B. Kantarci, and S. Andreescu, Health monitoring and management usinginternet-of-things (iot) sensing with cloud-based processing: Opportunities andchallenges, in2015 IEEE International Conference on Services Computing.IEEE, 2015, pp. 285292.

[7] Z. Yang, Q. Zhou, L. Lei, K. Zheng, and W. Xiang, An iot-cloud based wearableecg monitoring system for smart healthcare,Journal of medical systems, vol. 40,no. 12, p. 286, 2016.

[8] D. Azariadi, V. Tsoutsouras, S. Xydis, and D. Soudris, Ecg signal analysis andarrhythmia detection on iot wearable medical devices, inModern Circuits and Systems Technologies (MOCAST), 2016 5th International Conference on. IEEE,2016, pp. 14.

[9] O. Yakut, S. Solak, and E. D. Bolat, Measuring ecg signal using e-health sensorplatform, inInternational Conference on Chemistry, Biomedical and Environ-ment Engineering, Antalya, 2014, pp. 7175.

[10] C. Li, X. Hu, and L. Zhang, The iot-based heart disease monitoring system forpervasive healthcare service,Procedia Computer Science, vol. 112, pp. 23282334,2017.

[11] J. Gomez, B. Oviedo, and E. Zhuma, Patient monitoring system based on internetof things,Procedia Computer Science, vol. 83, pp. 9097, 2016.

[12] D. M. Lloyd-Jones, M. G. Larson, A. Beiser, and D. Levy, Lifetime risk of de-veloping coronary heart disease,The Lancet, vol. 353, no. 9147, pp. 8992, 1999.

[13] P. Jousilahti, E. Vartiainen, J. Tuomilehto, and P. Puska, Sex, age, cardiovascularrisk factors, and coronary heart disease: a prospective follow-up study of 14 786middle-aged men and women in finland,Circulation, vol. 99, no. 9, pp. 11651172,1999.

[14] I. Yusoff, R. Yahya, W. Omar, and L. Ku, Non invasive cholesterol meter usingnear infrared sensor, innovation Commercialization of Medical ElectronicTechnology Conference (ICMET), 2015. IEEE, 2015, pp. 100104.

[15] T. T. Khan, N. Sultana, R. B. Reza and R. Mostafa, "ECG feature extraction in temporal domain and detection of various heart conditions," 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), Dhaka, 2015, pp. 1-6.

IoT Based Smart Energy Management in Residential Applications

Md. Rakibul Hasan
Department of Electrical and Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
rakibul.eeekuet@gmail.com

Eklas Hossain
Department of Electrical Engineering and Renewable Energy
Oregon Institute of Technology
Klamath Falls, USA
eklas.hossain@oit.edu

Hossain Mansur Resalat Faruque
Department of Electrical and Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
resalat_f@yahoo.com

Tipu Sultan
Department of Electrical and Electronic Engineering
International Islamic University
Chittagong, Bangladesh
ssultanstipu@yahoo.com

***Abstract*—The lack of credibility of data acquired through traditional means from a conventional metering system has reduced the efficacy of the energy management system. For overcoming the drawbacks of the conventional metering system, this paper proposes a smart IoT based energy management system (EMS), which not only ensures transparency and credibility of data but also offers flexibility to the consumers by offering the facility of remotely controlling their home appliances. Electricity consumption and its corresponding cost is calculated automatically and then the data is sent to a server in a periodic manner. For carrying out the task, NodeMCU has been utilized, which has the combined functionality of both microcontroller and Wi-Fi modules. By means of its microcontroller function, constant readings of energy meter are collected and saved in NodeMCU; then the data is transferred to a remote server on the internet via Wi-Fi. Consumers can observe real-time data on the server to gain a clear idea of electricity usage, therefore, they can reduce or increase the consumption as per their budget. This proposed system also provides the functionality of controlling the home appliances remotely over the Internet. Furthermore, electricity suppliers can produce bills from the server data automatically, making it a feasible solution. This paper also proposes and analyzes individual sub-metering of home appliances of high rating to observe and control their energy consumption. A detailed cost analysis of the system has been performed for ensuring its cost-effectiveness. According to the analysis, the total cost associated with the system is BDT 1064, which is fairly economical compared with existing systems. The add-ons of this proposed system can be also integrated with the existing prepaid meters of the market, eradicating its major drawbacks by providing added functionality.**

Keywords— IoT, Smart meter, Energy management, Home automation, Residential tariff.

I. INTRODUCTION

Bangladesh being a developing country, the energy demand is increasing by leaps and bounds. In order to meet with this rising energy demand, there has been a prolific increase in the implementation of alternative energy sources for generating more electricity, however, this is proving to be insufficient due to system energy loss. Therefore, a smart energy management system (SEMS) is necessary to ensure proper electrical power management, which will reduce the system energy losses. Furthermore, the integration of the SEMS with the internet enables the consumers to observe their consumption of electricity and take action according to their budget by remotely controlling the electrical equipment. It also provides a feasible mean to the electric suppliers for producing bills with minimal error.

With the advancement of technology, the accessibility of the internet is increasing day by day. Apart from being used for communication and entertainment purposes, the internet also provides a bunch of scopes in power system management, development, control, etc. Internet of things (IoT) is a cutting-edge technology that interconnects the 'things' over the internet [1]. These ‘things’ incorporate everything we can perceive of connecting-goods, buildings, appliances, machines, vehicles, plants, animals, and even human beings. It is estimated that the number of objects connected to the internet will reach 50 billion by 2020 [2]. IoT is a bidirectional approach, where objects export data to a distant server and the users can control the objects through the server by software programs and it is mainly done over Wi-Fi networks. The traditional electric power network has many limitations, however, by integrating with IoT it can be transformed into an intelligent power network [3]. The application of IoT can effectively improve the accuracy of real-time monitoring and control of smart grid [4, 5]. The informatization of power grid system can be sped up by utilizing IoT and thus is beneficial for effective management of the power grid infrastructure [6]. It can also be used for precise and accurate estimation of energy consumption [7]. Recent studies show that IoT can also be integrated with smart energy meters and can increase their accuracy [8-10]. IoT based smart energy meter has been proposed in [11-13]. However, it does not have the functionality of remotely controlling the home appliances. An efficient billing system can also be developed using IoT [14, 15]. It also possesses great potentiality in its application in the control system of smart homes [16-18].

This paper proposes a SEMS based on IoT, which ensures the credibility of smart meter data as well as provides the facility for controlling home electrical appliances remotely. The energy usage data of the consumer is stored in NodeMCU, which is then transferred to a remote server on the internet using Wi-Fi. NodeMCU is an open-source development kit composed of microcontroller and ESP8266 Wi-Fi module, where microcontroller part does the necessary calculations and control and ESP8266 helps in communication over Wi-Fi. The corresponding cost of the

consumed electricity is automatically updated, reducing the errors that occur in the conventional energy meter system. This proposed method also reduces the complexity of bill payment and the consumer can check his/her energy usage at any time. Furthermore, it gives consumers the ability to remotely control any home appliance. Hardware implementation of the system along with necessary software configuration has been carried out in this research. Additionally, a comparative cost analysis with the existing energy meters available in the market has also been shown in this paper.

The remaining part of the paper is organized as follows: in section II, the present condition of energy metering system in Bangladesh is represented. The tools and apparatus required for the proposed system and its working principle are discussed in section III. The problems that occur during the operation of the proposed smart energy meter and corresponding solutions are presented in section IV, which is followed by a cost analysis in section V. Finally, in section VI, the key features along with future scope are described as a conclusion.

II. Present Condition of Energy Metering System in Bangladesh

A common practice in Bangladesh is that a person from utility services goes to individual consumers to collect the energy meter data. The bill is prepared using those data and the bill paper, which contains all the information about the usage of electricity in that month, is sent to individual consumers for payment. Finally, consumers pay their bill according to the bill paper, which was sent to them by going to a particular station or a bank. Therefore, this process is very much hassling and requires a lot of time and steps, and since there is direct involvement of human in collecting the meter data, the credibility of this system is compromised.

There are system losses occurring in the power network due to technical and non-technical aspects [19, 20]. Technical reasons include losses in power generation equipment, transmission lines, distribution lines, etc.; whereas non-technical losses occur mainly due to customer level meter tampering. In order to reduce the losses occurring at the customer level, the prepaid meter has been introduced in some cities and the rest are on the way. Although the prepaid metering system offers more advantages than the conventional meter system, it also has some major drawbacks. The customers need to purchase a card to refill the account but purchasing the card is quite troublesome sometimes. The consumers also need to predict the monthly bill and recharge accordingly, which can create a situation of uncertainty. Furthermore, prepaid energy meters do not provide the consumer with the ability to control electrical appliances remotely.

III. IoT Based Smart Energy Meter

IoT-based smart energy meter solves the issues of prepaid energy metering by minimizing the complexities and mitigates the non-technical losses by ensuring the credibility of data. It also brings new important features, such as real-time viewing of consumption data and remote controlling of home appliances. A single phase static watt-hour meter is used to calculate the consumed energy and these data are extracted from the meter through a LED. Table 1 shows the specification of the watt meter used in this paper.

TABLE I. Specification of The Single Phase Static Watt Hour Meter

Supply System	Range
Voltage	0 to 240V (Phase to neutral)
Current	5A to 30 A
Frequency	50Hz ± 5%
Temperature	$-10°C\ to\ 50°C$

In order to perform IoT operation, a microcontroller along with Wi-Fi module is needed. The microcontroller sends the data to predefined internet server using Wi-Fi. NodeMCU is such a development board that provides combining the functionality of having both I/O ports like microcontroller and Wi-Fi module. This kit is based on the ESP8266 Wi-Fi module. It is programmed in an Arduino environment. Here, the microcontroller does all the calculations necessary, whereas ESP8266 provides the functionality of communication via Wi-Fi. The pin diagram of the NodeMCU is shown in Fig. 1.

Another required apparatus for the proposed system is the optocoupler, which provides isolation between relatively high voltage energy meter side and low voltage microcontroller circuit. The electrical signal from two isolated circuits is transferred using an optocoupler. Data from power circuit is fed to optocoupler and then microcontroller extracts the data from the optocoupler. The model used in this work is optocoupler – 4N25. Other components used for the proposed IoT-based smart energy meter are a 220V – 9V transformer to step down the supply voltage, IC 7805 to regulate dc voltage, diodes for rectification purpose, capacitors for smoothing the rectified dc, and some light emitting diodes (LED) to emulate the home appliances.

Fig. 2 shows the functional block diagram of the whole system. In the functional block diagram of a conventional meter, only the ac power supply, energy meter, and loads block exist. However, some additional blocks are needed to be added here to realize the IoT and energy management features. The energy consumption data is taken from the energy meter through the optocoupler. There is a LED in the energy meter named 'Cal', which blinks a certain fixed times per Kilo Watt

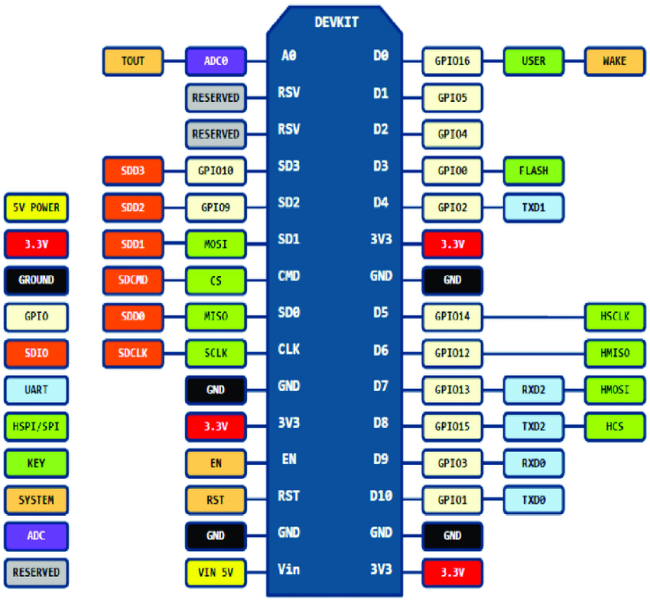


Fig. 1. Pin diagram of NodeMCU.

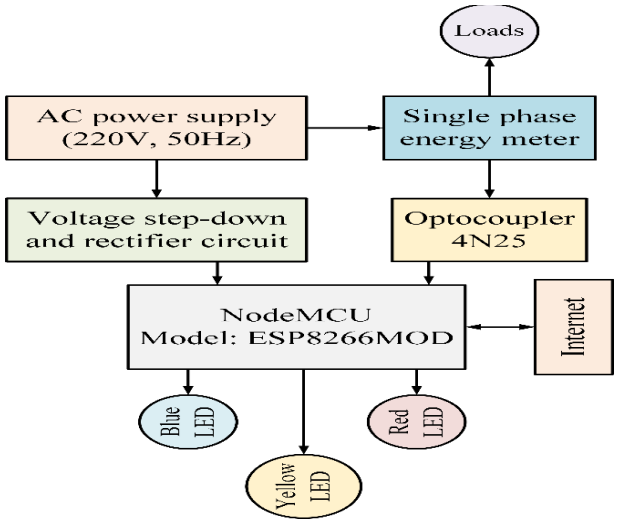


Fig. 2. Block diagram of the IoT based smart energy management.

Hour (kWh) for a particular energy meter. The meter used in this project has such a functionality, where the 'Cal' LED blinks 3200 times per kWh. This LED data is then fed to a microcontroller through the optocoupler. The number of blinking is observed by microcontroller programming and by using this data, the total electricity units consumed and associated cost are calculated.

The main feature of this system is to collect data over the internet. Therefore, the consumed unit of electricity and associated cost data need to be transmitted over the internet. This is done by writing the required Arduino code in the NodeMCU. In order to practically implement this system, NodeMCU needs internet connectivity and to store data on the internet, hence, a website domain is needed. However, in this prototype, the website is hosted locally and Wi-Fi hotspot connection is given to the NodeMCU to simulate the prototype. The Wi-Fi service set identifier (SSID) and password are written in the Arduino script. Therefore, NodeMCU gets connected to local Wi-Fi whenever Wi-Fi is available. The updated data are transmitted at a certain time interval.

The power supply of NodeMCU is taken from the main AC power supply. In order to use the 220V AC supply as 5V DC supply, voltage step down transformer and a rectifier circuit are used, which is shown in Fig. 3. In this circuitry, a step-down transformer is used to convert 220V AC to 9V AC. This 9V AC is then transformed into 9V DC by a full bridge rectifier circuit. The DC produced in the output of full bridge rectifier has a significant amount of ripples and in order to

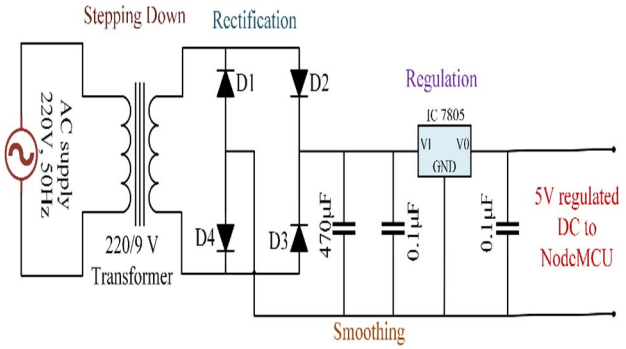


Fig. 3. Rectifier circuit to convert 220V AC to 5V DC.

reduce it, a smoothing capacitor is used in parallel. Finally, this 9V DC is transformed into 5V DC by using the 7805 regulator IC that is fed to the NodeMCU. The hardware implementation of the prototype is shown in Fig. 4.

Another important feature of this system is the control of home appliances over the Internet. A user of this proposed system can check the status of his home appliances and can take any action as he wants; that means he can turn on or off any home appliance remotely using the internet. In order to simulate this, three LED's are connected to NodeMCU and can be controlled over the Internet. The LEDs are connected to the general purpose input/output (GPIO) pins. Some predefined commands have been configured in the Arduino script to make the LED either on or off. When those commands are transmitted to the NodeMCU from the website, the LEDs will respond according to the specified commands and thus, they can be controlled easily. In real life application, this feature can be implemented by using a relay circuit. By controlling the relay, electrical appliances like fan, light etc. can be turned on or off as required.

When a user logs in to his account, he will see a webpage, which is shown in Fig. 5. This webpage has been developed using the basic HTML and bootstrap library which are open-source resources. Consumed unit of electricity and associated cost are shown on the webpage. The consumer can remotely

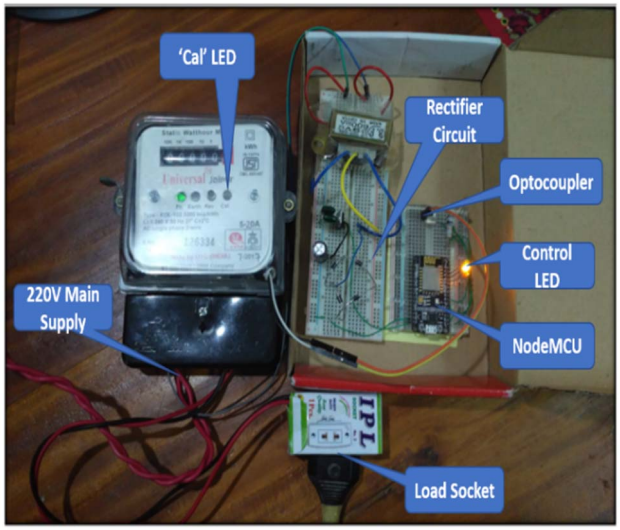


Fig. 4. Hardware implementation of the prototype of IoT based smart energy management system.

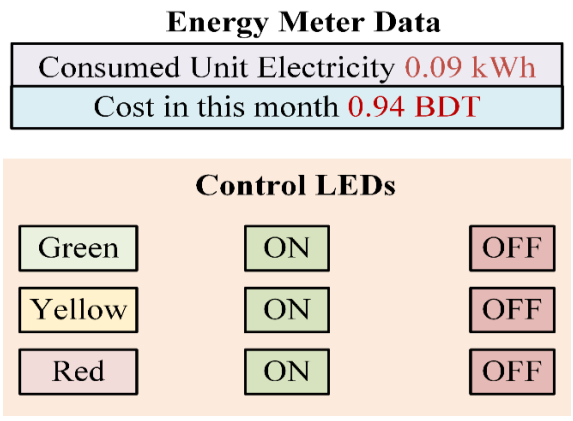


Fig. 5. Webpage showing the consumed electricity and their associated cost along with the feature of controlling the appliances remotely.

control any home appliance using this webpage. For this instance, in order to control the green, yellow, and red LEDs over the Internet, two buttons have been configured as ON or OFF for each LED and they can be turned on or off individually just by clicking on the corresponding buttons. A comparative analysis of the proposed system with existing works is tabulated in table II. The proposed system does not require any additional microcontroller as the NodeMCU has the dual functionality of microcontroller and a Wi-Fi module. Furthermore, the proposed system also has the feature of controlling home appliance remotely.

TABLE II. COMPARISON OF PROPOSED METHOD WITH EXISTING WORKS

Authors	**Microcontroller**	**Optocoupler**	**Functionality of controlling of home appliance remotely**
Abate *et al.* [11]	STM32F205RB	Not used	Not available
Sahani *et al.* [12]	ATMega328	MOC3071	Not available
Singh *et al.* [13]	ATMega328 and PIC16F877P	Not used	Not available
Proposed	Built-in microcontroller in NodeMCU	4N25	Available

The proposed system has been designed to determine the electricity consumption and calculate the corresponding cost for residential consumers only. Let us consider b is number of blinks of the 'Cal' LED in a month. As the 'Cal' LED blinks 3200 times per kWh, the total energy consumption (kWh) is

$$a = \frac{b}{3200} \quad (1)$$

If s and d are the service charge and demand charge per month in Bangladeshi Taka (BDT), respectively, then the associated cost of that month c can be calculated using the following equation

$$c = (a \times k) + s + d \quad (2)$$

Here, k is per unit price of energy in (BDT), whose value depends on a. According to Dhaka Power Distribution Company (DPDC), the value of k for different consumed unit (a) is shown in table III [21].

TABLE III. TARIFF RATE FOR RESIDENTIAL CONSUMERS

Consumed Unit (a)	**Per Unit Rate of Energy (k)**
0-75	3.80 BDT
76-200	5.14 BDT
201-300	5.36 BDT
301-400	5.63 BDT
401-600	8.70 BDT
601-above	9.98 BDT

The existing tariff rates in Bangladesh for different categories according to DPDC [21], is shown in Fig. 6. The variation of service charge s and demand charge according to DPDC [20] is presented in Figure 7.

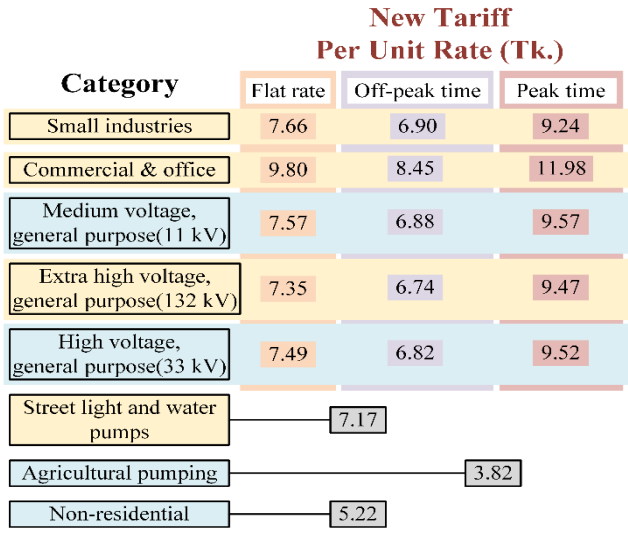


Fig. 6. Existing tariff rates in Bangladesh [21].

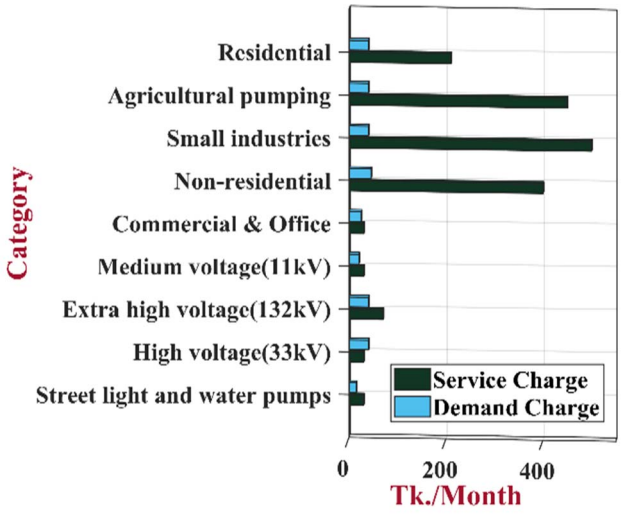


Fig. 7. Variation of demand charges and service charges in Bangladesh [21].

IV. PROBLEMS AND SOLUTION

NodeMCU needs power from the main supply to operate; thus when power interruptions occur, it will be turned off. However, there is no possibility of data loss as the data is stored periodically in the internal memory of the NodeMCU. Each time the NodeMCU starts, it extracts the most recently updated data from the memory. This system needs an Internet connection to send data to the web server and there might be the possibility of having no internet connection at a certain time. However, this problem is solved as NodeMCU sends the updated data whenever it regains internet connectivity. Data are always updated and stored in the internal memory of NodeMCU and is sent to the server whenever the internet is available. Therefore, there is no possibility of data loss. The exchange of data over the public domain such as the Internet is very much susceptible to unauthorized access [22]. Therefore, the system will be synchronized to a distant server and it can be programmed so that any undesirable attempt will

inform the consumer as well as the electricity provider. In this way, the risk of data theft can be reduced to a minimum.

V. Cost Analysis

There are various types of energy meters available in the market, their price ranging from BDT 2000 to 10000. Different prepaid energy meters available in the market and their corresponding prices are shown in table IV.

TABLE IV. Price of Available Prepaid Meters In Market

Product Title	Market Price
Liberty Prepaid Energy Meter	BDT 9880
Techno Prepaid Energy Meter Class 1.0	BDT 5070
Elmeasure Prepaid Energy Meter	BDT 5000
Meter by Yueqing Haixin Electronic Company	BDT 2540

Cost of different equipment required for the proposed system and the overall cost is shown in Table V. In our proposed design we used the Universal Energy Meter which is a conventional energy meter. Additional equipment such as, NodeMCU, optocoupler, and voltage step down circuitry are used as add-ons so that this conventional energy meter can function as IoT based smart energy meter. Furthermore, it has the facility to provide the consumers the ability to control their home appliance remotely. The aforementioned add-ons are also compatible with the existing prepaid meters available in the market. Therefore, the consumers are able to pay their bills without having to predict their monthly consumption of electricity with the added facility of remotely controlling their home appliance. From data presented in table III and IV, it can be said as the add-ons cost is very cheap, therefore, it is very much cost-effective integrate them with conventional or prepaid energy meters.

TABLE V. Costing of Different Equipments of The IoT Based Energy Mangement System

Equipment		Market Price
Universal Energy Meter		BDT 430
Add-ons	NodeMCU V-3 Development Kit	BDT 500
	4N25 Photo Transistor Optocoupler	BDT 15
	Voltage Step Down Circuitry	BDT 120
Total		**BDT 1065**

VI. Conclusion

In this paper, integration of IoT with energy management system has been demonstrated to make a more effective and reliable system compared with the conventional energy management system. Instead of collecting data from door to door, the key feature of the proposed system offers a more accessible way of collecting data from a server through internet, which is automatically updated after a short time interval via Wi-Fi. The NodeMCU used in this system has the dual functionality of a microcontroller and a Wi-Fi module, which eliminates the need for an additional microcontroller. Furthermore, remote control of domestic appliances has also been made possible through this system. The total hardware implementation along with interfacing with the internet has been presented in detail. For ensuring the cost effectiveness, a detailed cost analysis has been conducted. The overall cost of the proposed IoT based smart energy management system is BDT 1065 and add-ons cost is BDT 635. Hence, the proposed system is cost effective and the add-ons can be also integrated with the existing prepaid meters, which can expunge its major drawbacks by providing added functionality. Therefore, the presented design in this paper, possesses high potentiality in its application in energy metering system of Bangladesh.

References

[1] S. N. Han, I. Khan, G. Myoung, N. Crespi, and R. H. Glitho, "Computer Standards & Interfaces Service composition for IP smart object using realtime Web protocols : Concept and research challenges," *Comput. Stand. Interfaces,* vol. 43, pp. 79–90, 2016.

[2] M. Swan, "Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0," *J. Sens. Actuator Networks,* vol. 1, pp. 217–253, 2012.

[3] G. Bedi, G. K. Venayagamoorthy, R. Singh, R. R Brooks, and K. C. Wang, "Review of Internet of Things (IoT) in Electric Power and Energy Systems," *IEEE Internet of Things Journal,* vol. 5, pp. 847-870, 2018.

[4] L. Li, H. Xiaoguang, C. Ke and H. Ketai, "The applications of WiFi-based Wireless Sensor Network in Internet of Things and Smart Grid," *6th IEEE Conference on Industrial Electronics and Applications*, pp. 789-793, 2011

[5] Miao Yun and Bu Yuxin, "Research on the architecture and key technology of Internet of Things (IoT) applied on smart grid," *International Conference on Advances in Energy Engineering*, pp. 69-72, 2010.

[6] Q. Ou, Y. Zhen, X. Li, Y. Zhang, and L. Zeng, "Application of internet of things in smart grid power transmission," *FTRA International Conference on Mobile, Ubiquitous, and Intelligent Computing,* pp. 96-100, 2012.

[7] F. Englert, P. Lieser, A. Alhamoud, D. Boehnstedt and R. Steinmetz, "Electricity-Metering in a Connected World: Virtual Sensors for Estimating the Electricity Consumption of IoT Appliances," *3rd International Conference on Future Internet of Things and Cloud*, pp. 317-324, 2015.

[8] D. Alahakoon, and X. Yu, "Smart Electricity Meter Data Intelligence for Future Energy Systems: A Survey," *IEEE Transactions on Industrial Informatics,* vol. 12, pp. 425-436, 2016.

[9] P. D. Talwar, and S. B. Kulkarni, "IoT Based Energy Meter Reading" *International Journal of Recent Trends in Engineering & Research (IJRTER),* vol. 2, pp. 586-591, 2016.

[10] E. Hossain, J. Hossain and F. Un-Noor, "Utility Grid: Present Challenges and Their Potential Solutions," *IEEE Access,* vol. 6, pp. 60294-60317, 2018.

[11] F. Abate, M. Carratù, C. Liguori, M. Ferro, and V. Paciello, "Smart meter for the IoT," *IEEE International Instrumentation and Measurement Technology Conference (I2MTC),* pp. 1-6, 2018.

[12] B. Sahani, T. Ravi, A. Tamboli, and R. Pisal, "IoT Based Smart Energy Meter," *International Research Journal of Engineering and Technology (IRJET),* vol. 4, pp. 96-102, 2017.

[13] A. Singh, and R. Gupta, "IoT based smart energy meter," *International Journal of Advance Research and Development,* vol. 3, pp. 328-332, 2018.

[14] P. Mahajan, S. Tatia, and P. Jadhav, "Smart Meter: An IOT Based Smart Meter for Live Electricity Monitoring and Bill Payment," *International Journal of Electrical, Electronics and Computer Systems (IJEECS),* vol. 5, pp. 49-51.

[15] N. N. Sasane, N. S. Sakat, K. N. Shital, V. R. Kaushik, and P. K. Pallav, "IOT based energy meter billing and monitoring system - A case study," *International Research Journal of Advanced Engineering and Science*, vol. 2, pp. 64-68, 2017.

[16] S. Chen, C. Lai, Y. Huang, and Y. L. Jeng, "Intelligent home-appliance recognition over IoT cloud network," *9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 639-643, 2013.

[17] W. T. Cho, Ying-Xun Lai1, C. F. Lai, and Y. M. Huang, “Appliance-Aware Activity Recognition Mechanism for IoT Energy Management System,” *The Computer Journal,* vol. 56, pp. 1020-1033, 2013.

[18] M. Wang, G. Zhang, C. Zhang, J. Zhang and C. Li, "An IoT-based appliance control system for smart homes," *Fourth International Conference on Intelligent Control and Information Processing (ICICIP)*, pp. 744-747, 2013.

[19] F. I. Bappy, M. J. Islam, A. K. Podder, D. R. Dipta, H. M. R. Faruque, and, E. Hossain, “Comparison of Different Hybrid Renewable Energy Systems With Optimized PV Configuration to Realize the Effects of Multiple Schemes,” *1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT),* in press.

[20] E. Hossain, I. Khan, F. Un-Noor, S. S. Sikander, and M. S. H. Sunny, "Application of Big Data and Machine Learning in Smart Grid, and Associated Security Concerns: A Review," *IEEE Access,* vol. 7, pp. 13960-13988, 2019.

[21] *Tariff Rate*. [online] Available at: https://dpdc.org.bd/article/view/52/Tariff%20Rates [Accessed Apr. 11, 2019].

[22] M. S. Rahman, M. S. Hossain, E. H. Rahat, D. R. Dipta, H. M. R. Faruque, and F. K. Fattah, "Efficient Hardware Implementation of 256-bit ECC Processor Over Prime Field," *International Conference on Electrical, Computer and Communication Engineering (ECCE),* pp. 1-6, 2019.

IoT-based Portable ECG Monitoring System for Smart Healthcare

Tamanna Shaown[1*]
Department of EEE,
BRAC University,
Dhaka, Bangladesh
tamannashaown@gmail.com

Imam Hasan[2]
Department of EEE,
Atish Dipankar University of Science & Technology
Dhaka, Bangladesh
imam.tasin@gmail.com

Md.Muradur Rahman Mim[2]
Department of CSE,
Atish Dipankar University of Science & Technology,
Dhaka, Bangladesh
muraduur@gmail.com

Dr. Md. Shohrab Hossain[3]
Associate Professor, Department of CSE,
Bangladesh University of Engineering & Technology,
Dhaka, Bangladesh
shohrab29@gmail.com

***Abstract*—Sudden and unexpected death due to heart failure is a major cause of mortality among middle aged and elderly people. An efficient heart monitoring system can find out the malformation of heart conditions and that can also be helpful in diagnose at critical ambience. Sometimes the distance between patients and doctors is the main barrier that people do not have access to quality health services and thus having trouble for their regular health examine. IoT-based healthcare monitoring system is one of the manifested application areas in medical science. As a significant access in diagnose heart disease ECG observing system is widely used. In this paper, a progressive method for ECG monitoring system based on Internet of Things (IoT) has been proposed. In this study, a system is designed to frequently monitor the Electrocardiogram (ECG) signal collected from patient's body using wearable sensors and the data is stored into the database which can be accessed by authorized personnel only. When any malformation is found an automatic email is sent to the users and doctors for analyzing about the critical conditions of the patients and provides emergency health assistances. In order to verify the authenticity of this system tests have been implemented on several patients and the report shows that, this system is dependable and efficient for collecting real time ECG data which can be very helpful in diagnose heart diseases. This IoT-based low cost device can be reliably used to reduce the risk of disability and mortality rate due to cardiovascular diseases.**

Keywords—Internet of Things (IoT), Smart Healthcare, ECG Module, IoT Cloud, Graphical User Interface (GUI).

I. INTRODUCTION

The Internet of Things (IoT) has many purposes in medical areas such as distant patient monitoring, dietary program, testimony of chronic diseases, elderly people care etc. The elderly population is rising continually both in rural and urban areas, so observing their health regularly with the desire of staying at home and at low cost is one of the basic appeal. Healthcare personnel can assess, diagnose and prescribe elderly people by gathering medical information from remote regions by using IoT. The significant part of IoT can be used in so many medical equipment, sensors and imaging devices to primarily determine the patient's health statuses and to deliver them proper care in a shortest possible time. IoT-based healthcare services can be implemented at low cost hereby maximizing the user contentment. Thus, the main purpose of IoT-based healthcare system is its cost effectiveness and secures interconnection for individuals, clinics and different healthcare institutions.

Health conditions of aged people needed to be checked in regular basis which is a greater challenge both in medical research and hospitals [1].In medical institutes or hospitals, conventional 12 leads systems are used to collect ECG signals from patient's body. This device is too expensive to use at home or personal use and cannot be used as portable. It takes huge amount of time for routine checkup during each visit to hospitals and manual recording can also cause human made error. In such a manner, a low cost portable system for continuous ECG monitoring is highly required. Most of the available ECG monitoring devices use smart phones for secondary data processing [2]. The uses of smart phones in data transfer and processing have great consequences on regular use due to its limitations of power and computational effectiveness.

In our study, a wearable ECG monitoring system has been proposed and implemented occupying Internet of Things (IoT). This customized device can directly sent ECG data collected from patient's body to IoT cloud using *Wi-Fi*. *Wi-Fi* has wider coverage areas and higher data transfer rate than Bluetooth or ZigBee. The whole work is done by connecting Raspberry Pi with internet. Here, ECG sensor AD8232 has been used for ECG monitoring. Data have been processed in Raspberry Pi which uses processing software and a band pass filter is used for eliminating different noises to form an electrocardiogram (ECG) graph. An automatic email will be sent to the users and authorized doctors in case of abnormalities found in patient's body. This device is hazard free and cost effective for the remote patients as they do not need to visit the doctors regularly. We have compared our customized device with standard 12 lead ECG modular in generating heart rates and waveforms and the result shows that, our system can be used reliably.

The rest of the paper is organized in following sections. In section II, background and related works is discussed. In section III, system architecture of IoT based ECG monitoring system has been proposed and being explained. In section IV, shows the flowchart and implementation of IoT based ECG monitoring system. We conduct several tests on patients in order to verify the reliability of the system and it is discussed as experimental results and analysis in section finally we conclude the paper in section VI and also

discussed about the future scopes in the field of medical science using Internet of Things (IoT).

II. Background and Related Works

Internet of Things (IoT) has come into view as an important technology for health monitoring system. In professional medical institute ECG data is collected employing 12 electrodes due to their good performance in short term calibrating. But these devices are unlikely to be portable and patients often have to visit medical institutes for regular checkup which is also very expensive. Thus, there is a great demand in the domain of health care to support elderly people to care at home. So we have recommended allow cost ECG monitoring system and implement it so the elderly people can get precise care without travelling the hospitals frequently.

Spine, Pascoli, Iannaccone,et al. developed a low power wearable ECG monitoring system for multiple patient remote monitoring [3].The authors developed low-power ECG monitoring system with a wearable facility ZigBee protocol for data transfer from sensor but there was no real time notification in this system. Sneha and others [4] proposed an architectural framework using mobile to monitor the ECG of cardiac patients. The main drawback of the system was it cannot be controlled without a mobile application. But our device can be controlled by IoT.

An integrated wireless system was proposed for heart failure patients who can detect ECG signal, Heart Rate Variability (HRV) continuously using software based algorithm [1]. IoT can be used in identifying human diseases, since heart failure is a major problem now-a – days; ECG monitoring is necessary both in homes and hospitals [5]. In hospitals ECG data is collected using 12 electrodes. Basically the ECG instrument measures the difference in voltage potential between each of the electrodes. In medical equipment which consist 12 electrodes waveforms only 8 waveforms can be measured and rest of the 4 is extrapolated by measured waveforms. The equipment which is used in hospitals is not portable and also expensive for regularly use in home. So patients need to go to hospitals for their regular checkup. A low cost ECG monitoring system for home and hospital uses is highly needed. A low power mixed signal SOC for portable ECG monitoring using Bluetooth protocol has been introduced. This system is portable because electrodes are usually less than traditional methods [6].

Electrocardiography (ECG) for heart disease using medical e-services patients is a crying need for development countries. Because in traditional ECG monitoring process the collected data is saved and simultaneously sent to the doctors for further analysis. It might cause delay in emergency time and sudden heart failure cannot be detected from this system. In this paper sensor based system architecture is proposed for ECG monitoring based on Internet of Things. Finally based on the architecture we design an ECG based heart rate monitoring system .The collected ECG data will be directly sent to the database server using ***Wi-Fi*** module. ***Wi-Fi*** has been used here because it can provide larger cover areas and higher data rates. Wi-Fi is used here that it can send email to the doctors or relatives in case of patient's emergency health conditions. The proposed system is successfully developed and tested to examine the effectiveness and reliability of the system. This is a low cost smart device which can be used for patient's health monitoring so that doctors can take precautions at proper time to avoid sudden death due to cardiovascular diseases.

III. System Architecture of IOT Based ECG Monitoring System

The architecture of the IoT-based ECG monitoring system is illustrated in "Fig.1", which mainly consists of three parts, i.e., the ECG sensing network, IoT cloud, and GUI. The components used in ECG Sensing network are: ECG AD8232 Sensor, Raspberry Pi Model 2, and Arduino Uno.

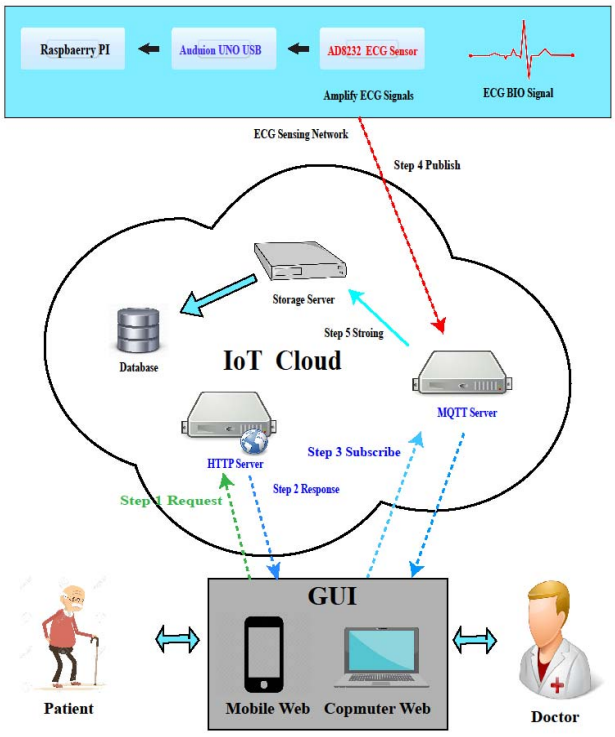


Fig. 1 Proposed Architecture of the IoT-based ECG monitoring systems.

A. ECG Sensing Network: ECG sensing network is set for assembling physiological data from the body surface and pass on these data to IoT cloud through a wireless channel. In our equipment wearable ECG sensor has used to gather data from patients body over long hours. Then the ECG signals are processed through amplification and filtering etc. to improve the signal quality. The ECG data gathered from sensors are transmitted to the IoT cloud via a specific wireless protocol such as Bluetooth, ***Wi-Fi***, ZigBee etc. [7]. With satisfying energy consumption all these three protocols can transfer enough data rates for transmitting ECG signals. Moreover, due to limited communication

ranges of Bluetooth and ZigBee,***Wi-Fi*** is used in our proposed system .Comparisons among various types of ECG sensing networks are given in Table I.

TABLE I.
COMPARISONS AMONG TYPICAL ECG SENSING NETWORKS SUCH AS ***WI-FI***, BLUETOOTH AND ZIGBEE.

Standards	***Wi-Fi* based ECG sensing network**	**Bluetooth based ECG sensing network**	**ZigBee-based ECG sensing network**
Protocol	IEEE 802.11	IEEE 802.15.1	IEEE 802.15.4
Coverage	20-200 m	20-30 m	2-20 m
Data rates	11-54 Mbps	3-24 Mbps	10-250 kbps
Power consumption	Medium	Low	Low
Terminal dependency	Data collection is independent of smart terminals	Smart terminals are needed for receiving and forwarding sensed data.	Smart terminals are needed for receiving and forwarding sensed data.

I. Raspberry Pi: Raspberry Pi is a low cost, small sized, powerful minicomputer which can interface with many devices. Here we used Raspberry pi model-3b.In this system we use Raspbian operating system which is Linux based. ***Wi-Fi*** is a built in characteristics in Raspberry- Pi. So we need not to connect Wi-Fi module for data sending.

II. Arduino Uno: Arduino Uno is a microcontroller board on ATmega328P (datasheet). It is used in this system for detecting ECG signals from AD8232.

III. ECG Sensor: ECG sensor AD8232 shown in "Fig. 2" is used for collecting ECG data from human body by using electrodes on the patient's skin. We collected signals from human body via AD8232 using Arduino Uno; it sent the collected signals to raspberry pi. Raspberry pi then processed the signal by using Processing 3.0 and shows the image finally on display.



Fig. 2. AD8232 ECG Sensor

B. IoT Cloud: With the help of IoT cloud in ECG monitoring system we can store data, modify data and all the patients information's are saved here. It can also send disease warning and protecting patients from getting injured.

C. GUI: Graphical User Interface (GUI) is used for data imagination management. It contributes easy entry of the data in the IoT cloud. Users can log onto the cloud to acquire visualized ECG data in real time. Generally mobile applications and web pages are the two kinds of GUI's are available for users to visualize ECG data. Although mobile app can ensure immediate response but web pages are the best options in terms of protection and up-gradation.

IV. FLOWCHART AND IMPLEMENTATION OF IOT BASED ECG MONITORING SYSTEM

"Fig.3," shows the flow chart of our proposed system. At first the user switch 'on' the device and the device will check whether all the modular are 'on' or 'off'. If the modular is 'off' then the system will remain 'off'. Otherwise, when the modular is on so the system will run and start to collect data from human body. After that, these data will be saved automatically in local database, when internet connection is found it will be analyzed the data in cloud database and show results for abnormal condition and send it to the users via ***Wi-Fi***. The system proposed and implemented in this paper take ECG data from patients' body and at any case of abnormalities an email is sent to the doctors so that, they can take necessary initiatives to prevent damages due to cardiovascular diseases.

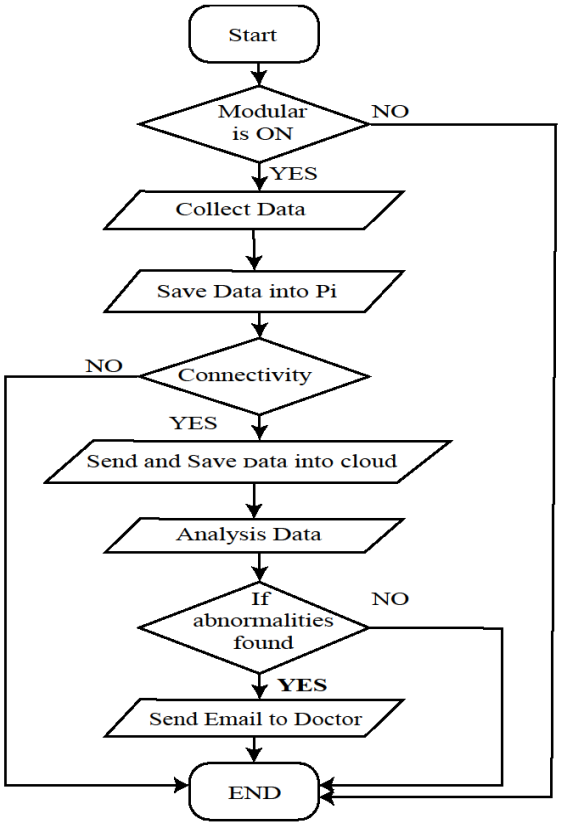


Fig. 3. Flowchart of the IoT -based ECG monitoring system

In medical institutes or hospitals conventional 12 leads systems are used to collect ECG signals from patient's body. This system cannot be used as portable. But for sudden analysis of patient's condition 3-lead system is

sufficient [8]. "Fig.4". shows the physical appearance of our ECG monitoring system using 3-lead system.

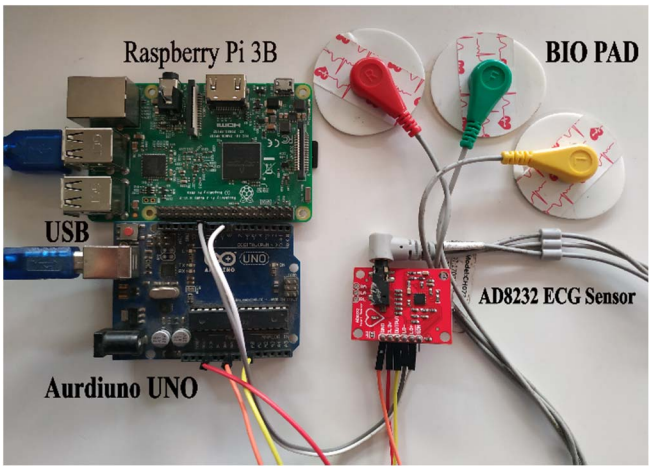


Fig. 4. Physical appearance of ECG monitoring system

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A 3-lead placement is adequate to analysis the primary features of ECG signal [9] compared to conventional 12-lead ECG monitoring device used in hospitals. In order to best sample the ECG signal, the electrodes need to be placed around the heart and form a triangle. In order to justify the performance of our system we have conducted several tests on some patients.

Typical ECG signals mainly consist of five type of waves such as the P wave, T wave, Q wave, R wave, and S wave, as shown in "Fig. 5". The intervals of these waves are used to diagnose several of heart diseases.

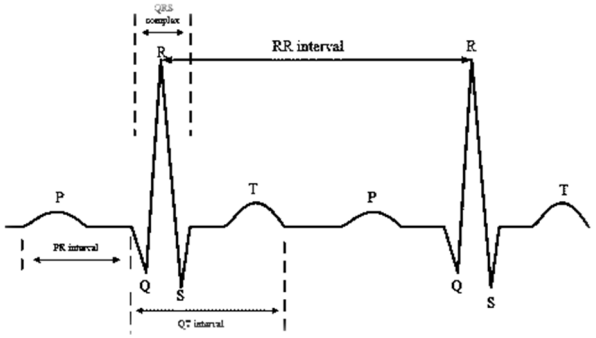


Fig. 5. Standard ECG Signal

- **RR** interval: As one of the most conspicuous characteristics, the R wave is often used to identify the period of an ECG signal. RR interval indicates the time interval between two adjacent R waves, which may become irregular in the event of some heart diseases, for example, the arrhythmia.
- **PR** interval: The PR interval measures the time between the beginning of the P wave and that of the QRS complex. It indicates the time the impulse takes to reach the ventricles from the sinus node.
- **QT** interval: QT interval represents the time between the start of the Q wave and the end of the T wave, which is related to the ventricular depolarization and repolarization. There is an increased risk of ventricular fibrillation or even sudden cardiac death if the QT interval exceeds the normal value.
- **QRS** complex: QRS complex is mainly associated with the Ventricular depolarization, which consists of three important waves, i.e., Q wave, R wave and S wave. By analyzing the QRS complex, certain diseases are likely to be detected such as drug toxicity and electrolyte imbalance.

According to [9] [10], the normal values of ECG parameters are listed in Table II.

TABLE II.
NORMAL VALUES OF KEY ECG PARAMETERS.

Features	Normal Range (Second)
RR Interval	0.6- 1
PR Interval	0.12- 0.20
QT Interval	0.32- 0.44
QRS Complex	< 0.12

Abnormalities are found when the two conditions are not valid. Finally, if any disorders are found then an emergency email has been sent to the users or doctors that, they can take emergency steps to prevent severe damage of the patients. This can reduce mortality rate and any kind of damages that causes due to heart failure.

To get the most proper data for sudden analysis of patient's condition 3 lead electrodes are placed in a triangular shape around the heart. "Fig.6", shows, the ECG data collected from a 40 years old healthy person. It is evident that the intervals between adjacent R waves (RR interval) are nearly the same, which shows no risk of developing arrhythmias.

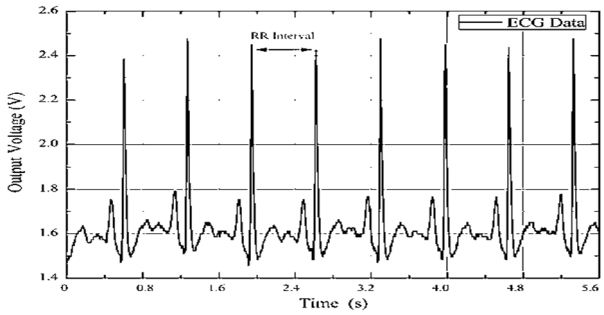


Fig. 6. Real time ECG signal of a healthy volunteer.

In order to analyze the key features of the measured ECG signal, two cycles of the signal are chosen as examples, which are shown in "Fig. 7".

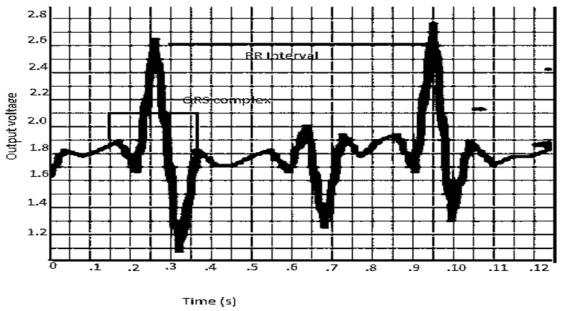


Fig. 7. Key features of the measured ECG signal.

In “Fig. 7”, it can be observed that, QRS complex is also regular indicating no abnormality in the processes of ventricular depolarization and repolarization. QT interval is also normal. Finally, the PR interval is also within the normal range, which implies that the process of atria depolarization is normal as well. So this person is not suffering from any heart diseases.

We have developed a prototype system and tested our device with a 12-lead ECG machine and the experimental result shows the reliability of our system

Percent Difference in “Equation (1)” is used to determine the relative difference between 12 lead and 3 lead portable ECG system.

$$\% \text{ difference} = \left| \frac{Expected - Actual}{\frac{Expected + Actual}{2}} \right| * 100\% \qquad (1)$$

Where, Expected = ECG value gained from 12-lead ECG machine and Actual = ECG value gained from ECG device prototype.

TABLE III.
SUMMARY OF PERCENT DIFFERENCE OF ECG PARAMETERS

ECG Parameter	Data of 12 lead ECG Device	Data of 3 lead ECG Prototype Device	Percentage Difference (%)
RR Interval	0.7709	0.7785	0.98
PR Interval	0.1232	0.1544	20.2
QT Interval	0.3496	0.3507	0.31
QRS Complex	0.0959	0.1040	8.104

From the Table III, we can observe that, the value of our prototype is quite similar with expected value collected from 12-lead ECG device.

To determine the efficiency rate of our device "Equation 2" is used to test the number of successful attempts vs. the total number of attempts.

$$\% \text{ reliability} = \left| \frac{number\ of\ successful\ attempts}{total\ number\ of\ attempts} \right| * 100\% \qquad (2)$$

We have tested our prototype device on 10 patients at the cardiology department of Delta Medical College and Hospital and we got 80% efficiency in detecting ECG conditions. From the samples collected from our prototype four trials (4, 8, 9, and 10) did not match with the 12-lead machine. The results have been shown below in Table IV.

TABLE IV.
TABULATED SUMMARY OF COMPARISON BETWEEN TWO DEVICES

Trial	Trial 12-lead ECG Device	Trial 3-lead ECG Device
1	Normal Sinus Rhythm	Normal Sinus Rhythm
2	Normal Sinus Rhythm	Normal Sinus Rhythm
3	Normal Sinus Rhythm	Normal Sinus Rhythm
4	Normal Sinus Rhythm	Sinus Bradycardia
5	Normal Sinus Rhythm	Normal Sinus Rhythm
6	Normal Sinus Rhythm	Normal Sinus Rhythm
7	Normal Sinus Rhythm	Normal Sinus Rhythm
8	Normal Sinus Rhythm	Sinus Bradycardia
9	Sinus Bradycardia	Normal Sinus Rhythm
10	Normal Sinus Rhythm	Sinus Tachycardia

V. CONCLUSION AND FUTUTRE WORKS

The implemented prototype of our IoT-based ECG monitoring system is highly efficient, low cost device and it reduces the time and cost for regularly visiting the doctors at healthcare institutions. We have collected data from different ages’ patients by using three electrodes placements in their body and achieved 80% accuracy in our system. The data collected from our device has been directly sent to IoT cloud using ***Wi-Fi***. The IoT cloud has been used for visualizing the data to users and store it for future analysis. By using the system ECG signal can be monitored continuously. The regular use of device is very helpful for preliminary detection of heart diseases and to reduce severe damage and mortality rate due to cardiovascular diseases. Similar to this user-friendly ECG monitoring System, additional health monitoring systems such as temperature measurement, Blood Pressure, Diabetes, etc., can be developed using IoT that will greatly help decrease existing health care problems to a certain amount.

ACKNOWLEDGEMENTS

This work has been supported by the Cardiology Department of *Delta Medical College and Hospital*, Darus Salam Road, Dhaka-1216.

REFERENCES

[1] Abadi, M., Subramanian, R., Kia, S., and et al., DECAF: MEG-based multimodal database for decoding affective physiological responses. IEEE Trans. Affect. Comput. 6:209–222, 2015.

[2] Miao, F., Cheng, Y., He, Y., and et al., A wearable contextaware ECG monitoring system integrated with built-in kinematic sensorsof the smartphone.Sensors 15:11465–11484, 2015.

[3] E. Span, S. Di Pascoli and G. Iannaccone, “Low-Power Wearable ECG Monitoring System for Multiple- Patient RemoteMonitoring’’ in IEEE Sensors Journal, July 1, 2016.

[4] S. Sneha and U. Varshney, "A Wireless ECG Monitoring System for Pervasive Healthcare," in International Journal of Electronic Healthcare, 3(1), 32-50, 2007.

[5] Kim H, Kim S, Van Helleputte N et al. (2014) A configurable and low power mixed signal SoC for portable ECG monitoring applications. IEEE Transactions on Biomedical Circuits and Systems 8:257-267.

[6] J. Dong and H. Zhu, "Mobile ECG detector through GPRS/Internet," presented at Computer-Based Medical Systems, 2004. CBMS 2004. Proceedings. 17th IEEE Symposium on, 2004.

[7] Palattella, M., Dohler, M., Grieco, A., and et al., Internet of Thingsthe 5G era: Enablers, architecture, and business models. IEEE J. Select. Areas Commun. 34:510–527, 2016.

[8] Wang Y et al. Design and evaluation of a novel wireless reconstructed 3-lead ECG monitoring system. In: Proc. IEEE 2013 Biomedical Circuits and Systems Conference (BioCAS), Rotterdam, Oct. 2013, pp. 362-365.

[9] ECGlibrary.com: Normal adult 12-lead ECG. In: Ecglibrary.com, http://www.ecglibrary.com/norm.php. Accessed 18, Sept. 2016.

[10] Gertsch M The Normal ECG and its (Normal) variants. In: The ECG manual. Springer, London, pp. 17-36,2009.

A Complete Device to Call Candidates and Time Management in Viva-voice

Mohammed Abdul Kader
Dept. of EEE
International Islamic University Chittagong(IIUC)
Chittagong, Bangladesh
kader05cuet@gmail.com

Rajvi Sutra Dhar
Dept. of EEE
International Islamic University Chittagong(IIUC)
Chittagong, Bangladesh
tanmoy.sd333iiuc@gmail.com

Md. Saiful Islam
Dept. of EEE
International Islamic University Chittagong(IIUC)
Chittagong, Bangladesh
si.iiuc@yahoo.com

Naeemul Islam
Dept. of EEE
International Islamic University Chittagong (IIUC)
Chittagong, Bangladesh
naymnaeem08@yahoo.com

Mohammad Abdur Razak
Dept. of EEE
International Islamic University Chittagong (IIUC)
Chittagong, Bangladesh
razzak123006@gmail.com

Yasir Arafat
Dept. of EEE
International Islamic University Chittagong(IIUC)
Chittagong, Bangladesh
yaeeeiiuc@gmail.com

***Abstract*—In this paper, a customized clock and alarming system is developed, which can manage time and call the candidates sequentially during important viva-voice or presentation. The main purpose of this work is to free the intricacy of viva-voce or presentation time and to make it captivating to both the examiners and the examinees. To perform this action, we have utilized microcontroller, 32X8 LED dot matrix display, 6 digit multiplex seven segment display, radio frequency (RF) transmitter-receiver (Tx-Rx) pair, IR receiver, IR remote and buzzer. Seven segment display unit is used to show time, group number and also to produce an alarm to alert the presenter about the time during the presentation. User can set time and group number, start countdown and control other functions of seven segment display unit by an IR remote. Dot matrix display unit shows group number and other information to the waiting candidates which communicate with 7-segment display unit by RF Tx-Rx pair. By using this system only the examiner of the viva-voce or academic presentation can smoothly conduct the session without the help of an assistant.**

Keywords—matrix display, microcontroller, digital learning, wireless control, RF Tx-Rx; EUSART

I. INTRODUCTION

In the age of technology, people are utilizing technology to make their lives more comfortable and accessible. Every aspect of life is deeply dependent on technology [1]. Technology can be used to make the academic viva broad or presentation easy. The current method is not a convenient way, because the utilizer needs adscititious human resources and effort to manage the overall process, which lags the whole process.

Consequently, it would be more convenient if we can control the electronic contrivance instead of any other utilizer utilizing a remote unit [2]. An apparatus which can deal with time management and calling someone is different from each other. The time management system is a vital part in all over the world. Whenever in time management system it requires some exhibit where the time can be exhibited. Additionally, in the calling process, it can be an intercom bell or another contrivance is needed. Whenever we require both the device as a single contrivance, it's not available. But the method is the most frugal mode of consuming time over all other designates. Much related work has been reported for the same function by different groups with different approaches. Many home contrivances can be controlled with a designed system utilizing a microcontroller as the heart of the circuit with an android predicated mobile phone. Here the mode of controlling inventions is by sending command wirelessly through Bluetooth [3].

IR remote control has an extensive application in the field of electronics. IR predicated remote control for controlling multiple home appliances with the microcontroller is additionally reported for the same function [4], [5]. Another approach is by GSM-predicated for home automation. This is done by sending a short SMS code from a mobile handset. Here it has a broader coverage area. So to control any household appliances from a distance place within the network area coverage sending a short SMS code will either ON or OFF the contrivances at home [6], [7]. N. Sriskanthan, F. Tan, A. Karande [8] exhibited show for home mechanization utilizing Bluetooth utilizing PC. This use of Bluetooth innovation in home computerization and systems administration condition. They propose a method, which contains a remote, versatile host controller and a few customer modules (home apparatuses). The customer modules speak with the host controller through Bluetooth devices. Moreover, Bhawna Saini proposed Smart LED Display Boards which is aimed at the colleges and universities for displaying a day to day information continuously or at regular intervals during the working hours [9]. Additionally, Most of the previous works are individual devices. But in our project, it is the combination of all other project and also customized countdown clock is shown where user can set any time value, which is not explained by any previous projects.

This paper represents a system, where the user can control various electrical appliances by IR unit. A real-time watch has been added to select the time for each candidate. If it is necessary to add extra time, it may be possible. The clock display shows the decremented count per second and min. If the minute reaches within three minutes, a warning signal comes through the market. If minutes and seconds are empty, the final warning sound comes up. Alert is consistent with the sound time, and this word appears with Buzzer. Another unit can be kept outside the room, where the candidate has his seat. This unit is also controlled by the IR remote which is placed in the hands of the examiner. 8x4 dot matrix displays the number of specific candidates, each change in the display, and the lightning must raise a sound. The external unit can be controlled by the IR remote, but it can communicate with the

clock display within RF communication. Here, an 8-bit microcontroller ATmega328P is used, and the program "ARDUINO" software is written [12]. In this way, they maintain the fearless schedule, proper sequence, and alarming protocol to conduct a viva or presentation. This will make their exam enjoyable and productive along with the examiner. So the primary motivation of this project is to develop a comparatively low cost and more reliable system to conduct viva or any presentation.

II. System Overview

The system mainly consists of two parts: the remote controller and two display units. The microcontroller of the remote control unit detects a press in a push button. If the time is pressed for management, the clock display unit is ready. Besides, with the external unit operating, it sends a specific 8-bit code for the RF transmitter for each push button. The microcontroller of the display unit receives the code sent through the RF receiver through the refreshing controller[11]and accordingly controls the display accordingly shown in Fig 1.

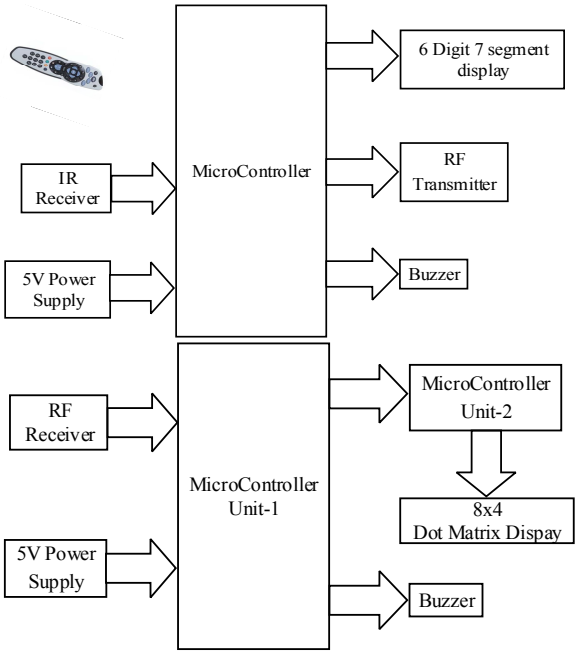


Fig. 1. Block diagram of the System.

III. System Desing

The methodology of developing the system is described in this section. The development process includes Interfacing 7 segment displays, receiving data from IR remote, Data transmission & receiving through the RF module, scrolling process in a dot matrix display and finally the complete circuit diagram of the system[9].

A. Interfacing 7 segment displays

At the very beginning, we have to convert BCD numbers into the 7-segment format and generate the respective decimal value for every BCD number. Then we create masking to separate the bits. After separate the bits, it sends to the appropriate output pin as the desired segments. The process of masking is the AND operation between two decimal codes, where we can separate each Binary value for the respective layers. For BCD value is 0000, the Decimal value is126 and for BCD 0001, Decimal value 48 are shown in "Table. 1".

TABLE I. CONVERTING BCD TO 7 SEGMENT FORMATS

Decimal Digit	BCD	g	f	e	d	c	b	a	Binary to Decimal
0	0000	0	1	1	1	1	1	1	126
1	0001	0	0	0	0	1	1	0	48

At first, we check the test value to light on the b segment only. So, we create AND operation between the Decimal value of 48 & 32. After AND operation, appropriate amount sends to the appropriate output pin as the desire segments are shown in "Table. 2". Then we need to select the display to show the desired value. In clock display unit we have 6 seven segment displays. When we need to show the value, these display selection line needs to high, and the rest of them goes to low. To show 1 in the first 7 segment display out of six 7 segments, data is a mask in the segment pin as 0110000 and selection pin S6 needs to high, others selection line goes to low at that time.

TABLE II. MASKING TO SEPARATE BITS

Segment	a	b	c	d	e	f	g
48	0	1	1	0	0	0	0
32	0	1	0	0	0	0	0
AND operation	0	1	0	0	0	0	0

B. Receiving data from IR remote

Here TV remote, an IR LED is utilized to transmit information to the contrivance. IR receiver in at clock exhibit picks out signals from remote through modulation. A mundane modulation scheme for IR communication is something called 38 kHz modulation. When we hit a key on our remote, the transmitting IR LED will blink very expeditiously for a fraction of a second, transmitting encoded data to the contrivances. The data word is reiterated as long as a key is pressed. In a data word which consists with 2 start bit, 1 toggle bit, 5 address bit, 6 data bits. In our remote panel utilize a space-coding method in which the length of the spaces between pulses of light represents a one or a zero which displays in Fig 2.

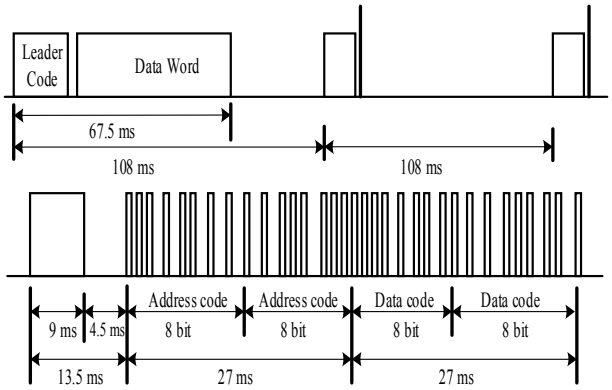


Fig. 2. Data format of IR remote.

C. *Data transmission & receiving through RF module*

Data transmission through the RF module, depends on the three units in Fig 3, which are Start bit, Clock bit and a Data bit. In Start bit occurs with every interruption. When an interrupt occurred, then the clock bit generated. Defines no of bit transferred for each code and synchronize the transmission between receiver and transmitter. In the time of the clock, a bit data bit can be sent to the receiver. The Status of each data bit d3d2d1d0. Same as Data transmission, Data receiving through RF module depends on the three units. When an interrupt occurs receiver unit start collecting data for a pulse. In clock bit, for the rising edge of the clock data is read in the data pin after that mask operation is used for separate the data value and sends the value to the specific data pins.

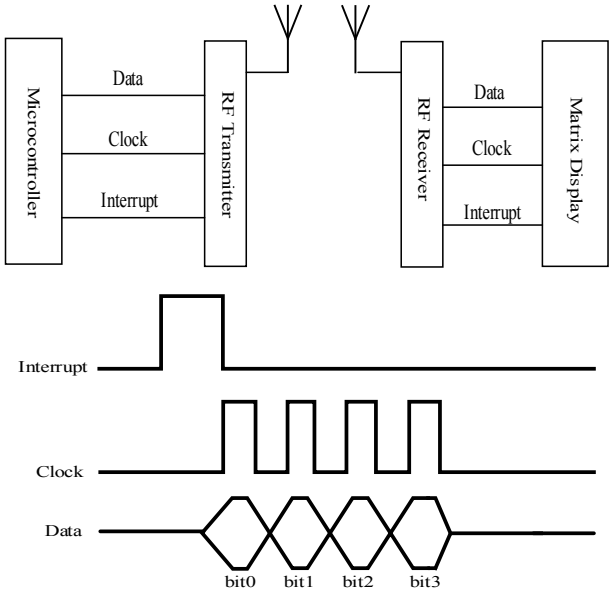


Fig. 3. Data transmission & receiving through RF.

D. *Scrolling process in dot matrix display*

We want to display the digit 1. We will first select the row as R1 data is required to scan the first row as 00011000 and deselect other rows by blocking their ground paths. Now, the first row is active. Next, select the row R2 and so on. Therefore, by scanning across the row quickly (> 100 times per second) and turning on the respective LEDs in each row of that column. In that way 64 bit required to show a character. For the scrolling process, we have used Max7219 display driver for left shifting. The persistence of vision comes in to play, and we perceive the display image as still. In Max7219, depends on the three units, Data IN, Clock and load. Data is required for the scanning process. The clock is used to receive the input data, and load pin requires loading the data to the device.

E. *Complete circuit diagram of the system*

The complete circuit diagram of the system is shown in Fig 4. Arduino Uno board is used here which has Atmega328 microcontroller. The OUT pin of IR receiver is connected with digital I/O pin 0 of Arduino Uno. The digital output pins 3-14 control the Clock display unit, digital output pins 10-13 controls the RF transmitter unit and finally digital output pin 2 controls the Buzzer unit. In receiving circuit, the OUT pin of RF receiver is connected with digital I/O pins 2-4 of Arduino Uno. The digital output pins 10-13 control cascade microcontroller unit and pin 8 controls the Buzzer unit. In cascade microcontroller used for controlling the Matrix display unit in digital I/O pins of 10-13. Logic HIGH in such a pin executes the specific operation. This is the principle to create a closed path in the load circuits: a single press in the reset button, both the display unit ready for the initial stages. The user can adjust the timer and calling the specific candidates by up and down button. When click the start button both the display units run perfectly.

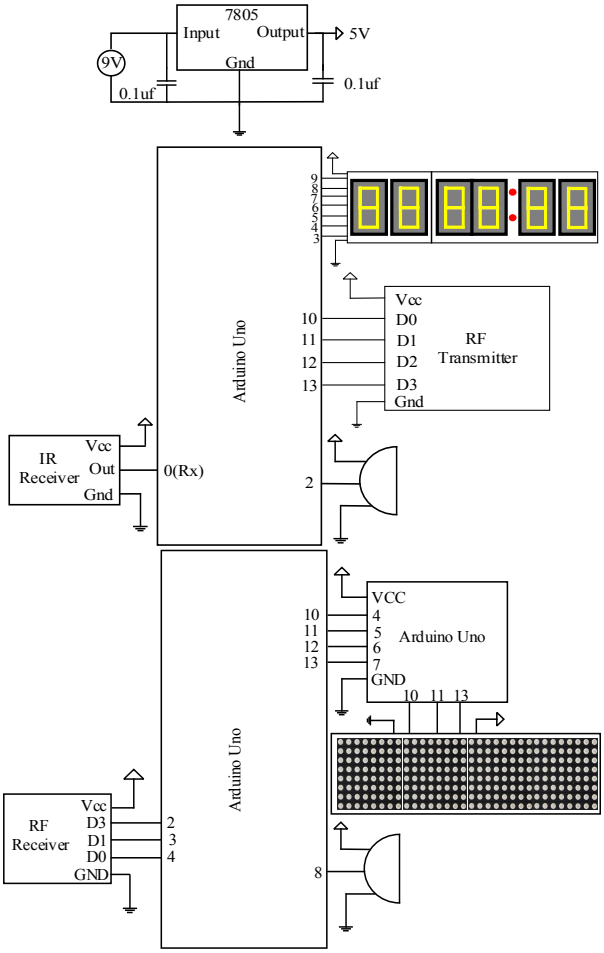


Fig. 4. Complete circuit diagram of the system.

F. *Flow Chart*

The flow chart of the whole system is shown in Fig. 5. it is divided into two parts. The first one is the sender and the second part is the receiver. In the transmitter part, after starting of the program data from IR receiver will be taken. If the data of the IR transmitter is 0x38863BE0, it active the clock system as 1min. Along with 0x38863BD0 as 2min and 0x38863BF0 as 3min. After that, it waits for starting the timer. If the data is 948321234, the countdown timer will start. When minutes & seconds close to zero or minute=3, the buzzer is active. If the data is 948321242, it will reset the clock display. If the data is 948321266 or 948321274, it transmitted increment or decrement logic state to the receiver part via RF communication. In receiver part Fig 6 and 7, after receiving the data from a transmitter, the system will be activated. If received data is greater than zero, the text of the group number is incremented than it will decrement the group number with the decremented logic state. In every new group

number, a buzzer will remain active and give the sound of the alarm.

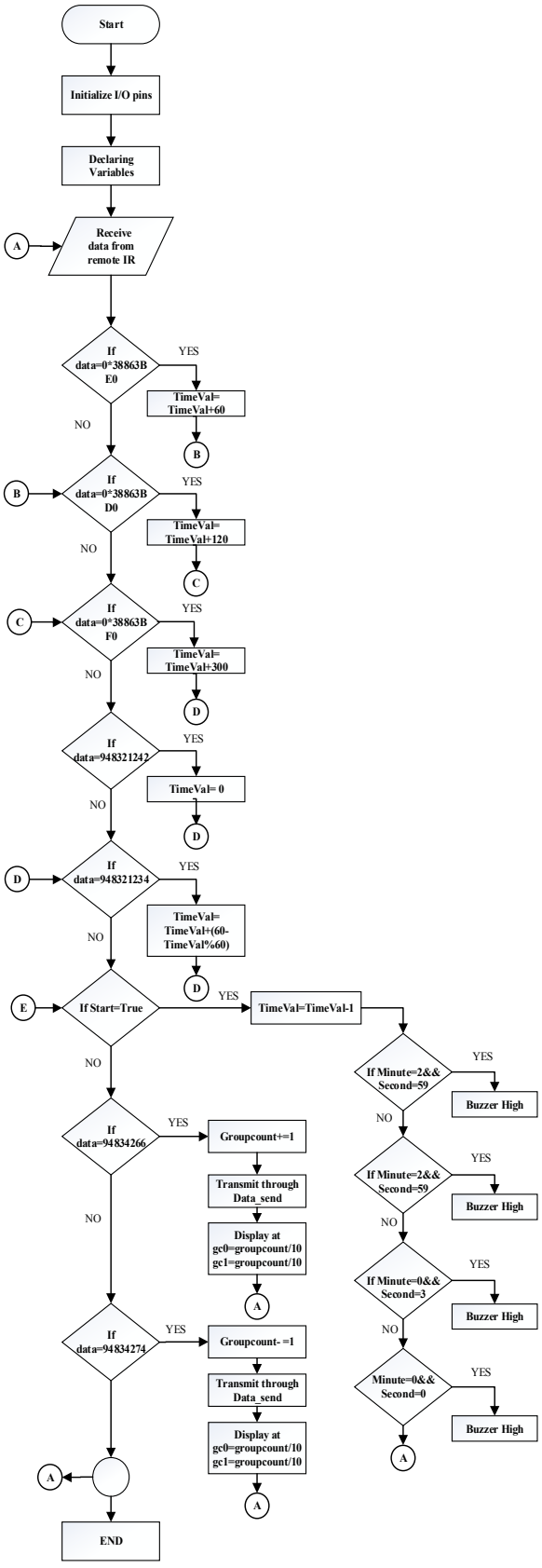


Fig. 5. Flow Chart of the system Transmitter.

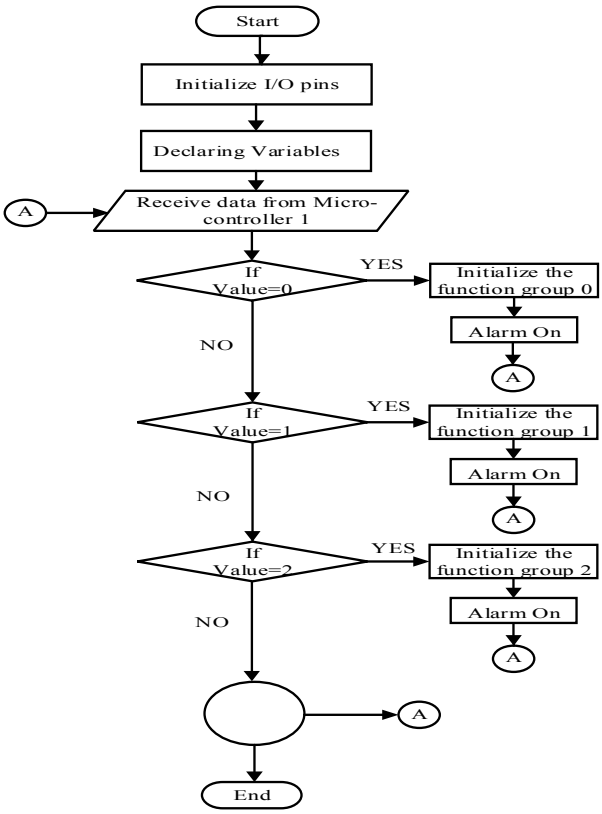


Fig. 6. Flow Chart of the system Receiver-1.

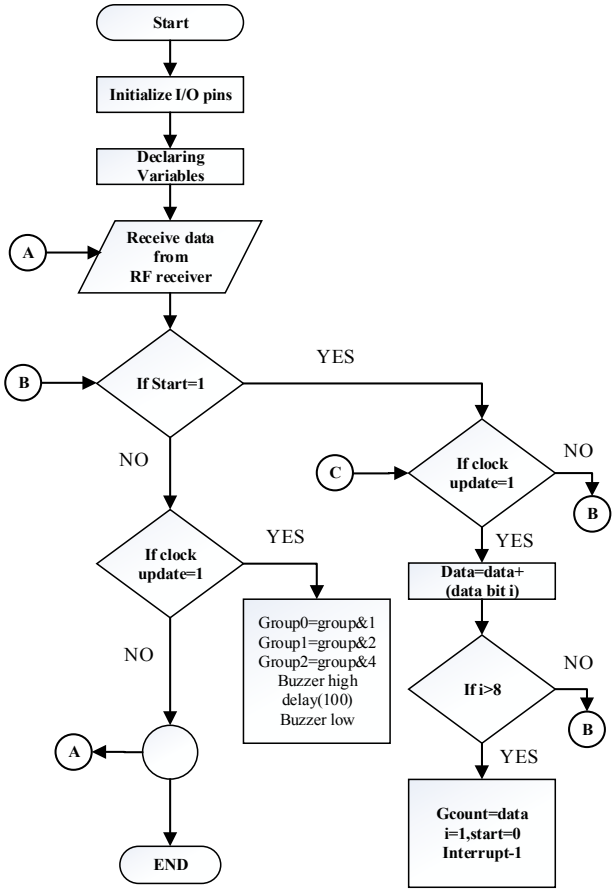
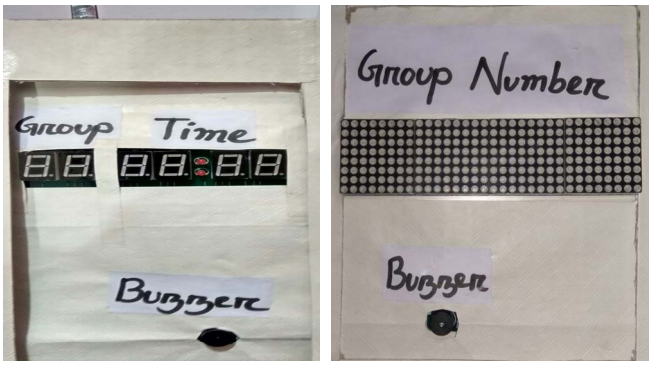


Fig. 7. Flow Chart of the system Receiver-2.

IV. Result and Implementation

The system is implemented as a prototype to check its eminence. Fig. 8, Fig. 9and Fig. 10, shows some images of the prototype. As the system is a wireless controlled system so it works for the IR & RF communication within a limited distance. The prototype of the system can work for the IR remote within 7m distance and the RF transmitter inside the room & receiver outside the room within 11-meter distance. The device can take power from a 9V battery providing current up to 1 amp or, it can be connected directly 220V supply through a 9V adaptor. In the prototype of the system, two display unit named as clock display and another one is group display unit. The buzzer is connected in each display unit. Initially, the system needs to reset. In this state, the reset command needs of the clock display and group display unit showing welcome text.



Indoor unit Outdoor unit

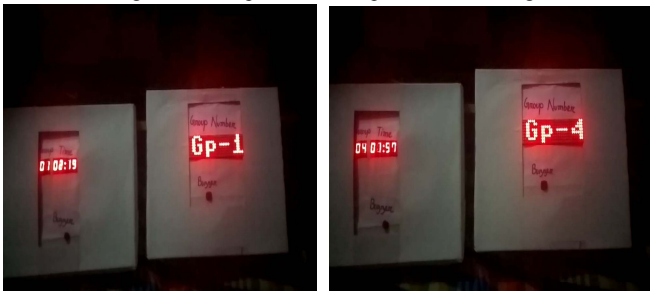
Fig. 8. Implemented prototype of system.



Group Changing Remote Panel

Time Managing Remote Panel

Fig. 9. Remote panel for Group and Time Settings.



Outside Display adjust with Indoor Group number

Fig. 10. Hardware implementation of the system.

V. Conclusion

The system has a unique advantage over other the conventional process. Controlling electronic appliances by RF or IR remote is exciting and enjoyable for all. Anyone can operate this device, and the total process executes in sequentially. Another advantage of this system is that user can manually adjust time and every time ends an alert sound reminds for the timing and also this is low power consumption electronic device. So there is no risk of short circuit or fire. If we replace any human through by our device, we can save more time in interview and presentation. Moreover, there have some limitation such as here start button is required to double press for starting the clock display and the maximum limit to operate the transmitter and receiver device is approximately 11 meter. Furthermore, Though RF-based communication has many desirable capabilities, it has some drawbacks too such as Antenna installation should be careful. Future work can be done adopting an extension of this other language besides English.

References

[1] M.A.Kader,M.M.Rahman,S.I.BHaider,& M. Islam, (2014, December). LED matrix based digital learning display for children with wireless control. In Computer and Information Technology (ICCIT), 2014 17th International Conference on (pp.397-400). IEEE.

[2] E. K. F.Chan,M. A. Othman, and M. A.Razak. "IoT Based Smart Classroom System." Journal of Telecommunication, Electronic and Computer Engineering (JTEC) 9, no. 3-9 (2017): 95-101.

[3] Y.G. Belgi,,P.G. Avatade,P.V. Deshmuk,A.M. Sakhare,A.J Shinde, & J.M. Patali, (2013). Android based appliances control system. International Journal of Emerging Technology and Advanced Engineering, 3(12), 681-683.

[4] A. F.Mitul,F. H. M.Rafi,M.M. Islam, &M. Ahmad(2012, December). Microcontroller Based Remote Control of Home Appliances. In Proceeding of the International Conference on Electrical, Computer and Telecommunication Engineering (pp. 01- 02).

[5] N.M. Santosh "Wireless Infrared Remote Controller for Multiple Home Appliances." International Journal of Electrical and Electronics Research 2.1 (2014): 25-35.

[6] A.Sindhuja, and B. Kiran Babu. "Remote Control of Electrical Appliance using Wireless Technology GSM." International Journal of Science and Research 2.4 (2013): 498-500.

[7] M. H. A.Wahab, N. Abdullah, A.Johari, & H.A.Kadir (2010). Gsm based electrical control system for smart home application. Journal of Convergence Information Technology, 5(1), 33-39.

[8] N.Sriskanthan, F.Tan, and A.Karande. "Bluetooth based home automation system." Microprocessors and microsystems 26, no. 6 (2002): 281-289.

[9] B. Saini, R. Devi, S. Dhankhar, M. Ziaul-Haque and J. Kaur, "Smart LED Display Boards", International Journal of Electronic and Electrical Engineering., vol. 7, no. 0974- 2174, pp. 1057-1067, 2014 (October 2016).

[10] M.A.Kader, M. Y.Haider, Karim, M. R., Islam, M. S., & Uddin,M. M. (2016, October). Design and implementation of a digital calling bell with door lock security system using fingerprint. In Innovations in Science, Engineering and Technology (ICISET), International Conference on (pp. 1-5). IEEE.

[11] N K Kaphungkui, “RF based Remote Control for Home Electrical Appliances”, International Journal Of Innovative Research In Electrical, Electronics, Instrumentation And Control Engineering,Vol. 3, Issue 7, July 2015.

[12] M. Caruso., D. Tommaso, A. O. Miceli, R. Galluzzo, G. R.Romano, P., Schettino, G. Viola (2015, October). Design and experimental characterization of a low-cost, real-time, wireless AC monitoring system based on ATmega 328P-PU microcontroller. In AEIT International Annual Conference (AEIT), 2015 (pp. 1-6). IEEE.

Smart Water System at House Using Arduino Uno and C# Desktop Application to Reduce Water Wastage and Energy Loss

Arif-Ul-Islam
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
a.rifislamm.cse@gmail.com

Shamim H Ripon
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
dshr@ewubd.edu

***Abstract*—Water is a vital element of nature. People would not think of a single day without it. Water becomes more important when there is scarcity of it. Mostly developing countries are facing this issue. Moreover, the people of these countries are not aware of wasting waters. Safety of water is one of the important solutions to this problem. A smart water management system is presented in this paper to remove the water wastage and energy loss at house. In this system, there are two motors. One motor is controlled by Arduino Uno to fill the water tank and another motor is managed by C# (c sharp) desktop application to distribute water to the users through an optimal way. To recognize water level, the research has not implemented any sensor. Instead, wires are used to sense the level. Water tank automatically starts to fill up with ground water by the first motor when water reaches down to a certain level, 2) the motor will be switched off and alarm goes on to notify the owner when the water gets up to a fixed upper level and 3) the owner gets notification that tank is full on the desktop and later he can switch on the 2nd motor to provide water to the users and to control usage of water through desktop application sitting in his home whenever users need. Moreover, application will show that on a single day how many times 2nd motor is set on. Thus this approach gives an automated and hassle-free model for the people in order to control the home water system and provides optimal water consumption.**

***Keywords*—Smart water system, Arduino, Automated system, Desktop application, Without sensor**

I. Introduction

Fresh water is generally used for drinking, household works, industrial uses, irrigation, and livestock and aquaculture etc. all over the world. These uses are increasing rapidly as the temperature changes and the population is growing. On the other hand, climate change puts a huge impact on water shortage. In this situation, water wastage can bring a biggest fall in the environment. It can be wasted on water supply system at home, plant-irrigation, industrial works etc. Many researches have been done to eliminate this problem. Authors of [1] proposed a system for saving water in plant irrigation. A system, designed for physical disable patients, had a good impact on saving water [2]. Another automated plant-irrigation system was proposed in [3].

This paper focuses only on the wastage of water at human houses. Today one of the problems is people have no idea when their water tanks gets filled up. Sometimes there are situations where the pump keeps on pumping water to the tank and the water starts spilling out from the tank. This makes wastage of energy as well as wastage of water. To reduce the wastage of water at house, the target of this paper is to build a smart and automated system where the owner of the house can monitor the system in his desktop application. The water tank will be automatically filled up with ground water. When the tank becomes full, the motor will be shut down and the owners get notified by the alarm as well as notification on his application. Later, he can provide the water when the users need by switching buttons on the application. Supplying water to the tank is occurred according to the water surface in the tank. The desktop application interface is easily operated. Detection of water is prime matter in this part. It can be carried out by sensors. But usage of sensor is expensive. To accomplish the identification of water level, wires are used. When wires get touches with these water, they send HIGH signal to the Arduino. This approach is less costly, user friendly, automated and fully controlled. It can be added to smart home automation system with a view to saving water and energy.

II. Related Works

Most researches on smart water system are proposed for water-irrigation and smart home technology. An IoT based automated system was discussed for water monitoring in [4]. By this model, user can monitor the water usage and wastage through mobile. M. Asadullah et al. presented a home automation model using sensors, Arduino, smartphone and bluetooth [5]. To track household water supply and billing, in [6] an automated model was showed. Their model had features of switching motor based on the water level and billing system. Arduino mega 2560 and sensors were used in this system. T. Mantoro et al. [7] proposed a saving water system for smart home bathtub according to the water level. They also used sensor for water level detection. In [8], authors

had a method to watch and estimate the consumption of the water through a web server. They used Hall Effect sensor for measuring water flow rate. Previous researches in this field rarely discussed about the water wastage at household tank (which supplies water to the users). It can make a huge loss of water and energy. Again most of the papers implemented sensors for detection of water and other related works which is quite costly. Therefore, there must be a system which is less expensive, automated and monitored by computer for ideal water consumption at house. Proposed approach is having these working features.

III. Proposed Model Description

The system has been implemented using Arduino Uno, DC motor, 100k resistors, wires, buzzer and computer. In this process, there are two motors, one motor is used for pulling water out of the ground and store water in the tank which is situated at the top of the house. Another motor is for providing water to the users. So this system consists of two parts.

- The tank is automatically filled by Arduino Uno based on the water level using the 1st motor
- Desktop application controls the water distribution through the 2nd motor

The system flow diagram is presented in figure 1.

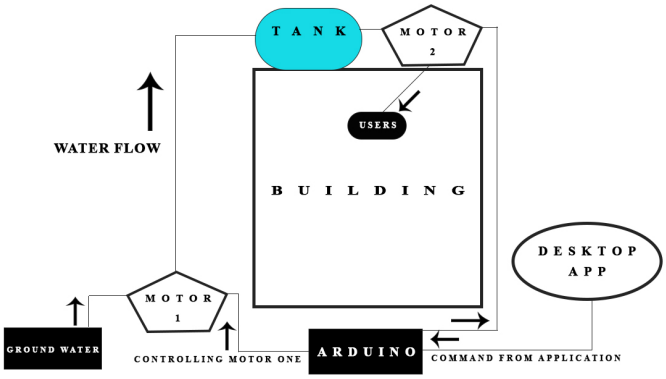


Fig. 1. Proposed System

A. A. Water level detection and motor 1 control

In this system, the identification of water level is done by setting parallel wires in the tank. Generally, ground water is impure and this kind of water is perfect for conducting electricity because it has some ions such as calcium and magnesium and these ions are good conductors. The idea is two wires are put in tank very closely. When water touches these two wires there generates some electricity and later we check the status of the wire in the Arduino. If status is HIGH, we get to know that water is now at that level. In this experiment, motor 1 works based on two water levels (upper and down). So, there must be four wires in total for level detection. Two wires in the top level and another two wires in the bottom level (Figure 2).

When the tank is set up for use, it is empty. That's why both down level and top level wires are LOW at that moment. Arduino Uno senses the signals from the wires and turns on the 1st motor to pull up the water from the ground to the tank (Figure 3).

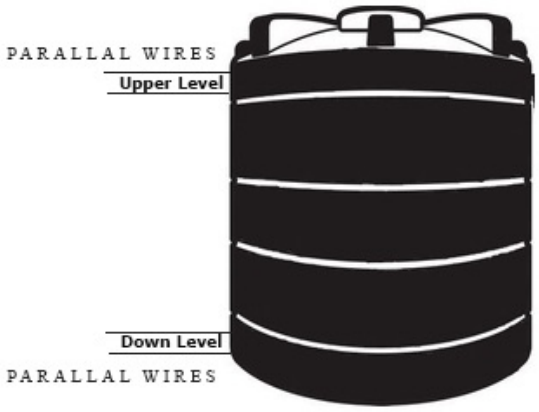


Fig. 2. Wires working as Sensor

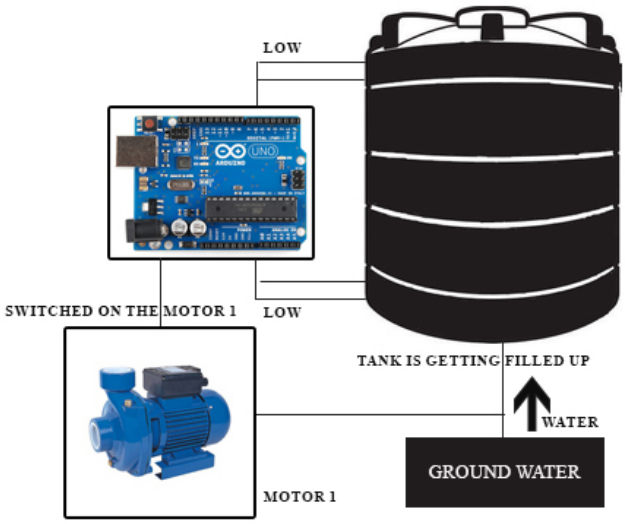


Fig. 3. First Phase- Tank is getting filled up with water

By this way, tank is going to be filled up with water by the motor 1 until the water touches the upper level. When the water surface is above the down line and below the upper line, down wires and upper wires will provide HIGH and LOW signals respectively (Figure 4).

After a fixed amount of time, tank will be full with water. During this time both upper and lower wires provide HIGH signal. While Arduino gets these signal, it will turn off the motor 1 (Figure 5). A buzzer which is connected to the Arduino for providing alarm to the house owner to indicate that there is sufficient water in the tank, goes on for few seconds.

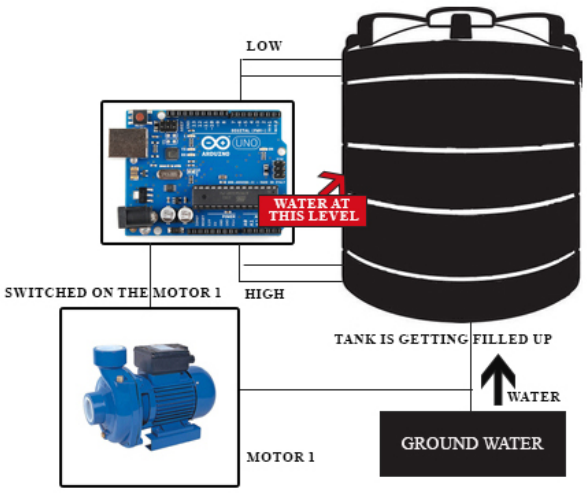


Fig. 4. Water surface between lower and upper level

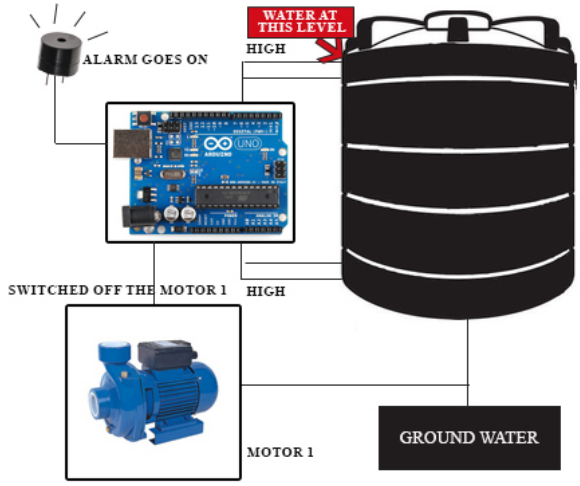


Fig. 5. Motor 1 is switched off and Alarm goes on

B. Desktop Application

To provide a hassle-free system for the house owner, this experiment has a desktop application. From the figure 5, we see the owner will be notified when the tank is full by the alarm. Later, he has ability to control the water distribution for the users. When he switches on the motor 2 through this application, the 2nd motor provides water to the building or the house (Figure 6). Desktop Application Interface is shown in figure 7. There are two buttons. These are for controlling 2nd motor. After consumption of water, when the tank becomes empty, upper and lower level wires will signal LOW. At this time, if there is no water and 2nd motor is still on, any damage can be happened. Also, suppose owner of the house unintentionally starts 2nd motor in this type of situation but Arduino let not him/her to do. To avoid any accidents, Arduino turns off the motor 2 and turns on the motor 1 whenever both pair of wires signals LOW. There rises a question, if Arduino shuts off the 2nd motor, then there should not be any OFF button in application. Suppose a situation, while there is enough water in the house and the 2nd motor is turned on, user continuously gets water but they need not it at that time. This can make water wastage. So to get rid of this problem, owner can turn off the 2nd motor by using OFF button during this situation. Also there is a counter of how many times 2nd motor is turned on a particular day. By this counter, owner can track the water usage. Full architecture of the proposed system is shown in figure 8.

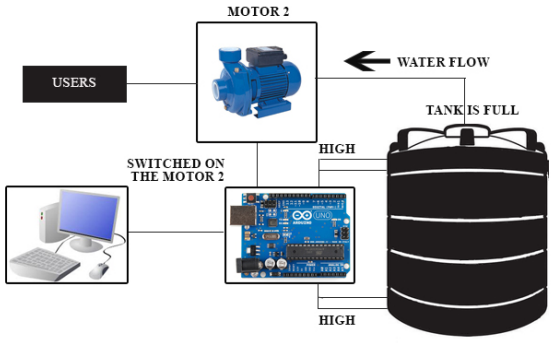


Fig. 6. Water distribution to the users by the Application

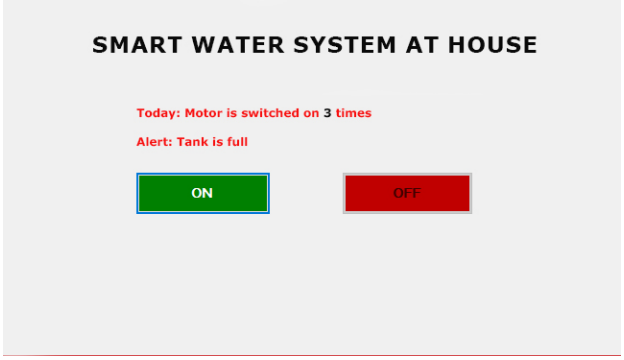


Fig. 7. Application Interface

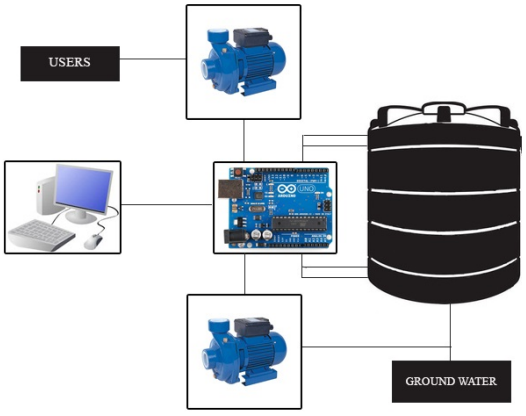


Fig. 8. Full Architecture of the Proposed System

IV. Hardware Model And Pseudocode

Figure 9 shows the implementation of hardware model.

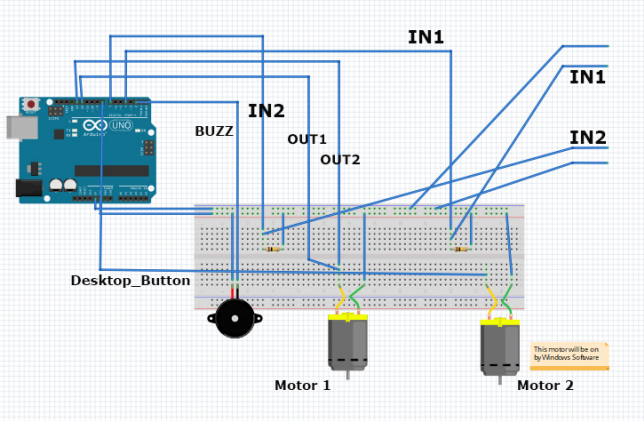


Fig. 9. Hardware Implementation

Algorithm 1 Algorithm

```
1:  procedure SmartWaterSystem
2:  loop:
3:      if (in1 == LOW and in2 == HIGH ) or (in1 == LOW and in2 == LOW ) then
4:          out1 ← HIGH
5:          out2 ← HIGH
6:          buzz ← LOW
7:      else if (in1 == LOW and in2 == LOW ) then
8:          out1 ← HIGH
9:          out2 ← HIGH
10:         buzz ← HIGH
11:         desktopButton ← LOW
12:     else if (in1 == HIGH and in2 == HIGH ) then
13:         out1 ← LOW
14:         out2 ← LOW
15:         buzz ← LOW
16:     if (seriaData == "O" ) then
17:         desktopButton ← HIGH
18:     else
19:         desktopButton ← LOW
20:     goto loop
```

In algorithm 1, the model seems to have a problem in a situation when the water surface reaches at the top level and IN1, IN2 gets high then motor1 is shut down. Later, while motor 2 is on, the water level gets below of the upper level. So based on the condition (on line 3) the motor 1 becomes on again. The motor 1 is going to be continuously on and off. This can destroy the system and any harmful damage can be occurred.

To avoid this issue, we can use FLAG variable. Algorithm 2 is the updated version of algorithm 1. In algorithm 2, at the beginning FLAG is declared to be false. While IN' and IN2 are LOW, then the tank is becoming filled up (on line 11) and

Algorithm 2 Revised Algorithm

```
1:  procedure SmartWaterSystem
2:      flag ← false
3:  loop:
4:      if (flag == true) then
5:          if (in1 == LOW and in2 == HIGH ) or (in1 == LOW and in2 == LOW ) then
6:              out1 ← HIGH
7:              out2 ← HIGH
8:              buzz ← LOW
9:          else
10:             flag ← false
11:     else if (in1 == LOW and in2 == LOW ) then
12:         out1 ← HIGH
13:         out2 ← HIGH
14:         buzz ← HIGH
15:         desktopButton ← LOW
16:         flag ← true
17:     else if (in1 == HIGH and in2 == HIGH ) then
18:         out1 ← LOW
19:         out2 ← LOW
20:         buzz ← LOW
21:     if (seriaData == "O" ) then
22:         desktopButton ← HIGH
23:     else
24:         desktopButton ← LOW
25:     goto loop
```

FLAG becomes true. As FLAG is true, tank is preceded to be filled up. Motor 1 gets stopped after IN1 and IN2 both are HIGH (on line 17). As flag is true and IN1 and IN2 are HIGH, the condition on line 9 becomes true, the FLAG is set to be false. Next time, the owner turns on the motor 1 but as FLAG is false the motor 1 will not be switched on until the water reaches down to the lower level. For controlling 2nd motor Arduino gets two input from serial port by desktop application. When "ON" button is clicked on the application Arduino gets "O" from serial port and when "OFF" is pressed, it gets "F" as input. Based on this input from serial port, Arduino change the status of the 2nd motor as follows-

- When Arduino gets "O", it turns on 2nd motor by setting desktopButon port as HIGH and
- Shuts down the motor by setting desktopButon port as LOW When it gets "F"

V. Conclusion

Countries such as Bangladesh, India etc. are having lack of water for their people. And general people of these countries are not concern about water wastage at household. They need a smart, automated and easy approach for operating household water system to reduce water loss as well as energy loss. Proposed system brings up this kind of solution and does not use any sensors which brings down cost. This system is effective for people of developing countries.

REFERENCES

[1] Zhang Feng, "Research on water-saving irrigation automatic control system based on internet of things," 2011 International Conference on Electric Information and Control Engineering, Wuhan, 2011, pp. 2541-2544. doi: 10.1109/ICEICE.2011.5778297

[2] H. Rashid, S. B. Osman, N. Hassan, I. U. Ahmed, R. Das and M. M. Karim, "A new design approach of home automation system for patients with physical disability to reduce water wastage and power consumption using renewable energy," 2017 4th International Conference on Advances in Electrical Engineering (ICAEE), Dhaka, 2017, pp. 770-774. doi: 10.1109/ICAEE.2017.8255458

[3] C. M. Devika, K. Bose and S. Vijayalekshmy, "Automatic plant irrigation system using Arduino," 2017 IEEE International Conference on Circuits and Systems (ICCS), Thiruvananthapuram, 2017, pp. 384-387. doi: 10.1109/ICCS1.2017.8326027

[4] S. Saseendran and V. Nithya, "Automated water usage monitoring system," 2016 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, 2016, pp. 0099-0103. doi: 10.1109/ICCSP.2016.7754501

[5] M. Asadullah and K. Ullah, "Smart home automation system using Bluetooth technology," 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), Karachi, 2017, pp. 1-6. doi: 10.1109/ICIEECT.2017.7916544

[6] T. Rahman, T. Ahmed, I. Hasan and M. A. Alam, "Automated household water supply monitoring & billing system," 2018 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, 2018, pp. 448-455. doi: 10.1109/ICISC.2018.8399113

[7] T. Mantoro and W. Istiono, "Saving water with water level detection in a smart home bathtub using ultrasonic sensor and Fuzzy logic," 2017 Second International Conference on Informatics and Computing (ICIC), Jayapura, 2017, pp. 1-5. doi: 10.1109/IAC.2017.8280602

[8] G. Gosavi, G. Gawde and G. Gosavi, "Smart water flow monitoring and forecasting system," 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, 2017, pp. 1218-1222. doi: 10.1109/RTEICT.2017.8256792

Towards Developing an Intelligent Automated Water Pumping System for Dhaka City

Kazi Md. Munim[1], Iyolita Islam[2], Moumita Sarker[3], Muhammad Nazrul Islam[4]
Department of Computer Science and Engineering
Military Institute of Science and Technology, Dhaka, Bangladesh
Email: [1]kazimunim2014@gmail.com,[2]iyolita2011@gmail.com,[3]moumita.sarker123456@gmail.com,[4]nazrulturku@gmail.com

***Abstract*— Cities where water supply system follows trivial analog system like Dhaka city, a twenty-four hours monitoring is required to start the water pump immediately when water is available in the main water supply line. A significant amount of water can be wasted only for the reason of overflowing the water tank. Thus, the aim of this paper is to propose an automated water pumping system to minimize the wastage of time, money, labor and water loss. A high-fidelity prototype system was designed and developed using the Arduino Uno, water pump, potentiometer and relay switch; and able to control its operation automatically with respect to the availability and amount of water in the main water supply line and the reserve water. An evaluation study was also conducted to assess the availability and reliability of the system. The evaluation results showed that the proposed system is reliable, available and cost-effective.**

Keywords—Artificial intelligence, water pump, water supply management, intelligent system.

I. Introduction

Dhaka city, the capital of Bangladesh, is one of the most densely populated areas within the country with around 13 million people. So, the people in Dhaka city require a huge quantity of water for different purposes everyday. Moreover, there is a large number of industries along with other organizations which require water supply for their operation. According to Dhaka Water Supply and Sewerage Authority (DWASA), considering water demand as 20 l/p/d (litre per person per day), total water demand in Dhaka city varies from 2300 to 2350 million litre per day (MLD), while the total production capacity of DWASA is 2450 MLD including both ground and surface water [1]. The DWASA is thus able to fulfill the current water demand. However, since the water level is decreasing day by day. So, it is obvious that water supply will not be sufficient against the demand and people will not get water on the water line for the whole day.

Again, the issue of ensuring enough water supply in Dhaka requires adoption of the most effective, efficient, relevant and appropriate systems which must be cost effective as well. As water is not available all the time in the main line, people always need to be concerned about the reserved water level so that the pump can be switched on as soon as the water is available. Moreover, people have to be conscious when the water tank is full and the pump needs to be switched off. Otherwise, a huge quantity of water can be wasted by overflowing the water tank. To perform all these, it is not easy for a person or an employee (which may cost) to check the availability of the water throughout the whole day.

Therefore, the aim of this paper is to propose and develop an intelligent cost-effective automated water pumping system. To attain this research objective, a prototype of the automated water pumping system was developed and an evaluation study was conducted to measure its performance.

The rest of the article is organized as follows: Section II presents related works, Section III discusses the implementation of the system, the evaluation of the system is discussed in Section IV and finally Section V contains conclusion and future expansion.

II. Literature Review

This section will briefly introduce related works focusing to the design and development of the water pumping and management systems.

Astrom [2] invented a water pumping apparatus which was basically used to develop a water pumping system and was suitable for home and farm use. On the other hand, Eusuff and Lansey [3] proposed a system named SFLANET which actually linked Shuffled Frog Leaping Algorithm and the water distribution system to develop a better water supply system that offered a better water supply system for the citizens. But the improved water distribution network had a very high computational complexity.

Oi [4] introduced Photovoltaic water pumping system using Photovoltaic modules and DC motor by simulating the model in MATLAB. Abdolzadeh [5] worked to improve the effectiveness of the system. He also evaluated the functionality of his system. Hsiao [6] coupled among a photovoltaic power source, a solar panel and a pumping system. Later, Kohle [7] analyzed the performance of Directly Coupled Photovoltaic Water-Pumping System.

Meierhoefer [8] proposed a model which used a device that boosted hydraulic pressure as well as the in-line water pressure. The model could take necessary steps by sensing the current condition of water demand and water discharge line. An additional pump was used here to support the pump pressure. Again, Ormsbee [9] examined different models including hydraulic network model, demand forecast model and optimal control models for both single and multi-source water-supply pumping system. He listed the important factors for optimization such as- operating cost, physical system

limitations, governing physical law, externally defined requirements, the maintenance cost of the pumping system etc.

A number of researches has been carried out to develop and design an optimized model of water supply system. Crawley [10] developed a multi-reservoir system with an experimental model which was basically a monthly optimized planning model with minimal system reliability. He also developed an experimental model to ensure that the model reduced 5% to 10% of pumping cost. Later, the system was adopted by the local authority named Engineering and Water Supply Department (EWSD). Similarly, some other researchers [11]–[13] also adopted the multi-reservoir concept in developing the water pumping system. Yu [14] focused to the pumping schedules as a solution to water distribution system. He also provided test example to validate his research work.

Jowitt [15] introduced a optimization procedure to determine the structure of optimal pump scheduling process or electricity tariff with available pumps. Basically, it worked by considering the unit and maximum electricity charges, efficiency of relative existing pumps, demand of the consumers and operational constraints of the network. Nitivattananon [16] developed another model for optimization of water supply system operation to generate the pump schedule in real time operation including monthly billing. It considered the major difficulties of a complicated tariff, discrete discharging of pumps, charging of demands etc. A simulation for a test showed that 20% of the actual cost could have been saved applying this model. Coulbeck [17] proposed a procedure to optimize the unit costs and demand charges, and to derive the maximum flow requirements for the tariff period.

McCormick [18] acquainted a stochastic dynamic program-based system to optimize the water-supply system. He showed that the scheduling for the pump switching on/off can be sometimes complicated when the maximum demand charges which can't be ignored, are involved in electricity tariff. In his paper, he considered some medium term maximum demand policies and some constraints in daily pump scheduling. He also proved the method better than using heuristic techniques because this model considered the demand for water is not settled. Shu [19] worked to optimize the power consumption in water pumping system by using hybrid genetic algorithm. His optimal controlled model was applicable for wide range of water supply system which could calculate the lowest cost after preserving the performance for each station. He was able to reduce power consumption by 6.04% than before which also reduced the pumping cost. Beckwith [20] contributed in Artificial Intelligence field by choosing a Genetic Algorithm approach of natural selection with a fitness function, to optimize the cost for the operation of a multi-storage, multi-source water supply system. He developed a genetic algorithm to ensure that the pumps in the system are adequately providing the required volume/amount of water in the water distribution system with minimizing the cost for the usage of electric energy by the pump/motors. Moreover, the algorithm considered some factors like- the efficiency curves and the flow limits of pumps in the system.

Reza [21] showed a microprocessor-based process to control the pumping schedule which senses empty or full tank by controlling on/ off signal. The system included PIC with sensors to sense the water level with LED lights as indicator. Stene [22] constructed a model for hot water supply system and evaluated the system to measure the performance for some specific operating conditions, while Zongxuez [23] carried out a simulation based study to measure the performance of the water supply system in Fukoka city.

The literature review showed that though a few works have been carried out related to the water pumping system, but no work explicitly focused in the context of Bangladesh. In Bangladesh, especially in Dhaka city cost is not only the crucial factor but also the availability of water (i.e., preserving the water for future use) is a key concern. Thus, this research aimed to develop a smart water pumping system to be used in context of Dhaka.

III. Implementation of the System

A. Equipment

To develop the prototype of the proposed automated water supply system following hardware were used:

1) Arduino Uno: Arduino is an electronic platform to control according to the input and give output by hardware and software programmed for it. This device is cost-effective, easily accessible and used in developing a number of intelligent and assertive systems [24]–[27].
2) Water Pump: Water pump is an electric machine that turns electrical energy to mechanical energy.
3) Potentiometer: Potentiometer is a resistor of three terminals which can adjust the voltage drop. Potentiometer is usually used for controlling any electrical device.

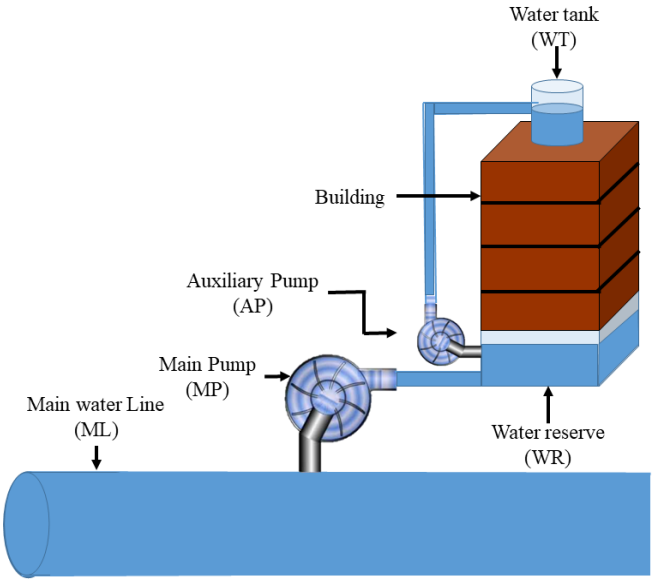


Fig. 1: Overview of water pumping system

4) Relay switch: Relay is a magnetic switch which is operated electrically and used for energising or accelerating the electrical signal to required state.
5) Two containers as upper water tank and reserve were used as well.

B. Prototype

A high-fidelity prototype of the proposed system was developed as showed in Figure 1. In this proposed model, Main Pump (MP) was placed with the main line (ML) to get water from the main line to the building. Water Reserve (WR) was placed under the building and the upper Water Tank (WT) was at the top of the building. When water was available inside ML, the system sensed it and automatically started the MP (if the WR was empty or under certain amount according to the expected amount). If WR was full and WT was empty or under certain amount, the system automatically switched on the Auxiliary Pump (AP). The starting of MP and AP depended on certain condition or amount of water in percentage. Thus, the system automatically worked to provide water in maximal beneficial and minimal costly way.

C. Working Principle

The working principle of this proposed water control model is quite simple and can be described as follows-

- Start the main pump when water is available in the ML and continue pumping until the WR is full.
- If 80% water is present in WR and MP is still pumping water and the WT is not full, the AP will start filling WT until it is full or MP is off.
- If water in WT is less or equal than 10% and if the water in WR is equal or more than 40%, AP will start automatically and fill WT up to 30%.
- If the WR is full, the MP will stop automatically.

The system can be depicted more clearly by the flowchart shown in Fig. 2. Decisions were taken at different phases of the system to start or stop pumping water automatically.

The percentage of the amount of water was measured by the concept of electrical circuit. In different levels of the upper or reserve tank, the wires of the different percentage labeled were clipped or attached with a conductor. If water reached any level then it makes a closed circuit so that

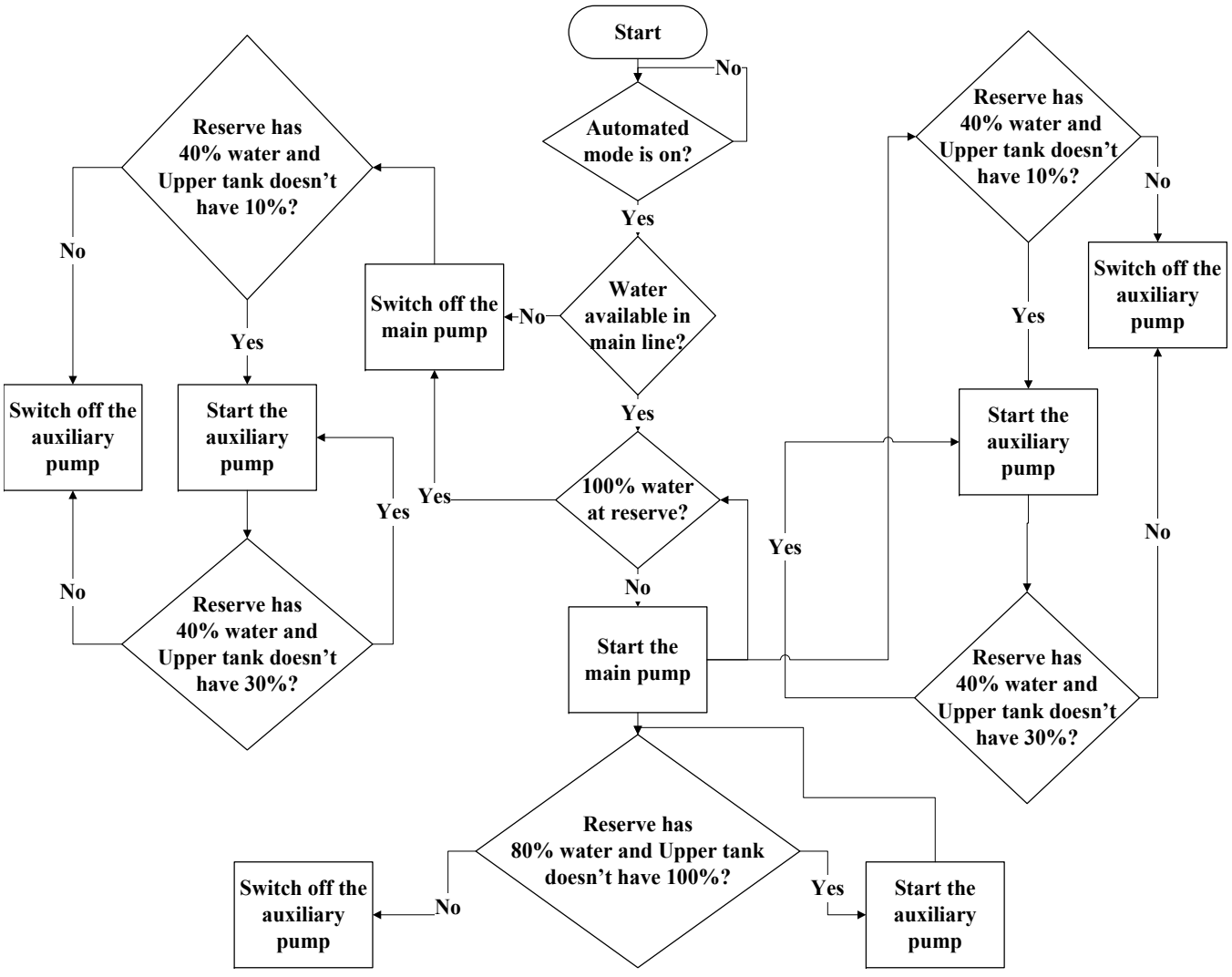


Fig. 2: Flowchart of automated water pumping system

Table I
THE OUTCOMES OF THE EVALUATION STUDY

Condition	Action	No. of Trials	No. of Success	Success Rate	Avg. Delay (second)
Water is available in ML	START MP	20	19	95%	0.9155
80% water in WR and ML is ON and WT is NOT FULL	START AP until the WT is FULL or MP is OFF	20	18	90%	1.141
Water in WT <= 10% and water in WR >= 40%	START AP to fill up WT up to 30%	20	16	80%	1.2605
WR is FULL	STOP MP	20	19	95%	0.8225

electricity can flow. The potentiometer used in the system, basically converted the water level to voltage drop. If there was no water in the WT, potentiometer showed 100% voltage drop for open circuit. If water rose up to certain level, the potentiometer showed certain voltage drop for closed circuit. From the reading of the potentiometer, the level or amount of water could be measured. A program was developed and loaded in Ardunio to analyse the all input data in order to provide decision of the switching on and off the motor.The main circuit implementation to develop the prototype is presented in Figure 3.

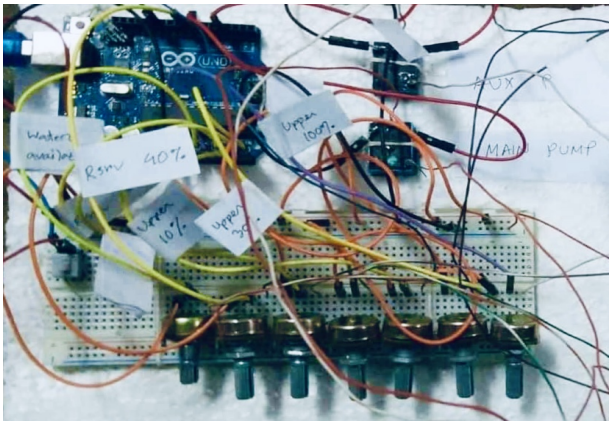


Fig. 3: Hardware components with circuits to develop the prototype

IV. EVALUATION OF THE SYSTEM

A. Evaluation Environment

The evaluation was conducted in lab environment. Two plastic containers were used: one as the upper water tank, WT and another as the water reserve, WR as showed in Figure 1. According to the working principle of the system, the system had to take actions depending on the water level in the both tanks i.e.: when water in WT is <= 10% and water in WR was >= 40%, then the AP should be started to fill up WT up to 30%. So, the container which was used as WR, was marked as 100%, 80%, 40% and 0% so that the working procedure could be easily monitored. For the WT container, there were traces to point 100%, 30%, 10% and 0% water height as well. The rest of the components were connected as they were demonstrated in Section III.

B. Evaluation Procedure

The system was initiated and started to collect data for the cases which have been described as working principle. A total of 20 trials were considered for each case and the response of the system was monitored which it was supposed to generate. From the responses, the number of success trials and delay for each trials were assembled separately. From the acquired values, the success rate percentage and average delay for each case were calculated. All the extracted data were compiled in Table I. Finally, total average delay of 1.0349 seconds and total success rate of 90% were determined for any case.

C. Data analysis and Findings

The study data was synthesized and analyzed to assess the system's *reliability* and *availability*.

a) Reliability- the results showed that the system responded well for different cases (see Table I). For the specific trials the delay may vary, but the average delay was not more than 1.26 seconds for each case. The system showed a total average delay of 1.0349 seconds for any response. The percentage of success rate to execute each condition varied from 80% to 95%. Thus, the system showed good performance in terms of the reliability.

b) Availability - the evaluation results showed that the system was consistently active (available) for the entire duration of conducting the experiment (see Table I). Though the system responded late for some cases, it was not switched off automatically in any case. In case, if there was no water usage for a long time, the system was still switched on and whenever there was any indication of water usage and WT needed to be filled up, the system started automatically to pump water to WT.

Moreover, a traditional water pumping system generally includes some common components i.e: water tank, water

Table II
APPROXIMATE COST TO DEVELOP THE SYSTEM

Equipment Name	Quantity	Cost in USD	
		Prototype System	Concrete Implementation
Arduino Uno	1	6	6
Water Pump	1	6	70
Potentiometer	7	2	2
Relay Switch	2	1	1
Circuit Cover Box	1	-	5
	Total Cost	14$	85$

pump and water passing line (pipe). Thus, except these items, the tentative cost to develop the proposed prototype system and the cost of concrete implementation has been presented in Table II.

This shows that only $85 (except the cost of water tank, main water pump and water passing line) will be required to develop the intelligent system while the prototype system required only $14. These cost data indicates that the proposed system is cost-effective with respect to its expected benefits regarding time, labor and water loss in the capital (Dhaka city) of Bangladesh.

V. Conclusion

Supplying adequate water is one of the crucial challenges to meet in major cities like Dhaka, in modern times. The major challenge of this model was to develop an intelligent cost-effective system so that it can be useful to the citizens of Dhaka. Thus, this article presented the design and development of an automated water pumping system. A set of logical conditions were integrated to develop the system so that the proposed water pump can work intelligently. The construction cost of the proposed system is very low and it requires a small amount of electricity to run the system. The system works automatically according to the logic applied in it and so no manual monitoring is required. The evaluation study and expenditure data showed that the system is reliable, available and cost-effective. In case of the implementation in real life, considering different aspects of using the proposed system like- per unit and maximum charges of water and electricity, efficiency of this system against the existing systems, electricity tariff etc., this simple system will be able to reduce the actual cost and effort. Future work may focus to implement the concrete version of the prototype model and then comparing this system with the existing systems.

References

[1] "Annual Report 2015-16, Dhaka Water Supply and Swerage Authority — https://dwasa.org.bd/annual-reports/."

[2] J. I. Astrom, "Water-pumping system," Apr. 11 1922, uS Patent 1,412,596.

[3] M. M. Eusuff and K. E. Lansey, "Optimization of water distribution network design using the shuffled frog leaping algorithm," *Journal of Water Resources planning and management*, vol. 129, no. 3, pp. 210–225, 2003.

[4] A. Oi, "Design and simulation of photovoltaic water pumping system," *California Polytechnic State University*, 2005.

[5] M. Abdolzadeh and M. Ameri, "Improving the effectiveness of a photovoltaic water pumping system by spraying water over the front of photovoltaic cells," *Renewable Energy*, vol. 34, no. 1, pp. 91–96, 2009.

[6] Y. R. Hsiao and B. A. Blevins, "Direct coupling of photovoltaic power source to water pumping system," *Solar Energy*, vol. 32, no. 4, pp. 489–498, 1984.

[7] M. Kolhe, J. Joshi, and D. Kothari, "Performance analysis of a directly coupled photovoltaic water-pumping system," *IEEE Transactions on Energy Conversion*, vol. 19, no. 3, pp. 613–618, 2004.

[8] N. S. Meierhoefer, "Water pumping system," Jul. 23 1996, uS Patent 5,538,396.

[9] L. E. Ormsbee and K. E. Lansey, "Optimal control of water supply pumping systems," *Journal of Water Resources Planning and Management*, vol. 120, no. 2, pp. 237–252, 1994.

[10] P. D. Crawley and G. C. Dandy, "Optimal operation of multiple-reservoir system," *Journal of Water Resources Planning and Management*, vol. 119, no. 1, pp. 1–17, 1993.

[11] Q. W. Martin, "Optimal operation of multiple reservoir systems," *Journal of Water Resources Planning and Management*, vol. 109, no. 1, pp. 58–74, 1983.

[12] Q. Zhong and K. E. Lansey, "Optimal operation of multiple reservoir systems," in *Optimizing the Resources for Water Management-Proceedings of the ASCE 17th Annual National Conference*. Publ by ASCE, 1990.

[13] J. W. Labadie, "Optimal operation of multireservoir systems: state-of-the-art review," *Journal of water resources planning and management*, vol. 130, no. 2, pp. 93–111, 2004.

[14] G. Yu, R. Powell, and M. Sterling, "Optimized pump scheduling in water distribution systems," *Journal of optimization theory and applications*, vol. 83, no. 3, pp. 463–488, 1994.

[15] P. W. Jowitt and G. Germanopoulos, "Optimal pump scheduling in water-supply networks," *Journal of Water Resources Planning and Management*, vol. 118, no. 4, pp. 406–422, 1992.

[16] V. Nitivattananon, E. C. Sadowski, and R. G. Quimpo, "Optimization of water supply system operation," *Journal of Water Resources Planning and Management*, vol. 122, no. 5, pp. 374–384, 1996.

[17] B. Coulbeck, M. Brdyś, and C. Orr, "Exact incorporation of maximum demand charges in optimal scheduling of water distribution systems," *IFAC Proceedings Volumes*, vol. 20, no. 5, pp. 359–364, 1987.

[18] G. McCormick and R. Powell, "Optimal pump scheduling in water supply systems with maximum demand charges," *Journal of water resources planning and management*, vol. 129, no. 5, pp. 372–379, 2003.

[19] S. Shu, D. Zhang, S. Liu, M. Zhao, Y. Yuan, and H. Zhao, "Power saving in water supply system with pump operation optimization," in *2010 Asia-Pacific Power and Energy Engineering Conference*. IEEE, 2010, pp. 1–4.

[20] S. Beckwith and K. Wong, "A genetic algorithm approach for electric pump scheduling in water supply systems," in *Evolutionary Computation, 1995., IEEE International Conference on*, vol. 1. IEEE, 1995, p. 21.

[21] S. K. Reza, S. A. M. Tariq, and S. M. Reza, "Microcontroller based automated water level sensing and controlling: design and implementation issue," in *Proceedings of the World Congress on Engineering and Computer Science*, vol. 1, 2010, pp. 20–22.

[22] J. Stene, "Residential co2 heat pump system for combined space heating and hot water heating," *International Journal of Refrigeration*, vol. 28, no. 8, pp. 1259–1265, 2005.

[23] X. Zongxue, K. Jinno, A. Kawamura, S. Takesaki, and K. Ito, "Performance risk analysis for fukuoka water supply system," *Water Resources Management*, vol. 12, no. 1, pp. 13–30, 1998.

[24] S. Z. Diya, R. A. Proma, M. N. Islam, T. T. Anannya, A. A. Mamun, R. Arefeen, S. A. Mamun, I. I. Rahman, and M. F. Rabbi, "Developing an intelligent waste sorting system with robotic arm: A step towards green environment," in *2018 International Conference on Innovation in Engineering and Technology (ICIET)*, 2018, pp. 1–6.

[25] M. N. Islam, M. R. Islam, S. M. R. Islam, S. A. Bhuyan, and F. Hasib, "Locsoldiers: Towards developing an emergency troops locating system in military operations," in *4th International Conference on Electrical Engineering and Information Communication Technology 2018 (iCEEiCT)*, 2018, pp. 264–267.

[26] N. S. Khan, S. Kundu, S. Al Ahsan, M. Sarker, and M. N. Islam, "An assistive system of walking for visually impaired," in *International Conference on Computer, Communication, Chemical, Material and Electronic Engineering 2018 (IC4ME2)*, 2018, pp. 1–4.

[27] M. Islam, M. Jahan, A. Ali, S. Rony, T. T. Anannya, F. I. Aziz, M. Bayzed, A. Yeazdani, and M. F. Rabbi, "Design and development of an intelligent seed germination system based on iot," in *Advances in Information and Communication Technologies for Adapting Agriculture to Climate Change II. AACC 2018*. Springer, Cham, 2018.

Bluetooth Network Based Remote Controlled Home Automation System

Md. Ismail Husain[1], Mahfuz Alam[2], Md. Golam Rashed[3], Md. Emdadul Haque[4], Mirza A. F. M. Rashidul Hasan[5], Dipankar Das[6]
Department of Information and Communication Engineering, Facualty of Engineering,
University of Rajshahi, Rajshahi-6205, Bangladesh.
Email: {[1]ismailhusainruice, [2]mahfuzalam5611}@gmail.com, { [3]golamrashed, [4]haque-ice, [5]mirza_ice, [6]dipankar }@ru.ac.bd

***Abstract*— With the advancement of wireless technology, wireless based automation is playing an important role in human life. Home automation is a house or living environment that contains the technology to allow devices and systems to be controlled automatically. The remote control is useful to keep home comfortable for general people and to support the elderly and physically impaired people. In this paper, we propose a system that will be able to control household appliances remotely. The system provides ON/OFF voice response via the speaker according to the user command on the appliances. We divide the overall system into two modules. One is a transmitter or Bluetooth remote and another one is receiver or relay module. Bluetooth remote consists of several push buttons to control the household appliances. The remote and relay modules are configured as master-slave to communicate between the modules. To produce the voice response, the recorded voices are store in micro-SD card. With the voice response, the user can easily understand which household appliance is controlling at a particular time. We implement the system to validate its applicability. Additionally, we compare and analyze statistically the general people's impression of the proposed system over a traditional system.**

Keywords*— *Home Automation, Bluetooth Network, Remote Control, Relay Module, Voice Response, Home Appliances.

I. INTRODUCTION

We are living in an era of advanced technology, where every part of our daily lives is related to the science of craft in one way or another. Housing and lifestyle have been impacted by modern technology such as automated door locks, security cameras and lighting control, which makes our life much easier, organized and safer. Bluetooth technology is one of the most popular wireless technology used to control the household devices wirelessly. Nowadays, home automation is one of the major applications of Bluetooth technology. Operating over globally available unlicensed frequency (2.4GHz), it can link digital devices within a range of 10m to 100m at the speed of up to 3 Mbps depending on the Bluetooth device class [1].

Nowadays human beings are addicted to use modern equipment specially in home environment. Conventional wall mounted switches located in different parts of the house makes it uncomfortable for the user to go near them to operate. It becomes more difficult for the elderly or physically impaired people to use it. The home automation system is used to control the indoor and outdoor lights, fans, ventilation, air conditioning, electrical and electronic appliances and so on using a control systems. Various types of home automation systems like Bluetooth controlled [2-7], internet controlled [8-12], Android-based [13-16], Infrared (IR) remote controlled [17] are available at present. It helps people to reduce work and better household management.

Bluetooth and Android-based home automation system consists of a receiver end at which Arduino board and Bluetooth device are interfaced [2-7]. As a transmitter, the cell phone application is used to send ON/OFF command signals to the receiver. Thus, by receiving these commands from cell phone application given by the user the appliances can be turned ON/OFF remotely. Internet of Things (IoT) technology is used to control various electrical appliances remotely over the internet [8]. The smartphone android application with user configurable GUI front end is used for the real-time scenario.

In IR remote controlled home automation system, the remote sends commands to control home appliances. After receiving the signal from the remote, Arduino sends the related signal to relays which are responsible for switching ON or OFF of the home appliances through a relay driver [18].

Usually, it is difficult for visually impaired people to operate the android phone and IR transmitter which require line of sight between transmitter and receiver. So, the above systems are not easily applicable for elderly and impaired people and need a better home automation system that easily operable by the user. In this paper, we propose a home automation system that controls the household appliances by Bluetooth remote and produces the corresponding voice response from the receiver of the appliances. By the voice response, the impaired people can easily make sure that the appliance is active or not now, according to the user command.

This paper consists of six sections. Section I discusses the introductory concept of the proposed system. Section II discusses the recent works on wireless home automation system. The proposed system designing procedure is discussed in Section III. Section IV discusses about the implementation and cost analysis. Section V discusses about the system evaluation results. Finally, Section VI concludes the works and discusses some direction of future works.

II. RELATED WORK

In literature a number of remote-controlled home automation systems have been found. For example, Piyare et al. [2] provided full function remote control home automation system in which Arduino Bluetooth (BT) and cell phone communicate wirelessly using Bluetooth technology. The home appliance was controlled by using Symbian OS cell phone application and the Arduino Bluetooth board was connected to these appliances. Java-based application cannot support the Symbian OS cell phone and it is only supporting the python language scripts. Note that nowadays mostly smartphone applications are developed in Java. Like this, another study [3] presented a home automation system using Bluetooth and android application. It was not feasible to control more than 4 home appliances and it was designed to control only for 4 lights.

An internet-based smart home automation system [4] proposed and implemented that can be controlled remotely upon user authentication. Another internet-based home automation system [10] is designed for the rural environment using Raspberry Pi (RP) and web page. After login in the webpage, the user can control the appliances using webpage via the internet. However, using this system user can control the appliances from anywhere through the internet but the user is not sure that the appliances are controlled and it cannot use the impaired people.

In [5] a low cost and user-friendly remote-controlled home automation system is designed using Arduino board, Bluetooth module, smartphone, ultrasonic sensor, and a moisture sensor. A smartphone application is used in this system to control the household appliances and monitor the sensor values. In this system Bluetooth wireless connection was established between smartphone and Arduino Uno. The smartphone cannot operate by the visually impaired people, so this system is not applicable for disable people. We proposed a system for visually impaired people using Bluetooth network. The impaired people can easily operate the Bluetooth remote control of the proposed system. By pressing the button of the remote, the visually impaired people can control household appliances.

The infrared remote can be used for this proposed system, but the coverage area of infrared is sort than Bluetooth and it required line of sight. Thus, the Bluetooth network-based remote-controlled home automation system is proposed in this paper.

III. SYSTEM DESIGN AND WORKING PRINCIPLE

A. *Proposed Home Automation System*

The proposed system has two main parts hardware and software. The hardware part consists of some electronic devices. Software part consists of an Arduino integrated development environment (IDE). The overall system divided into two sections.

1. The transmitter or Bluetooth remote.
2. Receiver or relay module.

The transmitter or Bluetooth remote consists of a keypad, microcontroller, and a HC-05 Bluetooth module. The keypad in transmitter consists several push buttons. The receiver or relay module consists of a microcontroller, a HC-05 Bluetooth module, 4-channel relay module, micro-SD card reader module and speaker, etc. The number of channels in the relay module depends on how many appliances are controlled.

Fig. 1 illustrated the block diagram of proposed home automation system.

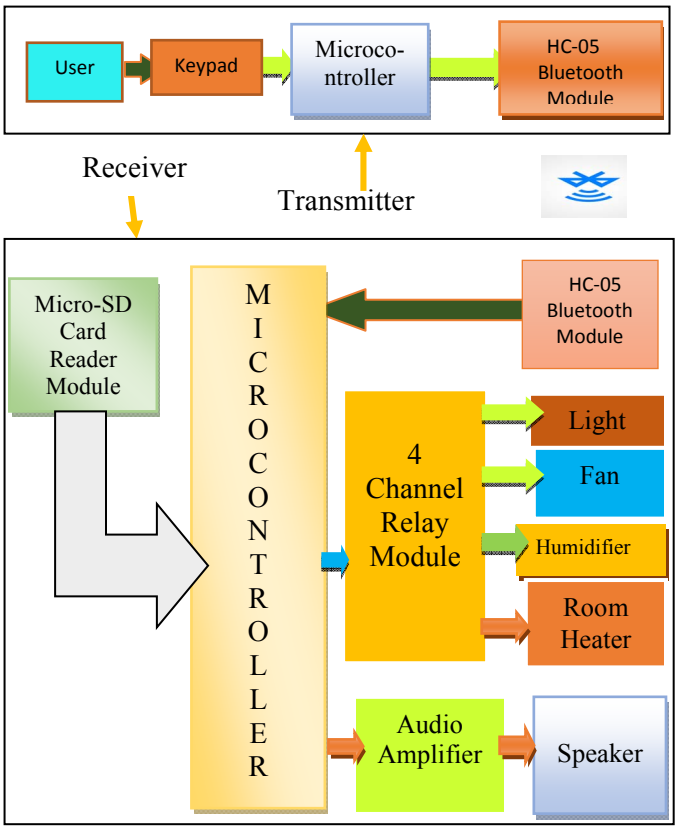


Fig. 1. Block diagram of proposed system.

B. *Hardware Components*

The following components are required to design the overall system.

- Arduino Uno
- Push button
- HC-05 Bluetooth module
- 4-channel relay module
- Micro-SD reader module
- Micro-SD card
- Speaker
- LM 386 audio amplifier.
- A Power source and DC adapter

C. *Master-Slave Configuration of HC-05 Bluetooth Module*

HC-05 module is an easy to use Bluetooth SPP (Serial Port Protocol) module, designed for transparent wireless serial connection setup. Before the design of the proposed system,

we need one HC-05 Bluetooth module as master and another is a slave to communicate between transmitter and relay control receiver device. HC-05 Bluetooth module has two modes of operation, Command Mode where we can send attention (AT) commands to it and Data Mode where it transmits and receives data to another Bluetooth module. The default mode is Data Mode and works as a slave. To configure the master mode, we need to change the default configuration. Using the AT command, we can change the default configuration. To set the modules into Command Mode, we need to connect the KEY pin high before applying power to the module.

D. Design of Transmitter or Bluetooth Remote Circuit

The transmitter or Bluetooth remote circuit is illustrated in Fig. 2. It consists of a master HC-05 BT module, eight push buttons and power source with Arduino Uno. In this circuit, the BT module transmits different signals according to pressing the push button. Each push button has a specific digital signal. These push buttons are numbered from left to right, button 1 is used to turn ON the light, button 2 is used to turn OFF the light. Similarly, button (3, 4), (5, 6), (7, 8) are used for fan, humidifier and room heater respectively for ON and OFF. The system could be extended for any numbers of appliances by increasing number of push buttons.

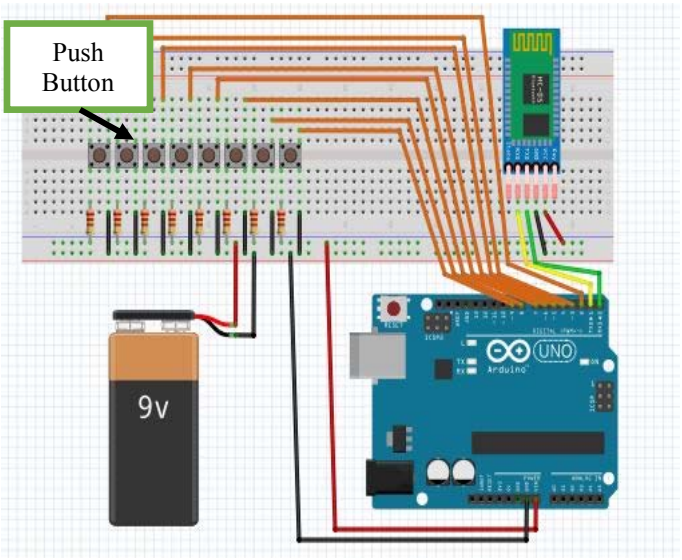


Fig. 2. Transmitter or Bluetooth remote circuit.

E. Design of Receiver or Relay Module Circuit

Circuit diagram of the receiver or relay module is illustrated in Fig. 3.In the receiver HC-05, the BT module acts as a slave, so it receives the transmitted signal. According to the transmitted signal, the channel is ON or OFF of the 4-channel relay module. The household appliances are connected to the relay channel. The voice response section is one of the parts of the receiver circuit. It consists of LM386 audio amplifier, Micro-SD card reader and speaker.

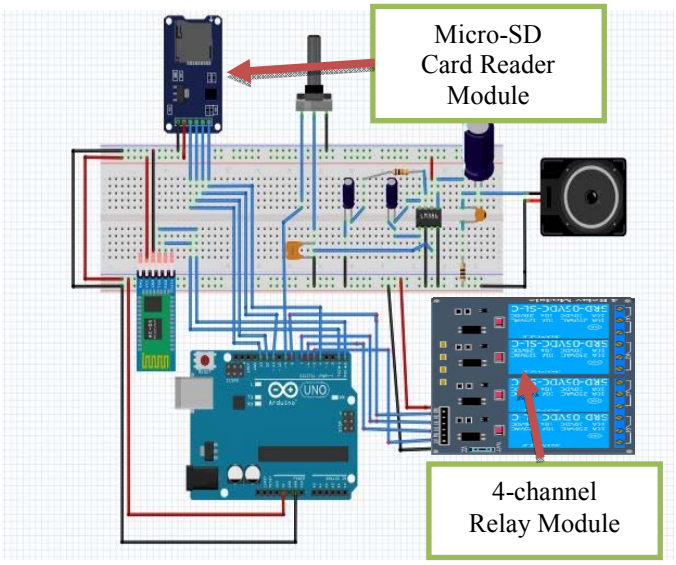


Fig. 3. Receiver or Relay control circuit.

IV. IMPLEMENTATION

A. Experimental Setup

The designed transmitter or Bluetooth remote is illustrated in Fig. 4. The user can control the household appliances using the module within the Bluetooth communication range. The device is portable, so the user can carry easily. The use of this device is simple. When a push button is pressed, it transmits a digital signal via HC-05 master Bluetooth module to the relay module to control the household appliances.

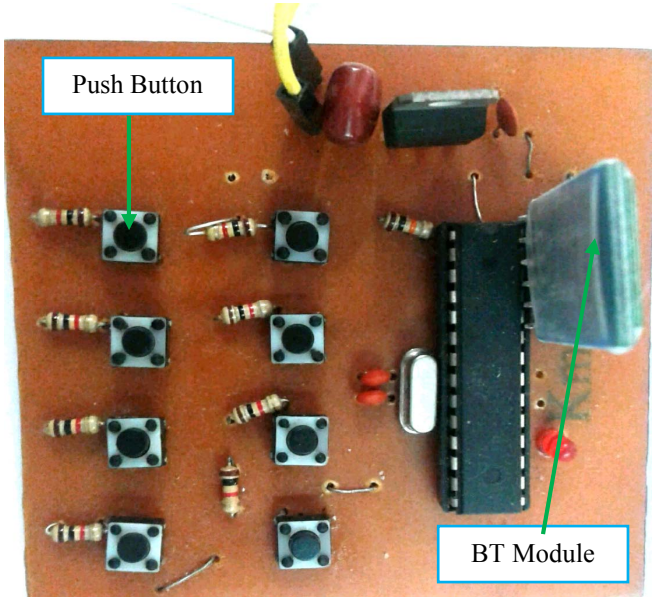


Fig. 4. Transmitter or Bluetooth remote.

The receiver or relay module is illustrated in Fig. 5. We use four appliances (light, fan, humidifier, and room heater) and show with four lights for primary implementation.

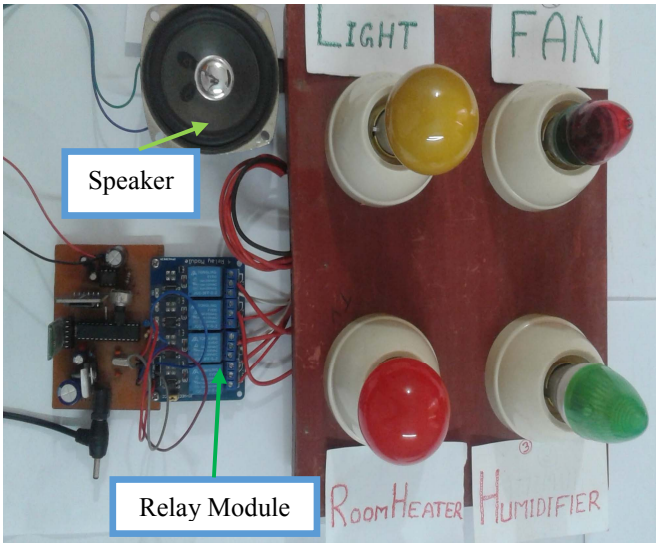


Fig. 5. Receiver or Relay module.

In the relay module device, the pre-recorded voice response is provided via the speaker. For example, when the user presses the button 1 (note button 1 for turn on light), then voice response gives as "Thank you, you on the light" two times repeatedly. Similarly, the response will be provided for fan, humidifier and room heater, when button 3, 5 and 7 are pressed respectively. When a user presses the button 2, then voice response gives as "Thank you, you off the light" two times repeatedly. Similarly, the response will be provided for a fan, humidifier, and room heater, when pressing button 4, 6 and 8, respectively to off the devices. Table I summarizes the voice response of all the 8 buttons.

TABLE I. SWITCHING SUMMARY OF THE FOUR HOUSEHOLD APPLIANCES

Name of Appliance	Button Number	Work	Transmitted signal	Voice response From receiver
Light	Button 1	ON	a	Thank you, you on the light
	Button 2	OFF	b	Thank you, you off the light
Fan	Button 3	ON	c	Thank you, you on the fan
	Button 4	OFF	d	Thank you, you off the fan
Humidifier	Button 5	ON	e	Thank you, you on the humidifier
	Button 6	OFF	f	Thank you, you off the humidifier
Room heater	Button 7	ON	g	Thank you, you on the room heater
	Button 8	OFF	h	Thank you, you off the room heater

B. *Cost Analysis*

We summarize the used equipment and cost to design the system in Table II. It is seen from the table that the designed home automation system is cost effective and economical. The cost of this system depends on the setup environment. In Table II we show the only cost to control four household appliances.

TABLE II. COSTS OF EQUIPMENTS

Name of Equipment	Quantity	Cost (BDT)
Microcontroller	2	2x175=350
HC-05 Bluetooth module	2	2×250=500
SD card reader module	1	240
4-channel relay module	1	250
Speaker	1	50
AC to DC adapter	1	75
Switch, Battery and other parts	-	700
Total cost	-	2165

V. SYSTEM EVALUATION

Our system is evaluated in two ways: Firstly, we conducted experiments to measure the effectiveness of the system. We control our home appliances under different distances from the receiver or relay module in real house environment. Secondly, we conducted an experiment with general participants to verify the usefulness to control the home appliances through evaluating their impression about the system.

A. *Percentages of Success*

Usually, the coverage area of the Bluetooth device is 10-100 m long from the receiver [1]. To measure the coverage area of the proposed system, we take the percentage of success in controlling the home appliances in different distance from the receiver or relay module. There is no line of sight between transmitter and receiver. Table III summarizes the result of the experiment.

TABLE III. MEASUREMENT OF COVERAGE DISTANCE

Distance of the Remote from Receiver (m)	No. of attempt	No. of Successful attempt	No. of failure attempt	Percentage of Success
2	15	15	0	100%
5	25	25	0	100%
8	25	21	4	84%
10	25	17	8	68%
12	25	10	15	40%
15	20	0	20	0%

The table shows that when the remote is used within the 8 m from the receiver, the proposed system is robust to control the home appliances. But, when the distance of the remote

from the receiver increases by 2 m then the percentages of success to control the home appliances drastically reduced to around 68%. When the remote was around 12 m away from the receiver then the performance of the proposed system is not satisfactory to control the home appliances. Thus, we can conclude from the above experiment that even though the coverage area of Bluetooth is 10-100 m, but the proposed system can robustly control the appliances within the distance of 8 m from the receiver.

B. System Evaluation with Participants

We conducted an experiment by involving general participant to verify the effectiveness of the proposed system.

1) Participants: In the experiment total of 32 unpaid participants (all are male students) were participated at the University of Rajshahi (mean age 22 years).

2) Experimental Case: In order to evaluate the effectiveness of the proposed system in controlling home appliances using Bluetooth network-based remote system, we measure the impressions of the participants about the proposed system and also measure the impressions of participants of the traditional system. Every participant used the proposed system and traditional system in controlling the home appliances in a living room environment.

a) With Proposed System (Sys-1): In order to control the home appliances, participants use the proposed system (explained in Section III).

b) With Traditional System (Sys-2): In order to control the home appliances, participants use traditional switchboard which is not remotely operable but fixed on a wall. In this case, to control the home appliances, participants should move to the swtichboard physically to control their home appliances.

3) Measurement: In the experiment, we evaluate the particiapnts' impression of both systems. We asked participants to fill out a questionnaire for each method (after using the systems). The measurement was a simple rating on a Likert Scale of 1 to 7. The questionnaire had the following three items:

- **Q1**: Did you feel that the system is convenient in controlling your home appliances?
- **Q2:** Did you feel that the system is helpful for elderly and physically impaired people in controlling their home appliances?
- **Q3**: How satisfied are you with the system in controlling your home appliances?

C. Results

The experiment was performed in a within-participant design, and the order of all experiment was counterbalanced. We conducted the repeated measure analysis of variance (ANOVA) for all measures

1) Impression: The mean (M), standard deviation (SD), *F*, *p-value*, and η^2 of the questionnaire measures for both system are shown in the Table-IV.

Table-IV
PARTICIPANTS IMPRESSION FOR QUESTIONNAIRES

	Q1		Q2		Q3	
	Sys-1	Sys-2	Sys-1	Sys-2	Sys-1	Sys-2
M	6.6	5.38	6.41	3.47	6.13	4.63
SD	1.28	2.56	1.22	4.06	0.37	3.27
F(1,31)	10.88		43.7		21.06	
p-value	0.00245		0.0000002		0.00007	
η^2	0.166		0.457		0.242	

For questionnaire **Q1** (Table-IV), the difference between the two systems (Sys-1, and Sys-2) was statistically significant ($F(1,31) = 10.88$, $p < 0.01$, and $\eta^2 = 0.166$). Thus, participants feel that Sys-1 is more convenient over Sys-2 in controlling their home appliances.

For questionnaire **Q2** (Table-IV), the ANOVA analysis also shows a significant difference in considering Sys-1 over Sys-2 for the elderly and physically impaired person in controlling their home appliances ($F(1,31) = 43.7$, $p < 0.01$, and $\eta^2 = 0.457$). Thus, participants feel that Sys-1 is helpful over Sys-2 for elderly and physically impaired people in controlling their home appliances.

To find the impression between the two systems for questionnaire **Q3** (Table-IV), we also performed the ANOVA analysis on the Likert scale. The difference between the two systems (Sys-1, and Sys-2) were also statistically significant ($F(1,31) = 21.06$s, $p < 0.01$, and $\eta^2 = 0.242$). It reveals that participants are satisfied in using our proposed system (Sys-1) in controlling their home appliances.

Thus, we conclude that the proposed system (Sys-1) is more convenient and effective than the traditional system (Sys-2) for controlling home appliances. It can also be stated based on the above analysis that our proposed system is also helpful for elderly and physically impaired people to control their home appliances.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have proposed the design and implementation of a low cost and user-friendly Bluetooth network-based home automation system, to control the home appliances remotely. The proposed system is implemented to control four appliances including light, fan, humidifier and room heater. The number of household appliances can increase by increasing the number of push buttons in the transmitter. The system is better to use for ON/OFF switching the appliances remotely without internet connection. The system is easy to operate by elderly and visually impaired people than other remote-control home automation system. Visually impaired people cannot run the android phone easily. The elderly and visually impaired people can easily control

the home appliances by pressing the push button of the remote module.

Future works can be focused on enhancing the performance of the system. The Bluetooth wireless communication system can be replaced with another wireless communication system to cover large areas.

REFERENCES

[1] The official Bluetooth website from Bluetooth SIG: http://www.bluetooth.com.

[2] R.Piyare and M.Tazil " Bluetooth based home automation system using Cell phone," 2011 IEEE 15th International Symposium on Consumer Electronics (ISCE), pp. 192-195, 2011.

[3] Ramlee, Ridza Azri, et aI., "Bluetooth remote home automation system using android application," The international Journal of Engineering and Science (The TJES), Vol. 2, Issue. 1, pp. 149- 153, 2013.

[4] Shirisha Tadoju and J. Mahesh " Bluetooth remote home automation system using Android application," International Journal of Advanced Technology and Innovative Research, Vol. 7, Issue. 10, pp. 1815-1818, August 2015.

[5] Muhammad Asadullah and Khalil Ullah "Smart home automation system using Bluetooth technology," IEEE, International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), pp. 1-6, 2017.

[6] N. Sriskanthan, F. Tan, and A. Karande, "Bluetooth based home automation system,'' Microprocessors and Microsystems, Vol. 26, Issue. 6, pp. 281-289, 2002.

[7] Kwang Ye01 Lee' and Jae Weon Choi' "Remote-controlled home automation system via Bluetooth home network," SICE Annual Conferencein Fukui, pp. 2824-2829, August 4-6,2003.

[8] Ali Ziya Alkar and Umit Buhur "An internet based wireless home automation system for multifunctional devices," IEEE Transactions on Consumer Electronics, Vol. 51, No. 4, pp. 1169-1174, Novembor 2005.

[9] Vikram.N, Harish K.S, Nihaal M.S, Raksha Umesh,and Shetty Aashik Ashok Kumar " A low cost home automation system using Wi-Fi based wireless sensor network incorporating Internet of Things(IoT)," IEEE, 7th International Advance Computing Conference, pp. 174-178, 2017.

[10] Ashokkumar B, Aswin Sriram M, Jones Michael J, Ramakrishnan R and Dr. Siva Subramanian S "Smart home development using relay circuit for rural environment," IEEE, Third International Conference on Science Technology Engineering & Management (ICONSTEM), pp. 34-37, 2017.

[11] Neng- Shiang Liang, Li-Chen Fu and Chao-Lin Wu. "An integrated, flexible, and Internet-based control architecture for home automation system in the internet era," Proceedings ICRA `02. IEEE International Conference on Robotics and Automation, Vol. 2, pp.1101-1106, 2002.

[12] Baris Yuksekkaya, A. Alper Kayalar, M. Bilgehan Tosun, M. Kaan Ozcan and Ali Ziya Alkar " A GSM, internet and speech controlled wireless interactive home automation system," IEEE Transactions on Consumer Electronics, Vol. 52, Issue: 3, pp. 837-843, August 2006.

[13] Kalyani Pampattiwar, Mit Lakhani, Rinisha Marar and Rhea Menon "Home automation using Raspberry Pi controlled via an android application," International Journal of Current Engineering and Technology, Vol.7, Issue.3, pp. 962-967 (June 2017).

[14] Tanish Sehgal and Shubham More "Home automation using IOT and Mobile App," International Research Journal of Engineering and Technology (IRJET), Vol. 04, Issue. 02 , pp. 694-698, February -2017.

[15] Ying-Wen Bai, Han-Wen Su and Wen-Cheng Hsu "Indoor and remote controls and management of home appliances by a smartphone with a Four-quadrant user interface," IEEE, International Conference on Consumer Electronics (ICCE), pp. 1-2, 2017.

[16] Deepali Javale, Mohd. Mohsin, Shreerang Nandanwar and Mayur Shingate " Home automation and security system using Android ADK," International Journal of Electronics Communication and Computer Technology (IJECCT), Vol. 3, Issue. 2, pp. 382-385, March 2013.

[17] Undavalli Vivek Chowdary, Kanduru Rohith, Pabolu Sandeep and Mutyala Ramu "Home automation system using IR sensors," International Journal of Electrical and Electronics Engineering (IJEEE), Vol. 4, Issue 6, pp. 11-16, Oct – Nov 2015.

Automated Railway Maintenance System in Context of Bangladesh

Nusrat Kabir Samia[1], Rifah Tamanna Usha[2], Anindita Kundu[3], Sadia Nowshin[4], Sharmin Sultana Sharmee[5], Sanjida Nasreen Tumpa[6], Major Md Fazle Rabbi[7] and Col Md Mahboob Karim[8]
Department of Computer Science and Engineering, Military Institute of Science and Technology
Dhaka, Bangladesh
Email: [1]nsrtkbr71@gmail.com, [2]rtamanna.usha@gmail.com, [3]anindita.gkundu@gmail.com, [4]sadianowshin2@gmail.com, [5]sharmee.sharmin@gmail.com, [6]tumpa.sanjida@cse.mist.ac.bd, [7]rabbi@cse.mist.ac.bd and [8]mahboob4146@yahoo.com

***Abstract*—Bangladesh railway system mostly relies on manual maintenance system. At present, a trolley car is used in which people go in person to examine the railway system depending only on their vision. Mostly, these checking include detection of rail line cracks and obstacles, existence of sleepers in the proper places etc. However, the inadequacy of this kind of maintenance system is one of the vital reasons of railway accidents. The current procedure is also time consuming and costly. Besides, one of the leading causes of railway accidents is attributed to human error which can be reduced through automation. This paper attempts to improve the railway maintenance system and convert this into an automated one. Here, we are introducing a sensor based arrangement which reveals problems like surface crack, unaligned rail lines, obstacles on the railway and also inspects the correct position, presence and absence of the sleepers with the convenience of sending the location and distance to the proper authority.**

***Index Terms*—Automation, Bangladesh railway, Obstacle detection, Railway maintenance, Sensor based system**

I. Introduction

In Bangladesh, railway is a very important domestic transport that links the length and breadth of the entire country. It contributes to the economic development of the country in a large way [1]. In our country, railway transport is most widely used and commonly used and cheapest mode of transport for the local people.

Though Bangladesh Railway is stated as one of the safest transportation system, its service condition and improvement status may not go with the statement [2]. Since 1972, Bangladesh has been going through a large number of train accidents. These days we can observe frequently in news that, railway accidents are rather common. The approximate number of accidents from 2001-2010 was about 5938 [3]. Among these, most of the accidents were for the derailment of train [4]. About 5461 train derailments were taken place from 2001-2010, which is approximately 92% of the entire incidences. These huge rate of derailments involve significant economic losses, interruption to services and obviously the safety risk of passengers [3].

The rapid increase of railway accidents have almost turned the railway safety status into impoverished state but the taken actions are not up to the mark. It has always been reported and claimed after an incident that main drawbacks behind this reason is the lack of obligation and the inappropriate management in operating system by the authority. Without proper steps, no solution will appear and the problems will last forever.

Analysis of train derailments in Bangladesh railway, shows that 72.6% (2195 out of 3022) of derailment was occurred because of the defects of track when 15.5% due to human failure, 7.1% due to vehicle failure, 4.8% due to combined factors and remaining 0.1% due to others [5]. This is the reason to automate the train maintenance system over the manual and human dependent present one.

From the above statistics, it is again clear that among the causes of railway accidents, one of the vital reasons is human error. The defects of track are difficult to understand manually. Generally in present railway system, the Track-men walk along miles of railway lines almost eight hours per day carrying 20-25 kg of monitoring equipment, in all atmospheric conditions. They have to check that the track bed under the sleepers is packed correctly or not, fix the cracks and sleeper damage, make sure of the nut-bolts and joints, and check on the clips that hold rails to the sleepers. This method of review is not so dependable because most of the time, those cracks can not be found by the Track-men. Also, this type of supervising system is not acceptable because of slowness and absence of conciseness, as the observer has to recognize the critical situations for a long period.

Our present model is a minor attempt to find out a way to make the maintenance system automated. In this paper we will describe a method to detect the surface and side cracks of railway. As well as we will detect fault in sleeper and obstacle in railway with latitude and longitude of that location. This system is designed in such a way which will perform as a railway safety monitoring system. Though this model will not serve the purpose of actual commercial use, but it is sufficient to show the way through which we can proceed to make the railway maintenance system completely automatic which will go through rail track and will detect the cracks and defects on the sleepers and hence preventing the rail accident.

The rest of the paper organization is as follows: We have covered the relevant researches in this field in the section II. In section III, the conceptual design has been discussed. In section IV, the technical implementation and work-flow have been demonstrated including real prototype images. Finally,

TABLE I
COMPARISON OF OUR SYSTEM WITH OTHER APPROACHES

Author	Paper Title	Paper Proposal	Our System
Sedat Ying Li et al [6]	Railway Fastener Inspection by Real-Time Machine Vision	In this paper, they have presented a real-time railway fastener detection system with the use of a high-speed laser range finder camera based on Pixel Similarity-Based Approaches and Histogram-Based Similarity Approaches with complex calculations	Our proposed system is simple to implement and the calculations that we have implemented are easy to understand.
Dogru et al [7]	Rail Component Detection, Optimization, and Assessment for Automatic Rail Track Inspection	Developed a group of image and video by vision-based analysis which also included distance measurement instrument for better precision and efficiency to detect significant rail modules for example ties, tie plates, and anchors automatically.	We put an automated sensor based system which can detect surface crack, unaligned rail lines, obstacle and additionally inspect the accurate position, presence and absence of the sleepers in context of Bangladesh. Also it can send the location of the defected position with distance from the station to proper authority.
Salzburger et al [8]	In-motion ultrasonic testing of the tread of high-speed railway wheels using the inspection system AUROPA III	Their Proposed system implements the examination of the whole boundary of the tread while the train passes over stationary ultrasonic probes at speeds of up to 15 km/h which only solves the problem related to the circumference of the tread.	We propose an automated railway maintenance system which will be able to detect four different most common defects in railway.

with some limitations described in section V, section VI concludes the paper.

II. LITERATURE REVIEW

Many approaches have been made for railway quality inspection. Fastener inspection is a part of general railway inspection. Fasteners are used for creating non-permanent joints in rail line and sometimes which can be damaged. Generally, the inspection to check the joints is done manually which is time consuming and crucial mistakes could be made. But by a real-time railway fastener detection system which

uses a faster 3-D camera placed under a railway carriage, this problem has been solved [6]. In another analysis, a real-time automatic vision based rail inspection system was presented, which performs inspections at 16 km/h with a 20 fps frame rate [7]. Detecting wheel tread damages is important for the security of the whole train. To detect crack-like defects in the tread of railway wheels AUROPA III is a new ultrasonic based inspection system [8]. This system enables the whole circumference inspection of the tread while the train passes over stationary ultrasonic probes at speeds of up to 15 km/h.

A number of processes have been used to perform fault detection in railway vehicles. These include advanced filtering, system identification and signal analysis methods. By few researches, the problem was solved by using track-mounted sensors for this kind of monitoring purposes [9]. Another system was developed which includes: a position system, an unmanned self-propelled vehicle, a sensor and a control system [10]. The unmanned automotive vehicle is adapted to travel on a railway track ahead of the train and has a sensor which is configured to detect an unfavorable railway condition. The positioning system is built to determine the position of the self-propelled vehicle. The control system in communication with the sensor and the train, where in the control system communicates the adverse railway condition to the train.

Elastic wave analysis for broken rail detection has also been done [11]. In that analysis, the process is based on elastic guided wave propagation through the rail. Broken rail detection is possible by using an impact device or energy propagating from the train wheel in contact with the rail. There has been a long debate among train control professionals about broken rail detection: where it is necessary and how it is achieved. There has been little research in terms of the actual practical application of new technology or risk mitigation approaches since conventional track circuits have been utilized to detect broken rails [12].

Some years ago, two railway inspection trains (RIT) already equipped with ultrasound, had additional advanced eddy current techniques installed. Recently, a new RIT was equipped with a system that was designed, from the beginning, to employ a combination of these two techniques for non-destructive rail inspection [13]. The eddy current technique has been developed to enable identification of rolling contact fatigue (RCF) faults. The ultrasound process is aimed at measurements in the rail bulk volume, which are not feasible using the eddy current technique. Another development of non-destructive techniques (NDT) for the in-service inspection of railroad wheels and gauge corners was the main activity of the NDT division VIII. 4 at BAM over the last 2 years. For such various components, two different inspection techniques were chosen to fulfill the end-user requirements. Firstly, the inspection of the wheels rim and disk should be carried out without dismantling the wheels and using ultrasonic techniques [14]. On the other hand, the inspection of the railroad track surface at a train speed of about 70 km/h should be guaranteed using eddy current processes. Eddy current application focused

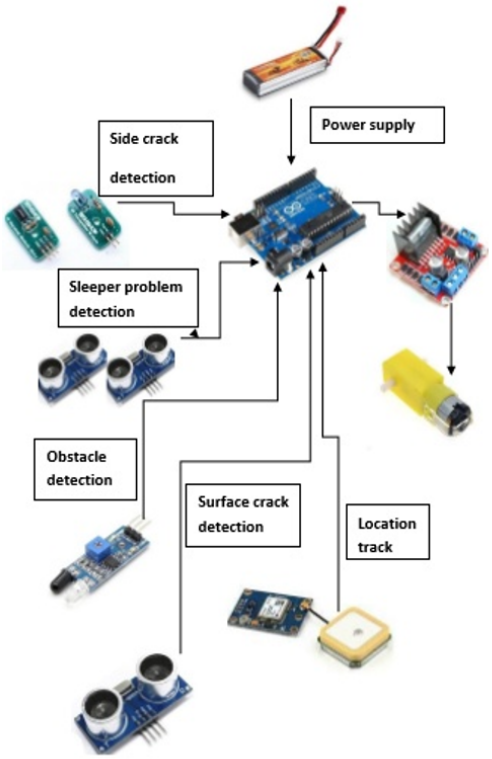


Fig. 1. System architecture of our proposed system

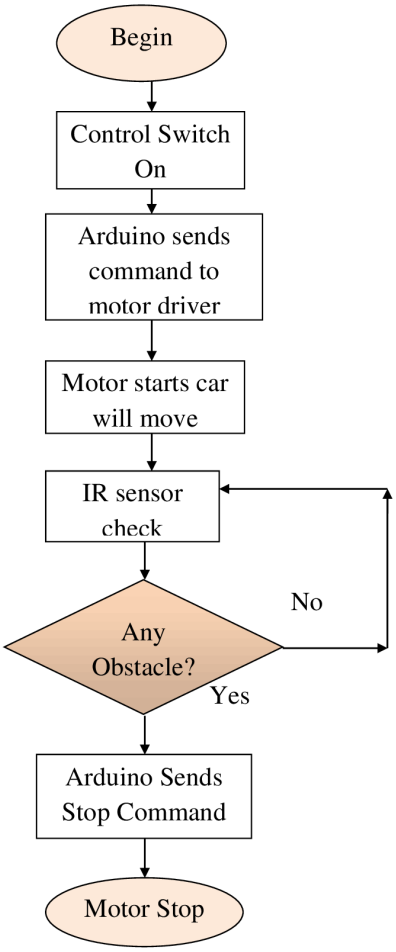


Fig. 2. Work flow of obstacle detection

mainly on the detection of head check defects occurring at the gauge corner of the rail.

The aim of a paper was to present a new vision based inspection technique for detecting special Rolling Contact Fatigue (RCF) defects that particularly occurs on rail head surface, meanwhile, an automatic detecting system is implemented, which consists of pre-processing, defect locating, defect identifying and post-processing subsystems [15]. To realize the defect locating sub-procedure, a simple and fast algorithm has been proposed, which adopts geometrical analysis directly on a gray-level histogram curve (the first-order statistical texture property) of the smoothed rail head surface image.

According to the researches so far related to railway maintenance system the major difference of other approaches and ours is the detection of slippers correct position and defect in slipper. Our proposed system is mainly based on different type of sensors which is less costly than many other proposed systems. In Table I we have shown a comparison between our approach and three relevant existing approaches. This comparison was based on only three approaches as they were pertinent.

III. Conceptual Design

In this section, we will describe about the main concept of our proposed maintenance car. Our car will detect different types of problems in railway system automatically. After detecting the problems, we will save the location of the problematic area into a local storage system. In addition, it will notify the proper authority so that problems can be repaired in time to avoid accidents. A system architecture of our maintenance car is given in Fig. 1. The main features of our system have been described below:

1. **Obstacle Detection**: Automated maintenance car will detect if there is any obstacle on the railway. Obstacles are one of the causes of train accidents. So, we will use a procedure which will detect obstacles on the railway and track the location. The flow of this feature is shown in Fig. 2.

2. **Surface Crack and Sleeper Problem Detection**: It detects the surface cracks of the railway which is one of the causes of train accidents. Here, we use a procedure which will continuously detect usual distance of the railway surface. When the measured distance is greater than usual distance it will detect the surface crack and save the location to database. It uses the same procedure for sleeper problem detection. It will be able to detect if there is any sleeper missing problem or sleeper alignment problem and will save those locations. The flow of this feature is shown in Fig. 3.

3. **Side Crack Detection**: It will detect side cracks and track

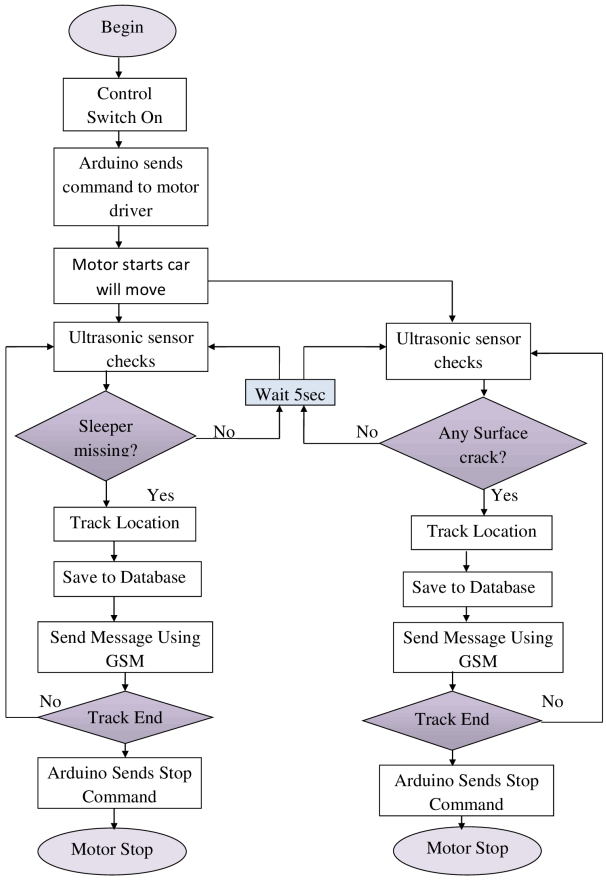


Fig. 3. Work flow of sleeper fault and surface crack detection

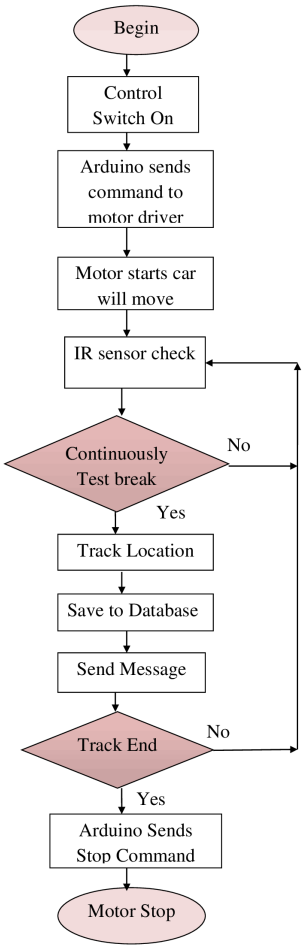


Fig. 4. Work flow of side crack detection

those locations. Here, we use a procedure which will receive a signal if there is any side crack, then it will save the location to the database. The flow of this feature is illustrated in Fig. 4.

4. **Storing Information**: Another important feature of our maintenance car is that it will be able save all the problem specifications along with location in a local storage, sd card.

IV. Implementation and Work Flow

The proposed system is implemented by IR sensor, Ultrasonic sensor, GPS, SD Card and Motor Driver Module which is demonstrated in Fig. 5 in the following manner:

IR Sensor module: Our maintenance car will detect if there is any obstacle on the railway with the help of IR sensor. Here, both the receiver and transmitter of IR sensor are set on the same side. The sensor will continuously send an infrared signal and when it gets any obstacle, the signal is reflected back to the IR receiver that captures the signal detecting the object.

IR Receiver and Transmitter Module Pair: Our maintenance car will detect side cracks with the help of IR sensor module. Here, IR transmitter and IR receiver are set on the opposite side. If there is any side crack, IR receiver will receive signal from IR transmitter. Thus it will be able to detect the side crack.

Ultrasonic Sensor Module: Our maintenance car will detect surface crack on the railway with the help of ultrasonic sensor. Ultrasonic sensor module can measure the distance to an object by using sound wave. The sensor of the maintenance car will continuously detect the distance of the surface of the railway and whenever the distance of the surface is greater than the usual distance it will be able to detect the crack of the surface.

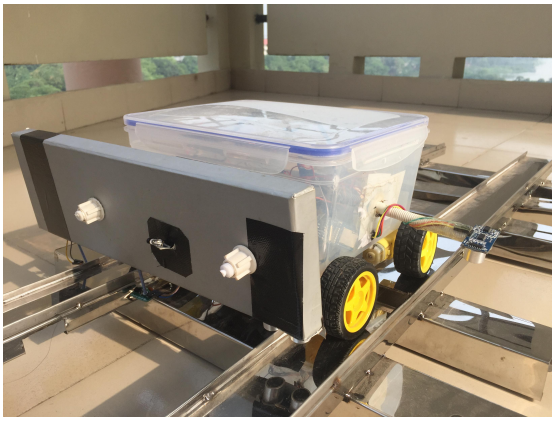
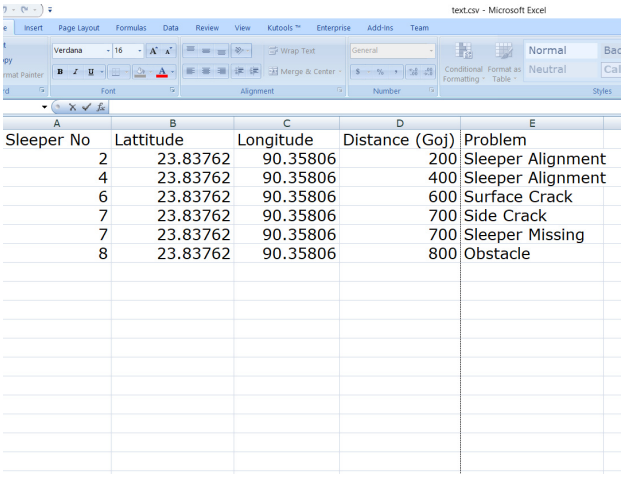


Fig. 5. Prototype of the proposed system



Sleeper No	Lattitude	Longitude	Distance (Goj)	Problem
2	23.83762	90.35806	200	Sleeper Alignment
4	23.83762	90.35806	400	Sleeper Alignment
6	23.83762	90.35806	600	Surface Crack
7	23.83762	90.35806	700	Side Crack
7	23.83762	90.35806	700	Sleeper Missing
8	23.83762	90.35806	800	Obstacle

Fig. 6. Data Storage

Ultrasonic Sensor Pair Module: Our car will detect the problem of sleepers with the help of two ultrasonic sensors. Here, we use the same feature of ultrasonic sensor as in surface crack detection but in a different way. Ultrasonic sensors will sense the distance of the sleepers continuously after a certain period of time when it is supposed to detect a sleeper. If the distance of both the ultrasonic sensors are same it will consider that there is no problem with the sleeper. But if the the distance of the ultrasonic sensors differs it will consider it as slipper problem.

GPS and SD Card Module: Whenever, it finds any error it will work as its principle and track the location with the help of GPS and save the location along with problem specication to the local storage (SD card). Here, we are saving the error specifications along with location in .csv file, shown in Fig. 6.

Motor Driver Module: Here, with the help of motor driver we are driving our car and controlling speed along with time synchronization.

V. Limitations and Future Plan

The implemented maintenance car has some limitations. It is not capable of detecting all types of error. It needs some additional maintenance cost. It can not be controlled remotely and if there is any problem with the maintenance car during automated investigation, it can not be detected remotely. The system can be improved in future by detecting other types of faults of railway and controlling the maintenance car remotely. The maintenance cost can also be decreased in future.

VI. Conclusion

After a number of investigations, it has been found that the existing railway maintenance system is infrequent and time consuming as well as costly. Another cause of railway accidents is human error for manual maintenance system. Our proposed system is mainly based on the use of different type of sensors which is less costly than many other proposed systems.

We have tried to improve the existing system through our project. Here, we have improved our system by converting the system into an automated one, tracking locations of specific problems with the help of GPS and reducing inspection time, human error. We hope Bangladesh Railway will be benefited through our proposed system in a large manner.

References

[1] M. Islam *et al.*, "Role of maintenance management of rolling stock in bangladesh railway: a study to assess effectiveness and efficiency," Ph.D. dissertation, BRAC University, February 2015.

[2] "Bangladesh railway (br), people's republic of bangladesh," last accessed on July 21, 2018, at 08:18:00PM. [Online]. Available: http://www.railway.gov.bd/

[3] C. Barkan, C. Tyler Dick, and R. Anderson, "Railroad derailment factors affecting hazardous materials transportation risk," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1825, pp. 64–74, January 2003.

[4] "Bangladesh railway, bangladesh railway information book-2009." last accessed on December 27, 2015, at 02:08:00PM. [Online]. Available: http://railway.portal.gov.bd/sites/default/files/files/railway.portal.gov.bd

[5] M. Shajedul Islam, "Study on factors for train derailments in bangladesh," Master's thesis, Bangladesh University of Engineering and Technology, March 2012. [Online]. Available: http://lib.buet.ac.bd:8080/xmlui/handle/123456789/434

[6] Ç. Aytekin, Y. Rezaeitabar, S. Dogru, and I. Ulusoy, "Railway fastener inspection by real-time machine vision," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 7, July 2015.

[7] Y. Li, H. Trinh, N. Haas, C. Otto, and S. Pankanti, "Rail component detection, optimization, and assessment for automatic rail track inspection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 2, pp. 760–770, April 2014.

[8] H.-J. Salzburger, M. Schuppmann, W. Li, and G. Xiaorong, "In-motion ultrasonic testing of the tread of high-speed railway wheels using the inspection system auropa iii," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 51, no. 7, pp. 370–372, July 2009.

[9] R. Ngigi, C. Pislaru, A. Ball, and F. Gu, "Modern techniques for condition monitoring of railway vehicle dynamics," *Journal of Physics: Conference Series*, vol. 364, no. 1, p. 012016, May 2012.

[10] S. M. Branka and R. T. Kelleher, "Railway inspection system," U.S. Patent US20 150 225 002A1, August 13, 2015.

[11] J. L. Rose, M. J. A. Jr., and Y. Cho, "Elastic wave analysis for broken rail detection," *AIP Conference Proceedings*, vol. 615, no. 1, pp. 1806–1812, 2002.

[12] D. F. Thurston, "Broken rail detection: Practical application of new technology or risk mitigation approaches," *IEEE Vehicular Technology Magazine*, vol. 9, no. 3, pp. 80–85, September 2014.

[13] H.-M. Thomas, T. Heckel, and G. Hanspach, "Advantage of a combined ultrasonic and eddy current examination for railway inspection trains," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 49, no. 6, pp. 341–344, June 2007.

[14] H.-J. M. H.-M. T. H. W. Rainer Pohl, A Erhard, "Ndt techniques for railroad wheel and gauge corner inspection," *NDT and E International*, vol. 37, no. 2, pp. 89–94, March 2004.

[15] L. Jie, L. Siwei, L. Qingyong, Z. Hanqing, and R. Shengwei, "Real-time rail head surface defect detection: A geometrical approach," in *Industrial Electronics, 2009. ISIE 2009. IEEE International Symposium on*. IEEE, 2009, pp. 769–774.

Wireless Home Automation System Using IoT and PaaS

Adiba Mahjabin Nitu
Dept. of Computer Science and Engineering
Hajee Mohammad Danesh Science and Technology University
Dinajpur, Bangladesh
nitu.hstu@gmail.com

Md. Jahid Hasan
Dept. of Electrical and Electronic Engineering
Hajee Mohammad Danesh Science and Technology University
Dinajpur, Bangladesh
jahidnoyon36@gmail.com

Md. Shahin Alom
Dept. of Electrical and Electronic Engineering
Hajee Mohammad Danesh Science and Technology University
Dinajpur, Bangladesh
ashahin200@gmail.com

***Abstract*—Home automation based on Internet of Things (IoT) is getting huge attention of mass people in recent years. Smart home eases and secures the management of the home appliances. This paper presents a low cost and reliable smart home system that assists the users to manage home appliances without the need of their physical presence. The proposed system includes smart door lock system using Radio Frequency Identification (RFID) card and password. It can store and display information of temperature & humidity of a home, and notifies the users of switching on/off time of light, fan and other home appliances using embedded micro-web server and an IoT platform. The user can also control the light intensity and fan speed based on LDR and temperature sensor values respectively. The system includes gas leakage & fire alarm. It has leakage gas removing & fire extinguishing facility and SMS notification system using IoT platform. The system uses real IP and RESTful API for controlling, monitoring and accessing the home appliances remotely from anywhere in the world using Android based smartphone app or web app. This system is user friendly and energy efficient.**

Keywords—IoT, PaaS, Home automation, Micro-web server, Teleduino, Altair SmartCore, RESTful API, Arduino.

I. INTRODUCTION

Internet of Things (IoT) is one of the mostly talked research topics in the recent years. IoT based home automation system is getting popularities day by day. Home automation is the controlling of the different home appliances using smart devices anywhere in the world without the need of physical presence in the home. Smart home system is very beneficial in everyday life as it saves electricity and reduces worries for home security of working people. Due to the secure, easily controllable and monitorable features, advanced home automation system design has become a core interest for many researchers.

Nowadays communication has become much easier due to the increased use of smartphone and laptop. People uses internet more than phone calls and SMS to communicate with each other. It is easy to process huge data faster and easier with the help of internet. Every day millions of devices are connected with internet in homes, schools, workplaces, and vehicles for helping people to live and communicate in a smarter way which is needed to analyze and translate via a gateway device. IoT platform can handle this kind of work. IoT platform is an important component of IoT that provides valuable application specific services in many application domains. Many kinds of IoT platforms are available nowadays [1] [2]. Platform as a Service (PaaS) is one of them, where the web developer has to take care only about the functionality of his application. During a development, he can use the API provided by the company to access important features of the PaaS, such as performing database operations, or managing users of his application [2]. An IoT platform is an integrated service that offers the user the things he needs to bring in online. It is capable of supporting millions of simultaneous device connections and easily allows the user to configure devices for machine-to-machine communication. In this paper, Altair SmartCore IoT platform [3] has been used which is based on the PaaS base IoT platform. This IoT platform has been designed in such a way where anyone can connect any type of electrical/electronic device. It can execute user command using it's REST API, manages devices where anyone can check the status of a load and can control remotely. This secured IoT platform can also send E-mail, SMS, and alarm. It can typically handle different tasks such as store data, visualize data and works as a middleware solution between hardware and application layers. The Altair SmartCore IoT platform does not apply any charge for connecting first two devices.

With the rapid growth of smart technologies and uses of electrical equipments (light, fan, etc.) and sensors (humidity, temperature, etc.), it has become very hard to control and get information from them manually. To getting control of that devices and measures the data from them, Arduino is an easy and cost effective choice. Arduino Ethernet Shield is an Arduino extension which is used to connect Arduino to the internet for sending and receiving data and works as a micro-web server. When a user sends data as a command through the internet, this micro-web server fetches the data from the internet and sends command to the Arduino.

This paper is based on the design and implementation of a flexible, secured, low-cost, advanced home control and monitoring system. This system includes services such as gas leakage, fire extinguishing and trespasser protection. If any unauthorized person tries to enter the home, the system sends an SMS to the authorized phone number. The system has sensors. When the gas sensor finds anomaly, the system starts alarm through buzzer and the kitchen exhaust fan turns on automatically. Again, when the fire sensor finds anomaly the system starts alarm through buzzer and an emergency pump turns on automatically. The system sends SMS to the user for both cases. The system consists of a Teleduino server [4], micro-web server and Altair SmartCore server [5] for

controlling and monitoring home appliances from anywhere and anytime. Each time with the change of state of any appliance, the information is stored instantly and the system notifies the user by E-mail. This stored data can be checked from anywhere by logging in Altair SmartCore web server. So, the proposed system allows authorized home owners to monitor and control the home appliances remotely using the 'Home Automation' android or web app through server real IP.

II. Related Work

Nowadays IoT makes technology more advanced. IoT is the network protocol integrated with physical devices, home appliances and other embedded devices which enables these things to connect and exchange data over a network without the presence of human being [6]. IoT has a novel impact on home automation system. Smart home is not a new term, it has been developed from over a decade. Home automation or smart home provides convenience, comfort, security and energy efficiency to its householders [7] [8]. Smart home consists of wireless technology such as Wi-Fi, RFID, Bluetooth and cellular networks. These have been utilized to various embedded systems which sense, control and exchange remote data.

The paper in [9] presents a smart home system where Bluetooth has been used as a medium of communication, without any internet connectivity. The physical devices such as sensors, actuators which can be accessed and controlled by smartphone using built in Bluetooth connectivity is implemented in [10][11].

To overcome the limitation of using Bluetooth signal (maximum range 100 m), researchers also introduced a popular technology called Wi-Fi based home control system [12] [13]. In Wi-Fi based home control system, a dedicated webserver and personal computer have been used, which is complex and not flexible.

GSM based cellular network has been applied in home appliances management system for longer range communication than Wi-Fi [14]. AT command is used to control and communicate with home appliances. The main drawback is that GSM has no suitable Graphical User Interface (GUI).

The authors in [15] have used MIT App Inventor, NearBus, and Arduino to develop a smart home prototype for controlling the home appliances without any monitoring capability.

Moreover, in papers [21][22][23], authors have designed Raspberry Pi based home automation system which is very costly. They have not used any IoT platform so their system is not capable of checking current status of home appliances. No security system is proposed there as well.

In paper [24], the authors proposed IoT based smart security and home automation system. They used PIR motion sensor-based security system and phone's digit pressing user interface for controlling home appliances using TI-CC3200 Launchpad board. Considering all these issues, an IoT based home automation system has been designed and implemented that can be controlled remotely from anywhere.

III. Description of The System

A. System Architecture

This section describes the details architecture of the proposed system. The system architecture is divided into five layers; remote login, server connectivity, real time database with notification system, home gateway and home environment.

Remote login allows the authorized user to login and access the system with 'Home Automation' app from anywhere in the world using Wi-Fi or 3G/4G network.

Server connectivity is controlled by Teleduino server, which links up between IoT platform and embedded micro-web server using a unique key. Teleduino converts Ethernet enabled Arduino into a versatile tool for interacting with devices over the internet. Ethernet enabled Arduino automatically connects itself to the Teleduino server when it is powered on.

Altair SmartCore web server stores data and works as real time database management system. When any appliances present state is changed, the server stores it and notifies the user by sending E-mail using Altair SmartCore Email API [3]. It also sends SMS for specific appliances (door, gas & fire sensor) it will send SMS using Altair SmartCore SMS API [3]. To get this facility, the user has to first create an account to the IoT platform.

Home gateway is an embedded micro-web server based on Arduino Ethernet Shield. The embedded micro-web server maintains and controls the home appliances using HTTP request by navigating Ethernet Shield's IP address through browser.

Home environment is directly related with home appliances. It consists of electrical equipments, modules and actuators. Module such as RFID [16] has been used for smart door lock system that ensures the entrance of authorized user only. This proposed system includes sensors such as temperature, and humidity for monitoring the home environment.

B. System Implementation

This section presents the explanation of working principle of the proposed system. The home gateway system is divided into two-parts; server connectivity and TCP/IP protocol. The server connectivity has three sections; Teleduino server, micro-web server, and Altair SmartCore server.

To implement the system, at first it is needed to establish a connection with Teleduino server, which creates a link between micro-web server and Altair SmartCore server. The user first sends initial standard HTTP command (using 'Home Automation' app) to embedded micro-web server through the Teleduino server. The micro-web server passes the command to the Arduino and the Altair SmartCore server.

The home gateway is connected to internet over TCP/IP. Arduino Ethernet Shield also supports TCP/IP stack. During the configuration with TCP/IP, the Ethernet Shield establishes connection with Local Area Network (LAN) using public IP address. Thus it creates an embedded system using TCP/IP protocol [17]. Once the home gateway is initialized, Arduino gets ready to execute any command. Then it enters into idle state until any further command is received from the remote

user. When the user sends HTTP command successfully from the 'Home Automation' app, Arduino Ethernet Shield decodes the command and Arduino takes appropriate action accordingly. At the same time the decoded command is stored in the Alter SmartCore server and notifies the user through E-mail. The complete process is illustrated as a flowchart in Fig. 1.

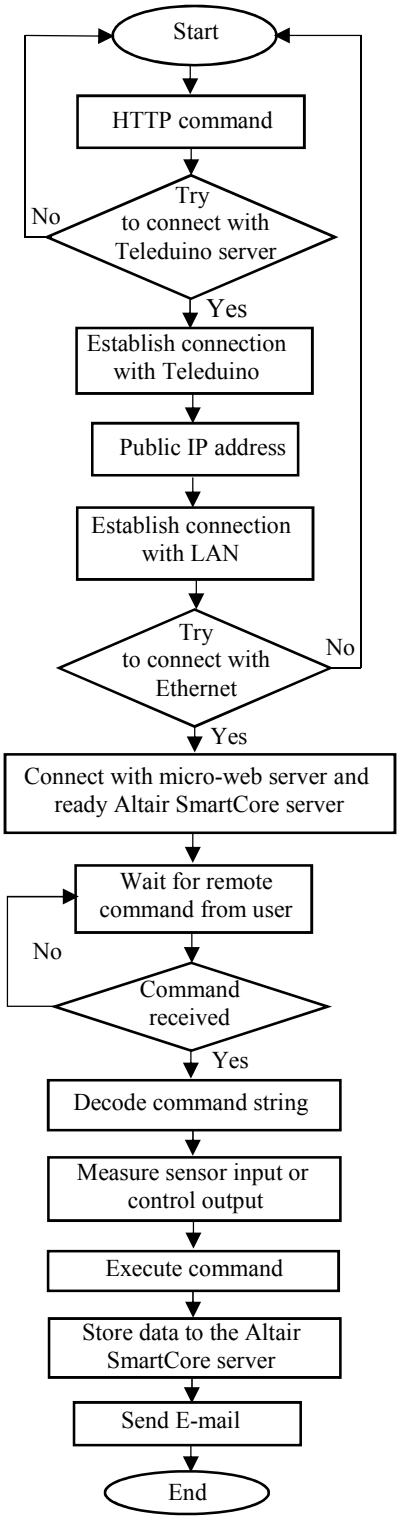


Fig. 1. Flow chart of E-mail notification process.

The home automation system includes some sensors and modules. The gas sensor detects gas leakage, fire sensor detects fire and keypad module is used for pressing password. The Arduino microcontroller is used to collect and analyze sensors data and controls the decision making process. The Altair SmartCore server stores the data and sends SMS. The complete process is shown in the following flowchart in Fig. 2.

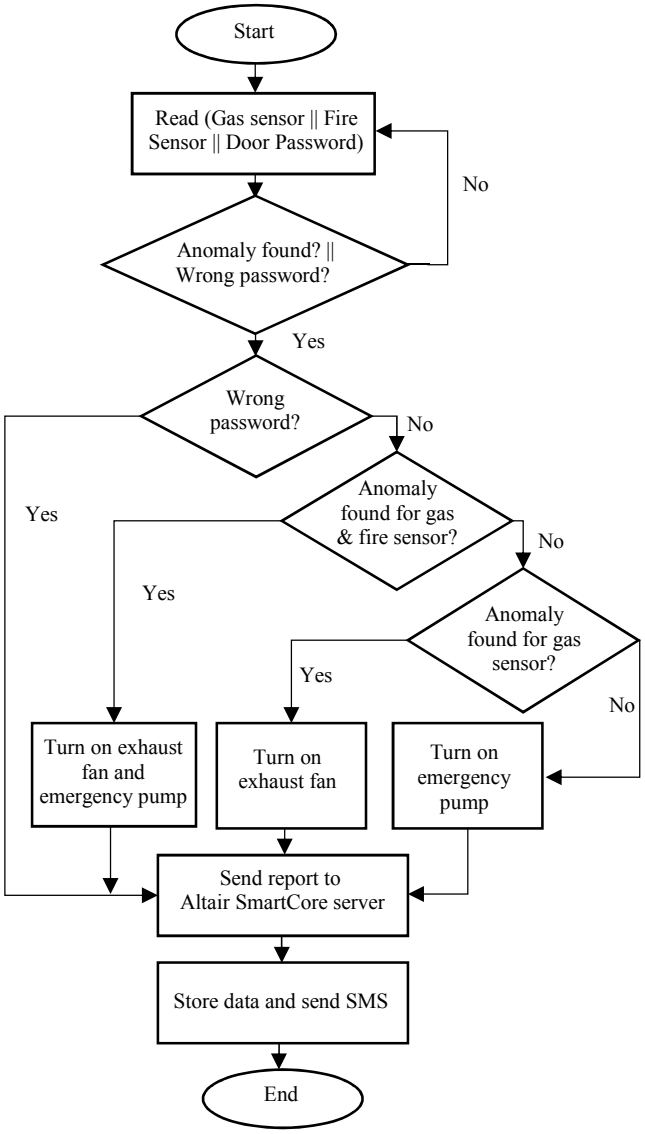


Fig. 2. Flow chart of SMS notification process.

C. *SystemFramework*

The home gateway framework is closely related with web services. The access and manipulation of web services have to be easy, direct, simple and interoperable. This eases the communication between the client and server applications of remote services. Web services are XML-based or JSON-based information exchange systems that use the internet for direct application-to-application interaction [18]. These systems can include programs, objects, messages, or documents. There are two main classes of web services; Simple Object Access Protocol (SOAP) and Representational State Transfer (REST). RESTful web service has a great deal with flexibility. It is much more stateless and lightweight than SOAP that offers functionality similar to SOAP based web services [19].

Therefore, this proposed system has used RESTful based web services. This service utilizes two standard operations: POST and GET. These two operations request and return

JSON response to communicate between the remote user and the micro-web server. JSON is much more lightweight for storing and exchanging data over client-server. It is easy for human beings to parse and generate messages simpler and faster than XML. For example, to turn a light on, an HTTP POST sends a request to micro-web server via Teleduino Server. After a successful operation the system sends response with success status response code "status": 200. The complete operation is illustrated as a data flow diagram in Fig. 3.

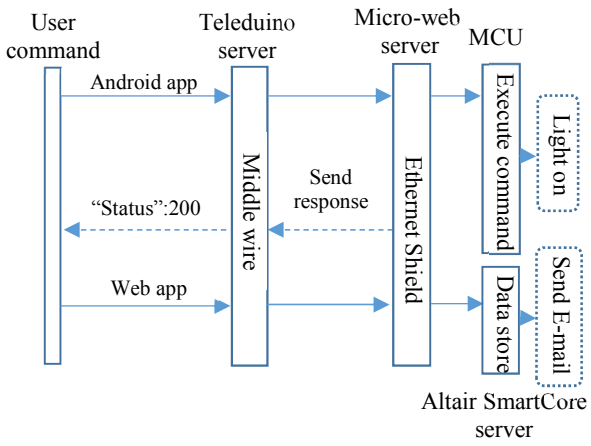


Fig. 3. Data flow diagram of web services.

D. Smartphone and Web Application

There are several platforms for developing smartphone application such as Android, Windows Mobile, iOS and others. About eighty eight percent people of the world use Android operating system as Android GUI is much simple and sophisticated [20]. Web application has been used productively because of the increased availability of information. These increases have led to heavier reliance on web-based application. Also, a web application can access anytime, anywhere through any web browser in any operating system with an internet connection. That is why this proposed system introduces both web application and smartphone application to control the home appliances. Fig. 4 and Fig. 5 illustrate the GUI of 'Home Automation' Android app and web application.

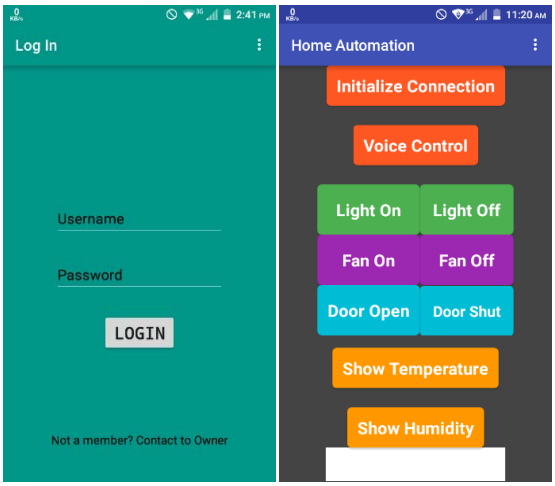


Fig. 4. Snapshot of 'Home automation' Android app with voice control option.

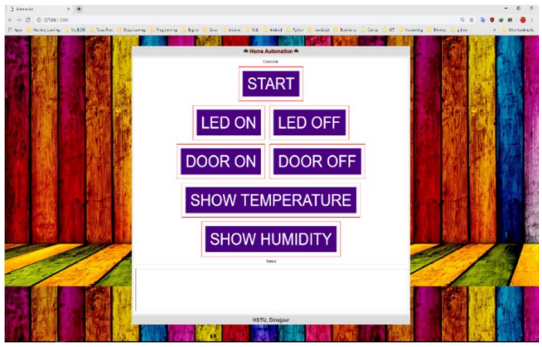


Fig. 5. Snapshot of 'Home automation' Web app.

To successfully control home appliances with home gateway, a user has to first connect with Teleduino server. When a user sends any command via 'Home Automation' app the micro-web server fetches it and forwards the request to the Teleduino server. The Teleduino server responses according to the command and sends back as JSON format to the user.

IV. Experimental Setup And Hardwire Configuration

The home automation system has used Arduino Mega as microcontroller board which works as the central control unit. Various modules including gas sensor, temperature and humidity sensor, RFID module, keypad, LCD, Relays and light/fan are interfaced to the Arduino board for testing the project. The experimental configuration and implementation of speech or manual control based home automation system are shown in Fig. 6 and Fig. 7.

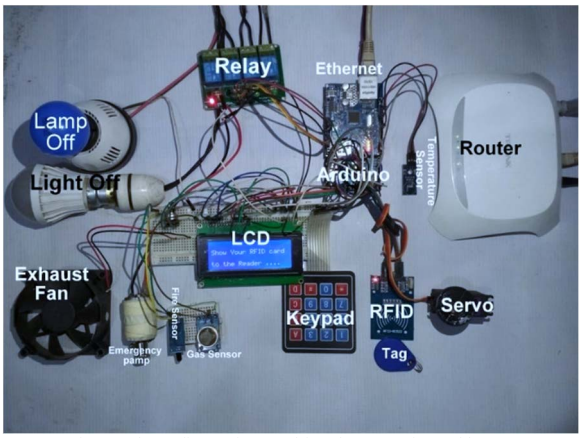


Fig. 6. Experimental configuration and implementation of home automation system with light and lamp off.

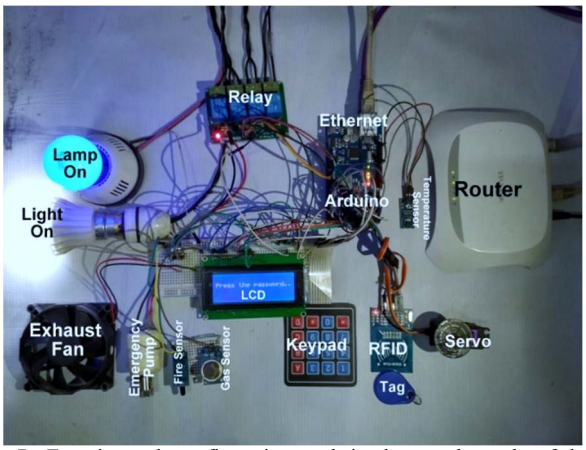


Fig. 7. Experimental configuration and implemented result of home automation system with light and lamp on.

This microcontroller board is interfaced with Arduino Ethernet Shield based on Wiznet W5100 Ethernet chip. The Arduino mega and Ethernet Shield are used to implement the micro-web server for home gateway control system. The Arduino mega 2560 is a microcontroller board based on the ATmega2560. A conventional light switch is integrated with Arduino using relays to demonstrate the switching capability. This system has sensors which detect input signals and actuators that perform responsive and adaptive functions. The experimental Setup of home automation system is shown as a block-diagram in Fig. 8.

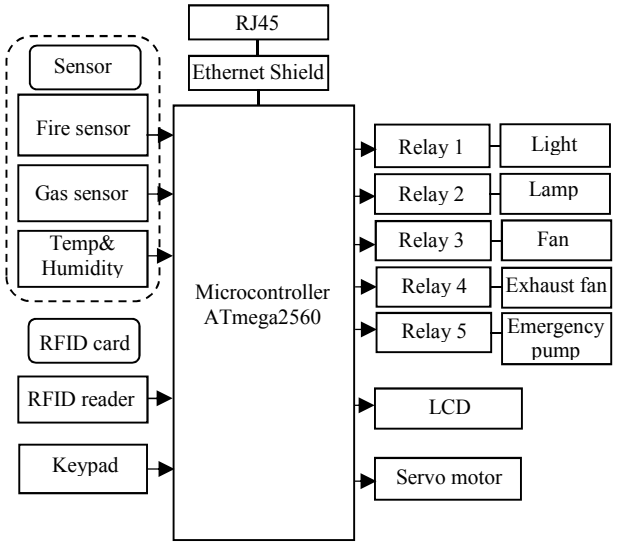


Fig. 8. Block-diagram of complete home automation System.

To configure this system, the user must connect with server through home gateway. To authenticate the user and to control the home appliances user has to first login and send initialize command through apps (smartphone/web) which will activate Ethernet Shield.

V. Comparison Study

A comparison study has been performed to measure the performance of the proposed system with other recent works. A table has been given below to compare the features of the existing works with the proposed system. The performance metrics cover many aspects such as real time database, instant notification through E-mail and SMS, control and monitor home appliances from anywhere and user interface. The TABLE I. shows that the proposed system is better than all others.

TABLE I. Comparison with Existing Works

Paper Ref. No	Performance Metrics				
	Real Time Database	*Instance Notification*	*Control home appliances from anywhere*	*Monitor home appliances from anywhere*	*User interface*
[21]	No	No	No	Yes	Web app
[22]	No	No	No	Yes	Web app
[23]	No	No	No	Yes	Web app
[24]	No	Yes	Yes	No	Digit of a phone
Proposed system	Yes	Yes	Yes	Yes	Android and Web app

VI. Conclusion

In this paper, an android and web based multi-functional advanced home automation system has been designed and implemented. The proposed system ensures easy and proper control of the home appliances from anywhere in the world. This system provides more security to a home by preventing intruders. It also protects a home from accident by removing leakage gas and fire extinguishing facility. This system is very beneficial for old ages, handicapped people and working people. The complete system has been tested and performed as expected. The GUI is user friendly and more practicable. The system ensures flexibility, reliability, energy efficiency, and a smart home.

VII. Scope for Future Work

In future, the home automation system would be made more functional. More home appliances will be included in the system. To improve the security of the system, facial recognition technique will be applied.

References

[1] M. Sruthiand B. R. Kavitha, “A Survey on Iot Platform”, vol. I, no.I, pp. 468–473, 2016.

[2] Clemens Pühringer, “Cloud Computing for Home Automation”, BACHELOR’S THESIS, Vienna University of Technology, pp 8-14, 2014.

[3] Altair SmartCore - Internet of Things Platform - Login. [online] Available at: https://www.altairsmartcore.com/ [Accessed 2 Feb. 2019].

[4] “Teleduino”, Arduino Internet / Web Control with Teleduino. [Online]. Available: https://www.teleduino.org/. [Accessed: 16-Oct-2018].

[5] “Internet of Things Platform” , Altair SmartCore. [Online]. Available: https://www.altairsmartcore.com/. [Accessed: 16-Oct-2018].

[6] Y. Liu, "Study on Smart Home System Based on Internet of Things Technology", In Informatics and Management Science IV. vol. 207, W. Du, Ed., ed: Springer London, pp. 73-81, 2013.

[7] D. Javale, M. Mohsin, S. Nandanwar, and M. Shingate, "Home Automation and Security System Using Android ADK", International Journal of Electronics Communication and Computer Technology, vol. 3, pp. 382-385, 2013.

[8] T. Perumal, M. N. Sulaiman, K. Y. Sharif, A. R. Ramli, and C. Y. Leong, "Development of an Embedded Smart Home Management Scheme", International Journal of Smart Home, vol. 7, pp. 15-26, 2013.

[9] Mohamed Abd El-Latif Mowad, Ahmed Fathy, Ahmed Hafez, "Smart Home Automation Control Using Android Application and Microcontroller", International Journal of Scientific & Engineering Research, Vol. 5, pp. 935-939, May-2014.

[10] R. Piyare and M. Tazil, "Bluetooth based home automation system using cell phone", In proceedings of 15th IEEE International Symposium on Consumer Electronics(ISCE), pp. 192-195, 2011.

[11] M. Yan and H. Shi, "SMART LIVING USING BLUETOOTH-BASED ANDROID SMARTPHONE", International Journal of Wireless & Mobile Networks (IJWMN), vol. 5, pp. 65-72, 2013.

[12] A. Z. Alkar and U. Buhur, "An Internet based wireless home automation system for multifunctional devices", Consumer Electronics, IEEE Transactions on, vol. 51, pp. 1169-1174, 2005.

[13] U. Sharma and S. R. N. Reddy, "Design of Home/Office Automation Using Wireless Sensor Network", International Journal of Computer Applications, vol. 43, pp. 53-60, 2012.

[14] R. Shahriyar, E. Hoque, S. Sohan, I. Naim, M. M. Akbar, and M. K. Khan, "Remote controlling of home appliances using mobile telephony", International Journal of Smart Home, vol. 2, pp. 37-54, 2008.

[15] Sarmla Tharishny, Saravanan Selvan, Umayal, Pratap Nair, "Android based Smart House Control via Wireless Communication", International Journal of Scientific Engineering and Technology, Vol. 5, pp. 323-325, May-2016.

[16] Mohsen Darianian, Martin peter Michael, "Smart Home Mobile RFID-based Internet-Of-Things Systems and Services", In proceedings of International Conference on Advanced Computer Theory and Engineering, 2008.

[17] HongRiLi, "Research and application of TCP/IP protocol in embedded system", In proceedings of 3rd IEEE International Conference on Communication Software and Networks, pp 584-587, 2011.

[18] A. A. Abd El-Aziz, A. Kannan, "JSON Encryption", In proceedings of International Conference on Computer Communication and Informatics (ICCCI), Jan. 2014.

[19] F. AlShahwan, K. Moessner, "Providing SOAP Web Services and RESTful Web Services from Mobile Hosts", In proceedings of Fifth International Conference on Internet and Web Applications and Services, pp174-179, 2010.

[20] Global mobile OS market share in sales to end users from 1st quarter 2009 to 2nd quarter 2018, August 2018 [Online]. Available: https://www.statista.com/statistics/266136/global-market-share-held-by-smartphone-operating-systems/ [Accessed: 15- oct- 2018].

[21] Michelle D'souza, Nelsha Wilfred, Rochel Pereira, Thanya Rayen, Prof. Aparna Telgote, "Home Automation using Internet of Things", In proceedings of International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), pp. 559-561, 2017.

[22] S L S Sri Harsha,S Chakrapani Reddy,Prince Mary S, "Enhanced Home Automation System using Internet of Things", In proceedings of International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), pp. 89-93, 2017.

[23] Soumya S, Malini Chavali, Shuchi Gupta, Niharika Rao, "Internet of Things based Home Automation System", In proceedings of IEEE International Conference On Recent Trends In Electronics Information Communication Technology, India, May 20-21, pp. 848-850, 2016.

[24] Ravi Kishore Kodali, Vishal Jain, Suvadeep Bose and Lakshmi Boppana, "IoT Based Smart Security and Home Automation System", In proceedings of International Conference on Computing, Communication and Automation (ICCCA), pp. 1286-1289, 2016.

Design and Implementation of an IoT Based Home Automation

Md. Mohaiminul Islam
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Chittagong-4203, Bangladesh
E-mail: mohaiminul.jim23@gmail.com

Md. Nahiyan Farook
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Chittagong-4203, Bangladesh
E-mail: nahiyanfarook@gmail.com

S. M. G. Mostafa
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Chittagong-4203, Bangladesh
E-mail: mostafa_93eee_iiuc@yahoo.com

Yasir Arafat
Department of Electrical and Electronic Engineering
International Islamic University Chittagong
Chittagong-4203, Bangladesh
E-mail: yaeeeiiuc@gmail.com

***Abstract*— Advancement in latest technologies has been helping to accomplish our day to day activities in faster and smoother ways. The advanced technologies provide solutions on multi-tasking with added security features in different sectors. Moreover, automation processes have been making regular life style smarter with less time-consuming activities. At the same time, the number of internet users has been increasing day by day. Internet of things (IoT) is one of the latest and emerging technologies that includes wireless monitoring & controlling functions with the help of internet. An IoT based Home Automation System (HAS) uses computers or webpage or android app for monitoring various parameters to control e.g., different electronic home appliances. By using the HAS technology, the users can access multiple appliances over the internet. By monitoring the parameters like temperature, light intensity, and movement of a human being e.g., using temperature sensors, the electrical appliances such as fan & light can be switched ON/OFF via internet from anywhere of the world. This paper proposed an IoT based home automation system which is designed with an aim of cost reduction and user friendly operation of the system.**

Keywords—IoT, HAS, MCU, sensor, notification, server, Wi-Fi, Blynk.

I. INTRODUCTION

Now-a-days, the technology has been advancing promptly, and with the development of this technology, the use of the internet is becoming enormous. The use of internet is further extended to connect things or devices to interact with each-other intellectually so that it can provide useful and meaningful information. This phenomenon is called Internet of Things (IoT). Undoubtedly, it is a must technology for the modern world. We can't imagine a single day without internet. With the development in science and technology, many systems have been developed for the betterment of human life. Home automation system (HAS) is one of them.

The HAS can control lighting, atmosphere, amusement systems, and appliances. Moreover, it can incorporate security of houses such as access control and alarm systems.. The HAS usually interfaces the devices to be controlled to a central hub or "gateway." The UI either uses divider mounted terminals, tablet or personal computers, a cell phone application or a web interface to control the system [1]. The HAS can be composed and created by utilizing a unit controller which can control and monitor interconnected appliances, e.g., lights, control plugs, temperature and moisture sensors, gas and fire locator, and security systems [11].

Various communication technologies can be used for the HAS, e.g., RFID, ZIGBEE, Bluetooth, GSM and Wi-Fi [2]. Every communication system has some drawbacks. Some systems use conventional technologies. Some systems use modern technologies which also have some limitations in monitoring and controlling the components that are related to the automation system. In the RFID based HAS, the energy meter is credited through the subscriber controls RFID technology and the automation of different appliances via an Android mobile application [7].

In RFID based internet of things system and service, it presents a novel readout strategy for a different leveled remote ace slave RFID reader engineering of multi-standard NFC (Near Field Communication) and UHF (Ultra High Frequency) advances to construct a smart home automation system [8]. In the Wi-Fi-based HAS, the concept of internetworking of smart devices is crystalized. According to [9], a Wi-Fi based Wireless Sensor Network (WSN) can be used to monitor and control e.g., electrical parameters of a smart interconnected home. Speech based HAS is presented in [10] where two schemes are implemented. First one is Bluetooth technology, and the other one is GSM/GPRS technology. ARM11 Raspberry Pi microcontroller board is used in that system.

In the ZigBee wireless interface management module a ZigBee device module is fundamentally in charge of the correspondence between the host PC and ZigBee remote system, ZigBee home system data criticism, and so on [6]. In such a system it is necessary to establish a network and a sub-network for further control. The Bluetooth based HAS incorporates a host and a few customer modules. Considering the passageway cost of the Bluetooth, it will be more financially savvy to have different devices Controllers associated with a Bluetooth device [3]. However, this system ranging problem and also has a limitation of the number of connectors to the Bluetooth device. In GSM technology,

when an SMS is sent to GSM module by Mobile, then GSM receives that SMS and forwards it to Arduino. Typically, Arduino reads this SMS and extracts basic command from the received string and stores it in a variable. Then, Arduino contrast this string with e.g., a predefined string. If matches happen then Arduino forwards the signal to relay via relay driver for switching ON and OFF the home appliances.

Moreover, relative outcome additionally prints on 16x2 LCD by utilizing suitable commands [4]. As the system is network dependent, it can easily get interrupted [5]. If a user makes any mistake in writing message format, the message would not be acknowledged by the Arduino device [5]. Hence, the HAS cannot be applied economically and efficiently in those systems. To overcome the limitations, this paper suggests a low cost, energy efficient system for the people.

Furthermore, the HAS can be accomplished by combining the GSM technology with short-range wireless communication technology, Bluetooth, and colse proximity ZigBee wireless technologies. The client **c**oupling has been accomplished using Android application. Th**e** Android application gets client's vocal information to portray an arrangement of printed directions. The portraid directions are transferred using SMS service to **a** telephone in house and exchanges the paths to a PIC controller using Bluetooth. This unit has been referred as a remote unit which then transfers the directions to a dedicated ZigBee handset which then exchanges these directions to the main controller using ZigBee technology**.** The controller then translates the instructions and plays out the planned activities. Moreover, it transmits back the status data using the ZigBee technology. Th**e** framework can provide remote access. Nonetheless, the frame may end up costly because of the utilization of various types of controllers. Additionally, the SMS framework can turn out to be not reliable [15].

In this proposed system, one will monitor various home appliances e.g., electrical appliances like light, fan, AC, etc., the condition of gas in our kitchen, the room temperature, and the presence or absence of residents in a house. For this automation purpose, the IoT technology is more reliable, cost-effective and provides more information. IoT is one of the promising advancements which can be utilized for associating, controlling and monitoring wise articles which are associated with the internet through an IP address [12]. The system can be controlled via voice recognition system. The voice command given by the owner is translated by the cell-phone utilizing Natural Language. The cell-phone goes about as a focal comfort; it figures out what task must be finished by which apparatus to satisfy the client's demand. The focal reassure may in like manner be either a work area application, web application or an advanced mobile phone application as almost the majority of the information exchanged can be handled by the cloud [13].

This paper aims to control the household electrical equipment through the use of the internet. A normal HAS comprises of 4 phases. Understanding the client condition by detecting, revealing the occasions to a brought together substance, concentrated element investigations and triggering the work process, work process will be executed and refresh client by any intelligent channels or even exercise over a home gadget (inciting) [14]. This system will be conducive for the person who leaves their house without monitoring the condition of electrical appliances. A person can control the state of electrical devices of his/her residence from anywhere of the world. If the dwellers go out, it becomes a great concern for the dwellers of the house. So it is essential for the dwellers to monitor the house. The proposed system in this paper can help in this regard by providing home automation such as controlling the current states of various electrical appliance, gas, and temperature using IoT technology. Anybody with the knowledge of the address of the web server can monitor the condition of home appliances and parameters in real-time.

Fig.1 shows the block diagram of the proposed home automation system. Every systems and sub-system will be depicted from where its operating procedure can be described. Various sensor modules will be connected with a microcontroller. The sensors will provide their outputs to the microcontroller. Then the microcontroller with having an adequate power supply and necessary processing and calibration shows this value via a mobile application. Again, the microcontroller connects to the internet using wifi module and sends the values to a remote IoT server where the dwellers or connected person can view the state from anywhere of the world.

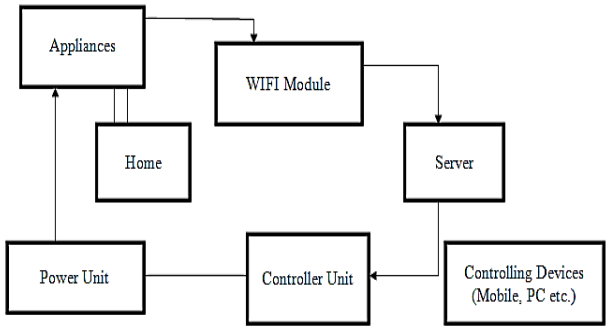


Fig. 1. Block diagram of the IoT based HAS

II. SYSTEM HARDWARE

The primary objective of this automation system is to design a HAS that is cost effective with higher reliability and which can be controlled from a long distance.. A smart control system can be developed with the use of IoT. The entire system consists of a microcontroller unit with Wi-Fi accessibility, a control unit, and a power unit. Fig. 2 shows a basic diagram for the whole of the home automation system. In the system, the heart of the system is Node MCU ESP8266, which is a microcontroller with Wi-Fi accessibility. This Node MCU ESP8266 is connected with the sensors and electrical appliances.

For controlling and monitoring the home electrical appliances, the user has to browse the particular web server; thus the user can control the status of the electrical appliances and in case of gas leakage, and if any motion detected by the sensor, it will give notification via a mobile application. So that the user need not browse the server all the times to know about sensors output. Different sectors of this HAs are:

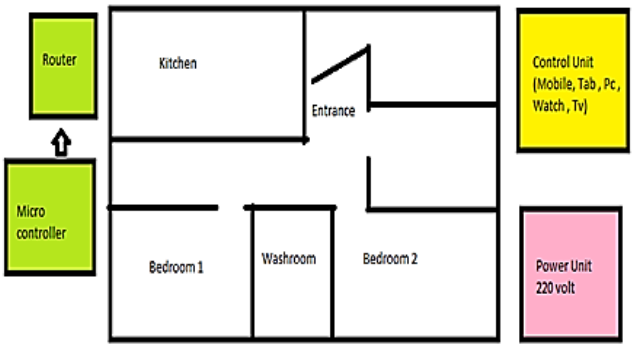


Fig. 2. Design of Home Unit

A. *Controlling Light, Fan or AC*

The States of various electrical appliances such as light, fan or AC can be monitored as well as controlled manually by a switch and automatically by the server.

B. *Open or Close a Door*

The door can also be controlled in this HAS. The door is connected with the Arduino microcontroller by a servo motor.

C. *Motion Sensing*

If the PIR motion sensor detects any motion, the sensor gives high output. Thus the microcontroller unit provides output to the controller unit as "Motion Detected." Fig. 3 shows the experimental setup of the proposed HAS. This HAS system consists of described appliances and sensors.

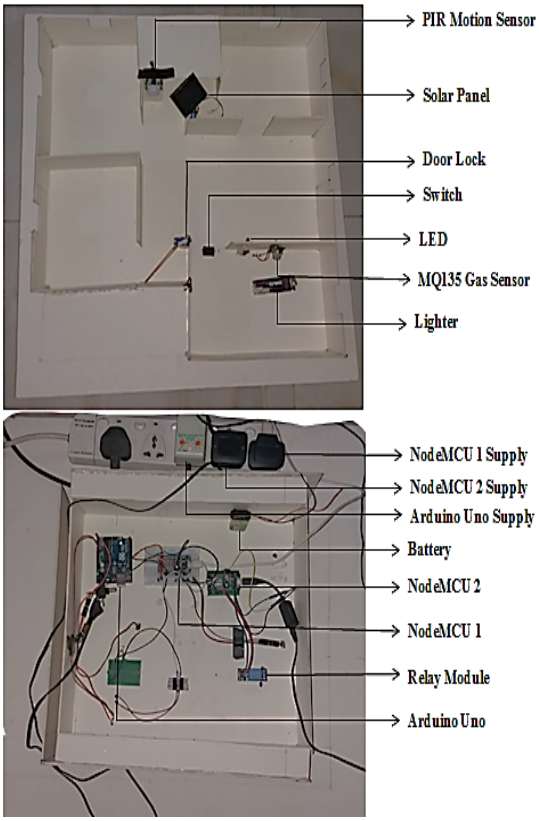


Fig. 3. Experimental setup for IoT Based Home Automation System

D. *Controlling Rotating Solar Panel*

Sun is a non-conventional source of energy. For which solar panel is placed up in the roof-top of the house. So according to the position of the sun, if the solar panel is placed, it can gather more power from the sun. By using a servo motor such placement is possible. By setting a definite angle according to the sun, the rotation can be made. By the use of the server, this rotation can be made from anywhere.

E. *Gas Leakage Warning*

In case of any inflammable gas leakage the MQ135 gas sensor senses the gas; thus there occurs a significant change in the conductivity of the sensor. So the sensor gives the output to the microcontroller unit, and from that, a notification arrived in our control unit such as mobile, tablet, PC, etc.

III. Hardware Programming and Application Development

A smart central controller Node MCU ESP 8266 is the fundamental component of the WSAN that is responsible to sort out and set up the remote monitoring system with control modules. A relating control module controls every machine. The home controlling unit is the main unit by which we can control the appliances of the system. The controlling devices can be mobile, tablet, laptop, pc or any other electric devices which has internet accessibility. Fig. 4 shows the flowchart of the proposed HAS. The smart central controller Node MCU ESP 8266 is a connection between the preceding and the following such that it merges the aggregated framework together. It is at the helm of laying out an RF 433 MHz WSAN e.g., in a home. These adaptable control modules are at the helm of commanding machines as well asspeaking with the principal supervisor. By using such a control framework, the occupants can employ mobile phone and/or PC to visually display home appliances and peform planned activities remotely, for example, turning on or turning off a lighting loads.

The software of the proposed HAS is isolated to server application programming and Microcontroller (NODEMCU) firmware. The server application proposed for the investigated HAS is used as online application constructed by using PHP, JAVASCRIPT, CSS and HTML. Server application programming is dependable of setup, arrangement and keeps up the entire home automation system. The server utilizes the database to keep a log of HAS parts, and we utilize JSON design spare system log. The Arduino programming, constructed employing C dialect, and also using IDE (Integrated Development Environment) coexist with the microcontroller. Arduino programming used in charge of gathering events from related sensors.

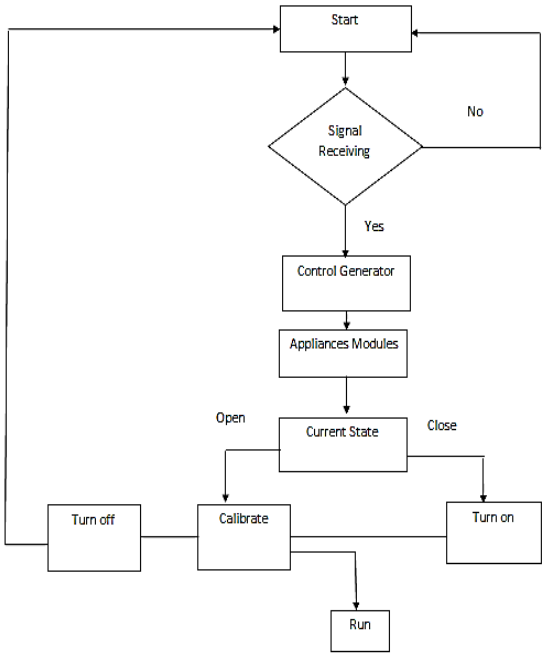


Fig. 4. Flow Chart of the Total System

IV. GRAPHIC USER INTERFACE

The essential requirement of HAS is to monitor and control the states of the electrical appliances using any mobile, PC, tablet, laptop, etc which are connected with the internet through the server. The electrical appliances and the solar panel controlled through the server. To make the system more effective, the sensor's output will be given via mobile application named as Blynk. Blynk software designed for such application to provide notification when the server gets the high output from the sensors.

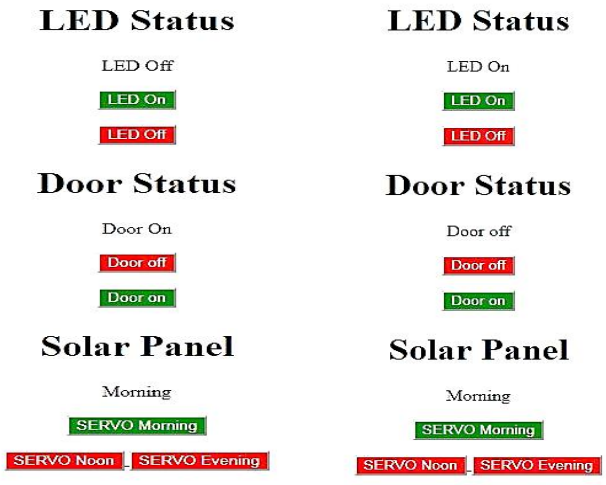


Fig. 5. Server Shows Different Appliances Status

Fig. 4 shows the server presents the status of LED, door and solar panel. When LED is off by tapping 'LED On' on the server the LED becomes on the state.

The door is automatically controlled via the server. While door remains off in state, by tapping 'Door On' on server the door opens.

The server shows the solar panel contains three states morning, noon and evening. According to the direction of the sun, the user can rotate the solar panel for capturing maximum power.

Whenever any flammable gas leakage occurs, gas sensor MQ135 becomes active, and it will notify via a Blynk application, Fig. 5 shows the result shows in the application.

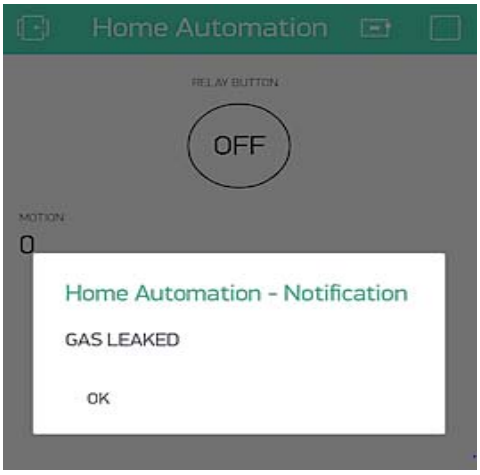


Fig. 6. Gas leaked Notification on Blynk Application

If the PIR motion sensor detects any motion, the sensor will also give notification via the Blynk application. Fig. 6 shows the notification of the application.

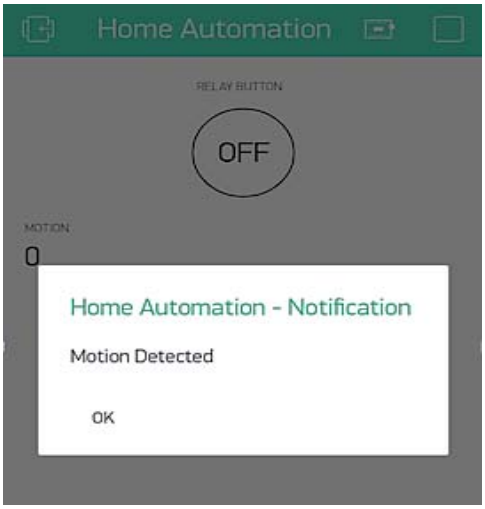


Fig. 7. Motion Detected notification on Blynk Application

By the use of server and the Blynk application, the HAS becomes more cost-effective, and it becomes easier to monitor and control the appliances.

V. CONCLUSION

In this work, the primary objective was to develop a HAS which can provide a more cost-effective solution with wide range of application. With the proposed HAS, more flexibility can be achieved as it uses the server to monitor and control appliances via internet access. The system can be installed into a house with an overall cost less than 100 USD. This system can also be useful for the handicap people to monitor the appliances. Various sensors can be used to perform applications such as water level indicating, rain alarm system, etc. This system can also be used for the industrial purpose so that the total industrial system can be controlled from a remote place by creating a domain based website. For the future improvement of the proposed HAS,

an automatic phone call system can be added. In case of gas leakage and fire alarm, the respective authority would get phone call automatically which would make the whole system more reliable when the dwellers are not in the town.

REFERENCES

[1] IoT, 23.07.2018, retrieved from: https://en.wikipedia.org/-wiki/Internet_of_things.

[2] Bluetooth based automation advantages and disadvantages, 24.07.2018, https://www.quora.com/What-are-the-disadvantages-of-using-home-automation-via-Bluetooth.

[3] GSM based automation, 24.07.2018, retrieved from https://www.slideshare.net/MainakSinha1/gsm-based-home-automation-62745555.

[4] GSM based automation, 24.07.2018, retrieved from https://circuitdigest.com/microcontroller-projects/gsm-based-home-automation-using-arduino

[5] GSM based automation, 24.07.2018, retrieved from https://www.slideshare.net/MainakSinha1/gsm-based-home-automation-62745555.

[6] Su Zhihua, "Design of Smart Home System Based on ZigBee, International Conference on Robots & Intelligent System," 2016.

[7] Muhammad Nafees, "RFID based prepaid energy meter and home automation with reporting," 23rd International Conference on Automation and Computing, 2017.

[8] Mohsen Darianian , Martin Peter Michael, "Smart Home Mobile RFID Based Internet-of-Things Systems and Services," International Conference on Advanced Computer Theory and Engineering, 2008.

[9] N. Vikram , K.S. Harish , M.S. Nihaal , Raksha Umesh , Aashik Shetty , Ashok Kumar, "A Low-Cost Home Automation System Using Wi-Fi Based Wireless Sensor Network Incorporating Internet of Things(IoT)," 7th International Advance Computing Conference, 2017.

[10] Dhiraj Sunehra , Vemula Tejaswi, "Implementation of Speech Based Home Automation System Using Bluetooth and GSM," International Conference on Signal Processing, Communication, Power and Embedded Sytem (SCOPES), 2016.

[11] Muhammad Asadullah, Ahsan Raza , An Overview of Home Automation Systems.

[12] Kumar Mandula, Ramu Parupalli, CH.A.S. Murty, E. Magesh, Rutul Lunagariya, "Mobile-based home automation using Internet of Things(IoT)," International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2015.

[13] Paul Jasmin Rani, Jason Bakthakumar, B. Praveen Kumaar, U. Praveen Kumaar, Santhosh Kumar, "Voice-controlled home automation system using Natural Language Processing (NLP) and Internet of Things (IoT)," Third International Conference on Science Technology Engineering & Management (ICONSTEM), 2017.

[14] S. M. Brundha, P. Lakshmi, S. Santhanalakshm, "Home automation in client-server approach with user notification along with efficient security alerting system," International Conference On Smart Technologies For Smart Nation (SmartTechCon), 2017.

[15] Satish Palaniappan, Naveen Hariharan, Naren T Kesh, Vidhyalakshimi, Angel Deborah, "Home Automation Systems - A Study," International Journal of Computer Applications, Vol. 116, No. 11, pp. 0975-8887, 2015.

Android Based Home Automation Using AVR Microcontroller Through Wireless Communication

Muzahidul Islam
Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
muzahidruet13@gmail.com

Md Tarek Hossein
Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
tarek101008@gmail.com

Md. Ilias Rahman
Electrical & Electronic Engineering
Rajshahi university of Engineering & Technology
Rajshahi, Bangladesh
himelmohammad@gmail.com

Md. Mohtasim Billah
Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
mutasimruet13@gmail.com

Mohammad Al Mahfuz
Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
almahfuz.ete.ruet@gmail.com

Mohammad Hasibul Hasan Hasib
Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
hasib.eee.ruet@gmail.com

Abstract— **Developing human's life activity with enhancing safety and security protocol Android-based automation is a vast alternative. This paper presents the design and implementation of an Android APP based home automation system which achieves multiple objectives about Password-based access control, induction motor speed and light intensity control, any electrical appliance switching ON/OFF and all unnecessary electrical appliances turn off automatically by counting person if no person is present. The work is constituted with a microcontroller, IR sensors, and HC05 Bluetooth module for confirming the wireless environment using Bluetooth technology and a software interconnection is also shown for proper using. Finally, this project eliminates the traditional method of controlling AC appliance and the process of individual involvement which makes life easier and faster in daily households need for everyone.**

Keywords— **ATmega8, Microcontroller, Android APP, IR sensor, Zero crossing detector, Relay, TRIAC Opto-coupler, Bluetooth module, LCD Display, Lamp**

I. INTRODUCTION

With the development of technology, the behavior of human's life is changing very fast. Nowadays Android phone is one of the most commonly used devices around the world. its demand is increasing with time and because of its' multi-dimensional functionalities and most enhanced applications. Today home automation is one of the most important application of Bluetooth technology. There are some objectives involved when designing a home automation system. Our project provides high security, electrical safety, power savings and reliability.

A speech recognition system where speech detection device is a mobile device's microphone [1]. This system can only turn on/off the home appliances, not voltage control. Besides vocabularies, multiple users and noisy environments are the major factors that affect the performance of the system. M.Preethi and R.Dharmalingam implemented a fan controlling system using an Android APP through wireless Bluetooth [2]. It only performs single work and operates DC fan but our system comprises multiple approaches. M. N. Azni et al. proposed Android Application based home automation [3] which controls different home appliances through wireless using Raspberry Pi but this project is costly. Our project proposes a low-cost solution. Shah Modi designed an RF-based home automation system which can do on-off of 15 devices using a single remote [4]. This system only controls the speed of a DC motor but not an AC Motor. Neng has proposed, and architecture for home automation [5] where the system was based on a dedicated network. This system only helps how to compose home automation problems at the software stage and no hardware stages were considered. Yavuz and Hasan [6] introduced a telephone and PIC Microcontroller based remote control system. Here the operation is performed using a dedicated telephone line not via Bluetooth technology. Ahammed presented DTMF (Dual Tone Multi Frequency) based Home appliances control [7]. DTMF based system can control electrical appliances from long distance but two mobile handsets (one for calling and one for receiving the call) are required and there will be call charges imposed from the service provider to the individual for each call. Sriskanthan introduced a Bluetooth technology based home automation system that can control home appliances from a PC using [8]. However, the system cannot be interfaced with the cell phone. A cell phone based home automation through Bluetooth communication which allows the user to monitor and control (only ON/OFF control) but not controlling AC voltage, has been discussed in [9]. In [10] S. Dey et al. and in [11] K. Mandula et al. proposed internet of thing (IOT) based home automation that can operate electrical appliances remote place through internet but its response is very slow and always must have an internet connection.

Considering these matters, we designed a system that can mitigate the above drawbacks at the same time it has been characterized with Multi-function.

II. HARDWARE ARCHITECTURE AND IMPLEMENTATION

The system hardware consists of three main parts such as (a) Password based access control. (b) AC voltage controlling and any electrical appliances switching (ON/OFF control) using Bluetooth communication of android (c) IR based person counter circuit.

A. Password based access control using Android APPS

This part adopts an HC05 Bluetooth module, MC, L293D motor driver & DC motor and other indicators electronic component. Here HC05 Bluetooth module that is interfaced

to Microcontroller is used to receive the data sent from APP. Any data is sent from the Android APP for different objectives to the Microcontroller through USART using Bluetooth communication. When a user sends a correct password, the green LED turns ON and the MC allows the user to enter into a secured area, whereas on sending a wrong password, the red LED & buzzer turns ON and the user will not be allowed in the secured area. The instructions of every step will show on the LCD display so that a user can easily understand the whole system. The Block diagram of this part is given below:

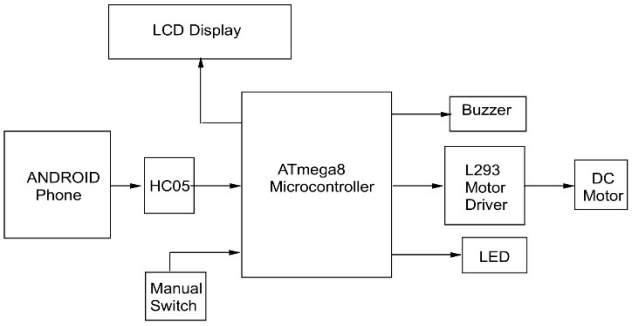


Fig.1. Block diagram of Android based access control system

B. *Android based AC voltage and electrical appliances control*

This part adopts a Zero crossing detector, MCU, Power circuit and other appliances controlling circuitry. The circuit diagram of AC voltage controlling is shown in fig.2.

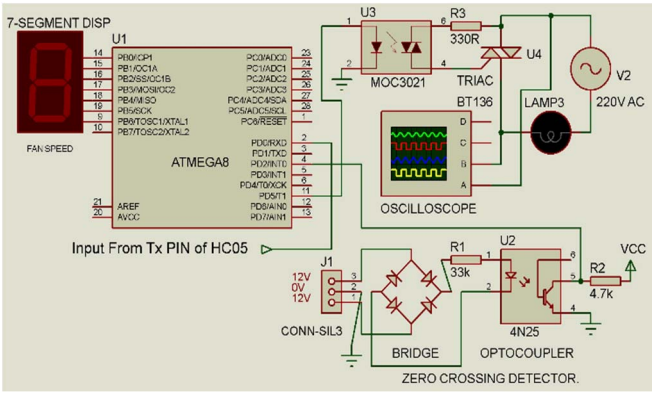


Fig.2. Circuit diagram of AC voltage control mechanism

Zero crossing detector which is made using an MCT2E optocoupler, bridge rectifier, and two resistors, produces very short duration positive pulses at each zero position of the sine wave input AC signal. These pulses are fed to the INT0 pin of MC through the optocoupler. The interrupt is produced at each zero position by the short duration. After that, the microcontroller generates control firing pulses per interrupt based on the receiving slider value of android app using a specified pin of output port which triggers the TRIAC as per requirement. TRIAC is a power electronic device that provides controlled output load power when triggered through the gate by using a phase angle control technique. The RMS value of AC voltage is controlled by the variation of firing angle (phase delay) of the TRIAC. Here MOC3021 is used for the isolation purpose between power circuit and control circuitry. The controlled output voltage is displayed in seven segment display from 0 to 9 corresponding the varied output ac voltage. Any electrical appliances on /off control mechanism is shown in fig.3.

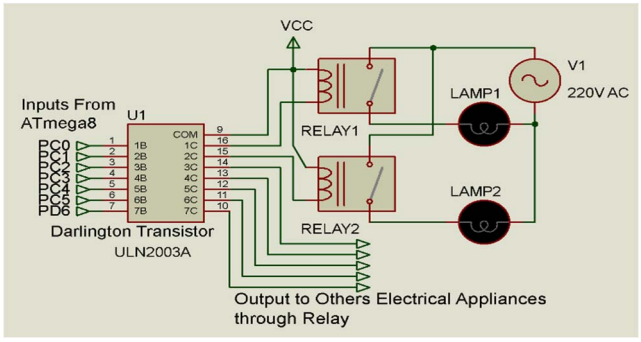


Fig.3. A connection of electrical appliances with ULN2003A via Relays

Here ULN2003A Darlington transistor which is connected directly with digital output pin (PORTC & PORTD) is used to operate the relays for providing required current.

C. *IR based person counter*

The arrangement of person counting is shown in fig.4 where two IR sensors are used to detect the presence of a human. The whole counting mechanism is performed by a microcontroller.

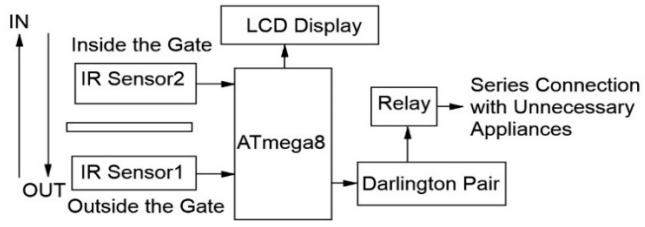


Fig.4. Block diagram of person counting mechanism

When a person enters the room, the counter value is increased, whereas on leaving, the counter value is decreased. If the counter value is zero (i.e. no person presents in the room) all unnecessary electrical appliances turn off automatically. If the counter value is greater than zero, the system will run in normal operation.

III. SOFTWARE DESCRIPTION OF THE SYSTEM

Here we have designed an android APP using online based open source android app inventor platform which is provided by Google, and now maintained by the Massachusetts Institute of Technology (MITC). App interface of Password based access control is shown in fig.5 and App interface of AC voltage and electrical appliances control is shown fig.6.

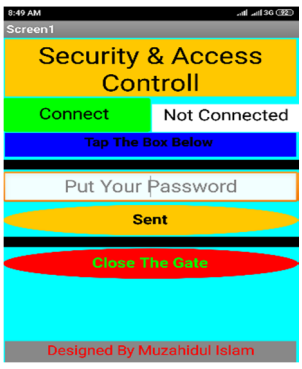


Fig.5. Screenshot of Password System APP interface

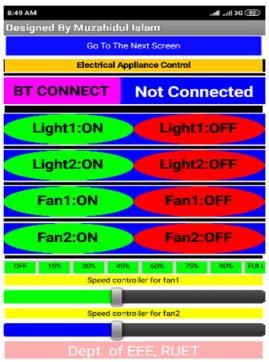


Fig.6. Voltage & Appliance control APP interface

when any push button on APP interface is pressed a specific ASCII value is transmitted to the MC through Bluetooth for the desired operation.

For Microcontroller programming, C language has been used for software design of the system. This program is written in Programmer's Notepad. The program has been written for three distinct output: Password system, Appliances control and Person counting. Fig.7 shows the basic flow chart of a Password system.

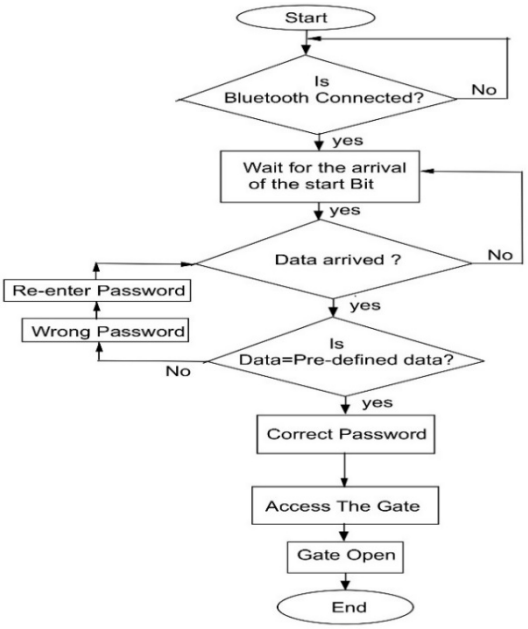


Fig.7. Flowchart of password based security

A password has been set in the MCU by software, which is 9 characters in length. When a user sends a password from the android app through Bluetooth to the MC, if this password matches the pre-defined password, Access will be granted otherwise denied. AC voltage controlling action in a flowchart is shown in fig.8

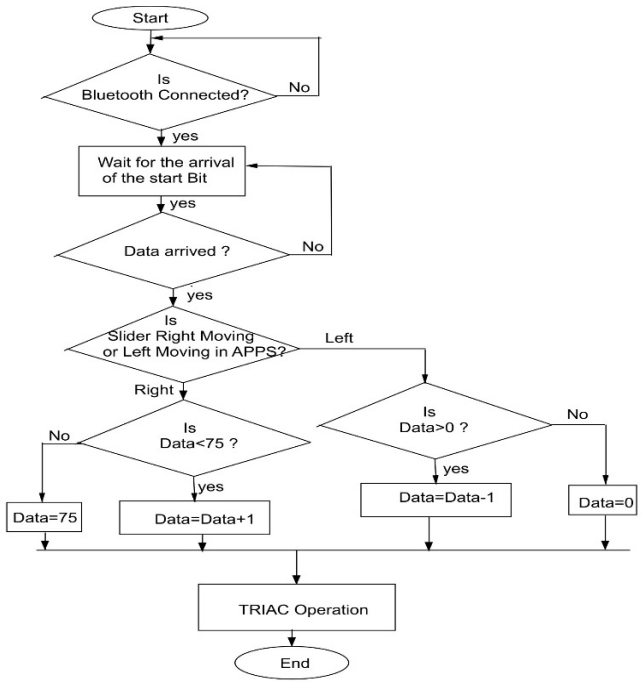


Fig.8. Flowchart of voltage controlling action

For controlling of AC voltage (i.e. a variation of RMS value of AC voltage), the Slider Thumb-position on Android APP interface is moved left or right. As a results, a changing ASCII value from 0 to 75 which is sent to the MC through Bluetooth communication is accumulated UDR buffer register. The UDR value is put in OCR2 through a globally defined variable for calling TIMER2 interrupt in every maximum value of OCR2, and MC provides control firing pulses that trigger the TRIAC via port pin. Then TRIAC provides a controlled output RMS voltage as per requirement.

For switching an electrical appliance, a specific character is sent by pressing a push button on the Android APP then the corresponding appliance is switched (ON/OFF).
For counting person, if the IR sensor1 detects the human first next IR sensor2, the counter value is added by 1. But if the IR sensor2 detects the human first next IR sensor1, the counter value is subtracted by 1. The counter value shows in the LCD display. If the counter value is greater than zero, one output pin of MC that goes to high turns on a defined relay. As a result, all unnecessary appliances is operated in normal mode. If the counter value is equal to zero, the same output pin goes to low, which turns off the all unnecessary appliances

IV. RESULTS AND DISCUSSION

Fig.9 shows the simulation result of the voltage controlling action in Proteus oscilloscope window. The controlled output curve shows the 10% voltage of input voltage. For the 10% output voltage, the firing pulses trigger the TRIAC at 9 MS delay from zero crossing of the input sine wave. Similarly, for 60% and 90% output voltages, the firing pulses trigger the TRIAC at 4 MS and 9 MS delay from zero crossing of input sine wave respectively. Fig.10 and fig.11 show the output curves of 60% and 90% voltage of input voltage respectively.

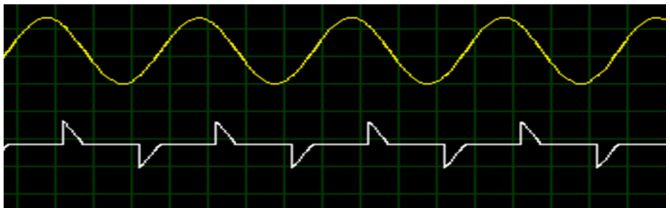


Fig.9. Input and 10% output voltage wave shape

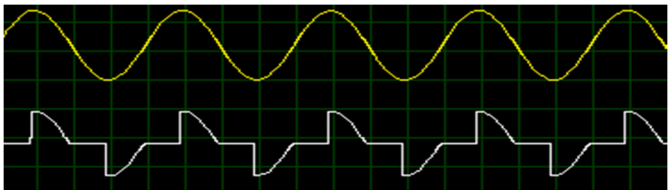


Fig.10. Input and 60% output voltage wave shape

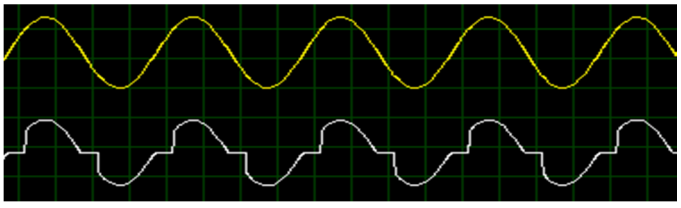


Fig.11. Input and 90% output voltage wave shape

The practical results of the whole project are shown in fig.12 and fig.13. Here two display has been used for password interfacing and person counting. Two light has been used to turn on/off and the other two light change the intensity which represents the controlling of the AC voltage. In the left side of fig.12, a DC motor attached model of a door is included. When the user password matched with the predefined password, the door opened and automatically closed after a predefined time. When the user password did not match with the predefined password, the system did not allow the user. When only the counter value was zero (no person present in the room), all unnecessary electrical appliances automatically turned off. Otherwise, the system is running in normal mode.



Fig.12. Complete setup with Test result

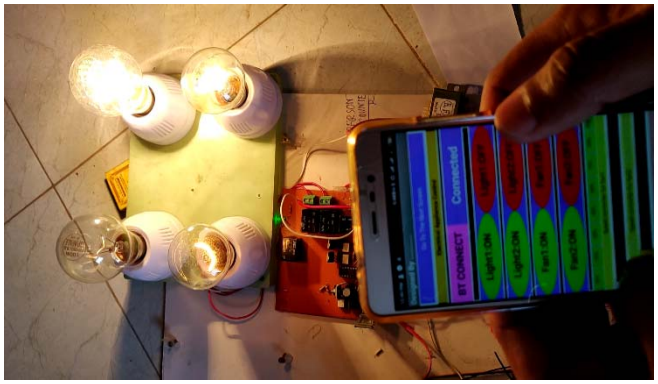


Fig.13. practical result of AC voltage controlling using Android APP

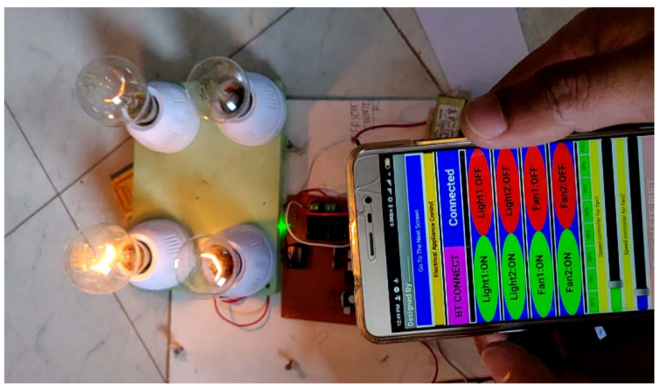


Fig.14. Switching ON/OFF and AC voltage controlling using Android APP

In fig.13 the lamp intensity which represents the controlling of the AC voltage is being controlled. When the slider position on APP interface is moved right or left, the controlled output voltage changed the intensity of light smoothly. In fig.14 one lamp turned OFF by pressing OFF marked push button and the intensity of other lamp decreased by moving the slider position to the left side.

V. CONCLUESION

In this paper, we have described a design and implementation of stable, flexible, safe and non-contact solution to the home automation. This system uses an Android APP to control the AC voltage (speed of induction motor) and the electric relay as well as it operates the password based access control through Bluetooth communication. The complete setup is successfully exposed to home appliances by controlling the intensity of light and switching on/off. This APP can control multiple numbers of appliances and the number of controlling appliances can be increased without any external circuitry. It provides high security and modernization of home automation. This system can reduce electrical hazard and misuse of electrical power. The work presented in this paper can further be extended for commercial, industrial applications and educational institutions.

REFERENCES

[1] K. Okokpujie, E. Noma-Osaghae, S. John and P. C. Jumbo, "Automatic home appliance switching using speech recognition software and embedded system", *2017 International Conference on Computing Networking and Informatics (ICCNI)*, Lagos, 2017, pp.1-4.

[2] M.Preethi and R.Dharmalingam "Controlling Fan Using Smart Android Device Via Wireless Bluetooth Microcontroller", *International Journal of Advances in Computer and Electronics Engineering*, Vol. 02, January 2017, pp. 24 -29.

[3] M. N. Azni, L. Vellasami, A.H. Zianal, F.A. Mohammed, N.N. Mohd Daud, R. Vejasegaran, N. W. Basharudin, M. Jusoh, K.N.F. Ku Azir,P.L. Eh Kan, "Home automation system with android application", *2016 3rd International Conference on Electronic Design (ICED)*. August 11-12, 2016, pp. 299-303.

[4] J. Shah, B. Modi and R. Singh, "Wireless home appliances controlling system," *2014 International Conference on Electronics and Communication Systems (ICECS)*, Coimbatore, 2014, pp. 1-6.

[5] Neng- Shiang Liang; Li-Chen Fu; Chao-Lin Wu. "An integrated, flexible, and Internet-based control architecture for home automation system in the internet era", *IEEE International Conference on Robotics and Automation*, Vol. 2, pp.1101-1106, 2002.

[6] E. Yavuz, B. Hasan, I. Serkan and K. Duygu. "Safe and Secure PIC Based Remote Control Application for Intelligent Home", *International Journal of Computer Science and Network Security*, Vol. 7, No. 5, May 2007.

[7] Ahammed, M. and Banik, P. (2015), "Home appliances control using mobile phone", *2015 International Conference on Advances in Electrical Engineering (ICAEE)*.

[8] N. Sriskanthan and Tan Karand. "Bluetooth Based Home Automation System", *Journal of Microprocessors and Microsystems*, Vol. 26, pp.281-289, 2002.

[9] M. O. A. Bader , I. M. A. Iman, H. M. A. Mahdi, H. A. A. Sami and S.S. Mohamed, "Design and Implementation of a Reliable Wireless Real-Time Home Automation System Based on Arduino Uno Single-Board Microcontroller", *International journal of control, Automation and systems*, vol. 3, no. 3, pp. 11-15, 2014.

[10] S. Dey, A. Roy and S. Das, "Home automation using Internet of Thing", *2016 IEEE 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*. 20-22 Oct. 2016.

[11] K. Mandula, R. Parupalli.,C. Murty, E. Magesh and E. Lunagariya (2015). "Mobile based home automation using Internet of Things(IoT) "*2015 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*. 18-19 Dec. 2015

An Effective Combination of Microcontroller and PLC for Home Automation System

Mohammad Hanif
Dept.of Electrical & Electronic Engineering
Chittagong University of Engineering & Technology
Chittagong, Bangladesh
hanif.mohammad2010@gmail.com

Nur Mohammad
Dept. of Electrical & Electronic Engineering
Chittagong University of Engineering & Technology
Chittagong, Bangladesh
nur.mohammad@cuet.ac.bd

Bijoy Harun
Dept.of Mechanical Engineering
Chittagong University of Engineering & Technology
Chittagong, Bangladesh
bijoy.harun5@gmail.com

***Abstract*—In this paper, a smart home automation system is proposed to control home appliances and to make a home safer using electronic controller. Also, Grid-Eye sensors are used to detect the presence of the occupant in the home upon which the activation of any appliance depends. As a result, all the electric instruments of the home run automatically when necessary, and the residents do not need to turn on/off of any electric appliances like air conditioner, fan, light, water heater or any other home appliances. Also if a person wants to make on/off of any utilities, he can easily make it by using android apps or by using a wireless network which is interfaced with the system. In addition to home automation, when any mishap happens in the home, GSM sends SMS to warn people about the accident. If a person wants to know the status of his home utility system, when he or she out of the home, he can also use GSM technology. Proteus simulation software is used to test the effectiveness of the proposed designed system. To control large electric machines like water pump, elevator and generator, PLC module is used. In this case, we also consider different conditions to run large electric machinery automatically. This designed system provides more convenient and flexible control for a smart home or a building. If implemented, the model will be beneficial for the promotion and popularity of smart home automation system.**

Keywords— Grid-Eye sensor, microcontroller, GSM module, Arduino IDE software, Proteus simulation software, PLC.

I. INTRODUCTION

Smart home system and home safety are vital issues for our modern life. Electronic advancement moves us to this development. But in most of the smart home system, the presence of people is ignored or the residents need to turn on/off electric appliances used in their home. The smart home system is an automated home with control devices associated with lighting, heating, air conditioning, ventilation, and security system [1]. Sensor and microcontroller are the main parts of the smart home system. In this case, sensors collect data from the surroundings and send electric signals to the microcontroller. The microcontroller processes this signal, and depending on the installed program in the microcontroller, necessary actions are taken to control the utilities connected to it. So, we can say that microcontrollers are the central axis of this system, and are sometimes called "gateways". The microcontrollers are usually interfaced with a tablet, mobile phone, or computer; and the network connectivity of the systems is managed by IoT [2]. Using this system, the controlling of all the electric appliances is possible which makes a home automation system. Android apps are also used to build home automation since Wi-Fi or Zigbee is available nowadays.

When designing a smart home system, the system should be flexible, affordable and scalable for integrating new devices into the system [3]. Moreover, the smart home system should be tiredness and boredom reducer to us. An important concern of the smart home system is the reduction of energy consumption and saving money for poor people [4].

There are many smart home research works in the online. In this paper, sensors and microcontroller are considered to control light, air conditioner or other home appliances only when they need to operate. In other words, when people are not present in the home, the light or air conditioner do not need to operate whatever the condition of the surrounding. For this reason, Grid-eye sensors with other sensors are used to detect the presence of the people. The surrounding conditions are only considered by the microcontroller during the presence of an occupant. When the output of danger warning sensors used, in the system, upon a critical level, the microcontroller does not consider the output of Grid-eye sensor and sends a warning to the people. As a result, the home will also be saved. In this designed home automation system, power will be reduced at the lowest level since utilities only run when people present in the home.

In case of controlling water pump, generator and elevator of a building, we interface PLC with the system since it is more reliable in this case. Also, we consider different conditions like the water level in case of the water pump and the presence of people in case of elevator and generator to run these electric appliances.

II. LITERATURE REVIEW

A smart home is the network of physical devices that provide electronic sensors, software, and network connectivity with each other inside a home [1]. Sometimes we consider a significant feature of the smart home is remote controlling [5]. There are numerous research papers related to smart home or home automation. A query on Web of Science, Science Direct, and IEEE Explore found almost 1798 papers in 2017 related to smart home [1]. In 2006, Davidoff et al. [6] published a paper related to the smart home and emphasized on control by end-user programming. Since 2010, researchers have been analyzing smart home based on IoT applications using many approaches. Stankovic [7] discussed the notion of smart car, smart homes, smart phones, and smart devices. An imagination of future intelligent home has been presented in some papers [2][8] where Zigbee (Wi-Fi) technology was considered. For managing power consumption, Zigbee-based smart home system is designed and implemented [9]. The smart home system is also designed by using android apps where

wireless Bluetooth system was used to assist disabled and elderly people [10]. Analyzing an original architecture and integrating it in a system, a smart home system is designed where IoT used in a paper published in 2013 [11]. IoT based home automation is versatile and popular. Numerous IoT based home automation systems are designed using many sensors to control electric appliances and to decrease the energy consumption [12]. Home automation system was designed using Global System for Mobile Technology (GSM) by Teymourzadeh and Rozita [13] in 2013 and their designed system was tested on four loads and its accuracy was 98% [13]. Smart traffic light control system was also designed where the prototype was presented to control traffic congestion using PIR sensor to calculate the number of vehicle on the road even the number plate of a vehicle is captured in this system [14]. Using wireless sensors networks (WSNs) interfaced with IoT to build smart home system is also available nowadays [15][16]. To connect objects like mobile phones, internet televisions, and sensors to the internet where the devices build new forms of communication, IoT is widely used [17]. But the main problem of this system is the cost of the wireless sensors, and the Wi-Fi network needs to keep connected all the time for communication. C. Barz et al. [18] used TIA Portal V12 and Siemens PLC to design smart home automation system where lighting, heating, irrigation, security system were considered. Raspberry Pi-based home automation system is also available nowadays where Raspberry Pi is used as a communication gateway between different home appliances and mobile [19]. To restrict the home automation for certain people, the command of user upon which the activation of home automation depends is proposed in some papers. Many firms assert that a smart home automation system can save 30% energy usage without comprising any comfort of the residents [20]. Home automation system balances demand-supply of power inside a home. As a result, it will reduce energy consumption and increase efficiency [21]. Using robotic system in a large electric machine like generator will not only save energy consumption but also reduce carbon emission in the environment and will help us to reduce the major global concerns of climate change and energy security [22].

III. Methodology

In this paper, we consider a home automation system to overcome many practical difficulties. Among many difficulties, the first one is the absence of residents of a home. Smart home system or home automation is always designed based on the need of people such as when a light or air conditioner will start. When the temperature of a room increases above a certain level or the light intensity of a room decreases below a critical level, the air condition or light turns on automatically. It is controlled by the microcontroller. But, most of the cases, the temperature of a room rises high during the day when people usually go to their office and children go to their school. In this paper, it is considered. We not only think about the fan or air conditioner or light, but we also think about the elevator, motor and other utility systems. Also, safety is added to our design approach. Automatic motor running during the empty state of a water tank and automatic elevator controlling when people need to use it is also considered in this paper.

We have considered our home automation system in two parts. One is an interior home appliances controlling system using a microcontroller and the other one is exterior i.e. building electric appliances controlling system. In the interior home appliance system, we have considered microcontroller to control different appliances using Grid-eye sensors. On the other hand, for controlling building electric appliances, we have considered PLC.

A. Interior System Design

1) System architecture: We have used a light intensity sensor BH 1750 to interface it between microcontroller and light bulb. BH 1750 sensor is set in all the rooms of the home to control the on/off of the light. When the light intensity of a room decreases below a critical level the microcontroller sends a signal to turn on the light bulb. But when there is no person in the room, the microcontroller also sends a signal to turn on the light bulb which is not a desirable design for us. So, we have also interfaced a Grid-eye sensor to detect the presence of people in the room. The microcontroller considers two cases; one is the presence of people in the room and the other is the condition of the room. As a result, our system will be smarter and will save fuel or energy costs. Moreover, safety systems are also included in this design. Manually starting the/stopping system using android apps or using a smartphone can be easily interfaced with this system to control the home appliances when anyone wants to turn on/off home appliances if necessary.

The system consists of a microcontroller which is interfaced with a GSM module. All the utilities are controlled by this microcontroller. Also, some sensors like gas/smoke sensor, temperature sensor are used to keep home safe, and if any mishap happens, then SMS is sent to the user's mobile indicating the temperature or gas presence in the room. In the system, we also consider auto open/close of the doors and windows of the rooms. HC-SR501 Infrared PIR Motion sensor is proposed to use in this case. Also, power generator, elevator, water heater, motor all are considered to automatically on/off upon the presence of people in the home.

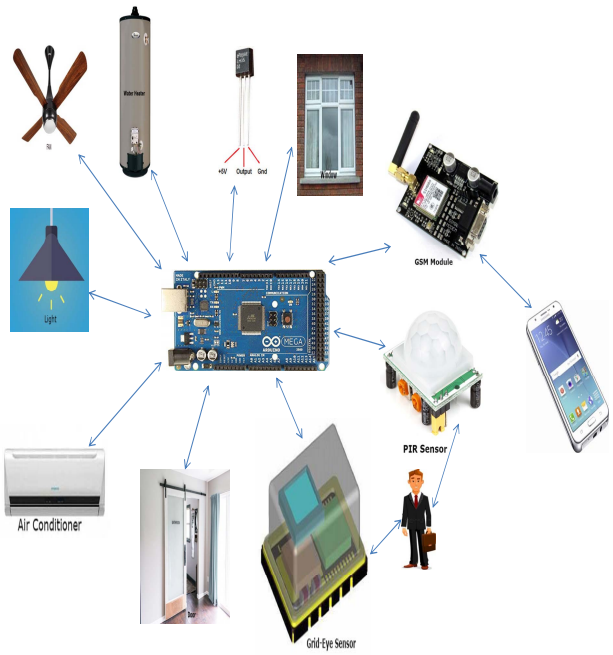


Fig.1. System architecture (inside a flat).

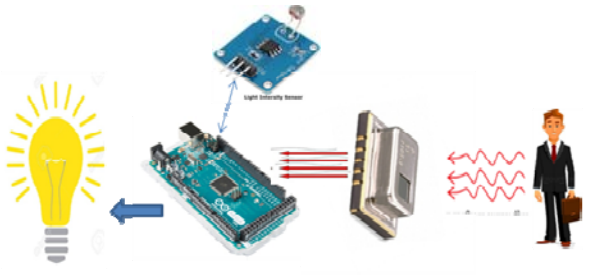


Fig. 2. Utility (Light) active.

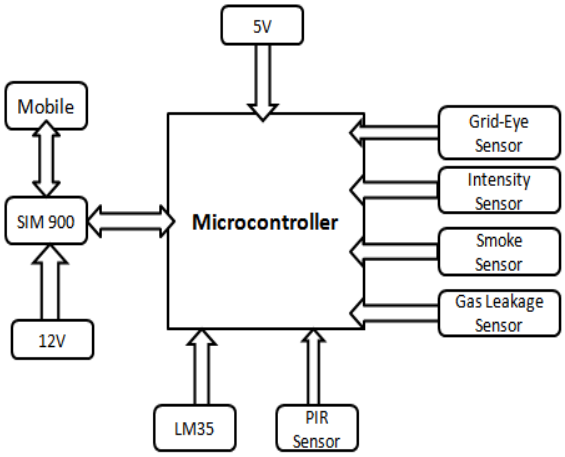


Fig. 3. System Diagram.

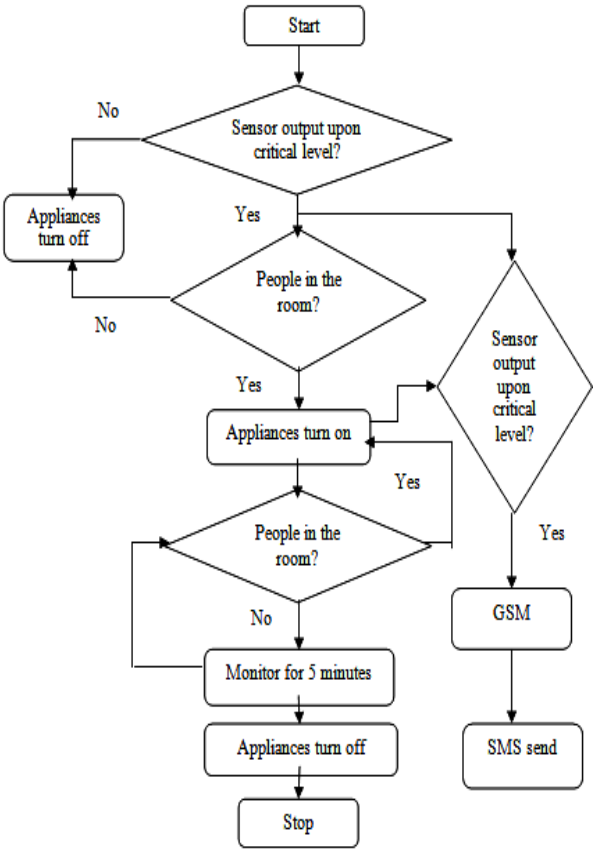


Fig. 4. Flow chart of the system.

2) 4-Room Model: We have implemented our smart home system by considering a 4 room model. In this model, we have used only one microcontroller which is put in the centre of 4 rooms. Four LM35 temperature sensors, four Grid-eye sensors are considered in this 4-room model. All of the 8 sensors are connected to the microcontroller. The microcontroller is programmed using Arduino IDE software. The lower limit of the LM 35 sensors is set at 18ºC and the upper limit is set to 40ºC in the microcontroller. The lower limit is on the condition of a fan, and the upper limit is for fire warning. The Grid-eye sensors are used to detect the presence of the people in the rooms.

When a person enters in room 1, the fan of this room switched on automatically since the temperature was 28ºC of the room on that day. The fan continues to remain on condition till the person present in the room. The fans of other rooms are in off condition since no one present in these rooms. When the person left room 1, the fan keeps running, and then it will stop after 5 minutes (Condition is set in the microcontroller). If the person now enters in room 2, the fan of room 2 becomes on, and it will also take 5 minutes to stop after leaving of that person. The same result will be found for the other 2 rooms. As a result, the person does not need to use android apps or Bluetooth or internet, in this case, to turn on/off of any electric appliances.

3) Programming the microcontroller: For the 4-room model, we have programmed the Arduino AT Mega 2560 microcontroller using Arduino IDE software. The programming is compiled following the flow chart of Fig.5. The four LM 35 temperature sensors have been connected with four analog pin A1, A2, A3, and A4 of the microcontroller. Also, four HC-SR501Infrared PIR Motion sensors instead of Grid-eye sensors are used which are connected with digital input pin 11,12,13, and 14. Both the LM 35 connection pins and PIR sensor connection pins are defined in the programming. When a person leaves from a room, the pin output of PIR becomes low. Since we set a delay (300000), the microcontroller monitors 5 minutes and makes the appliance turn off after it. We use "if ….else" loop for input to the microcontroller from LM35 sensors, and the upper critical value of temperature sensors is set to 40 ºC (inputPin >=40). A SIM 900 GSM module is interfaced, and AT+CMGF=1 command is used for sending SMS when LM 35 output goes over the upper critical level. The mobile number is set by using AT+CMGS command.

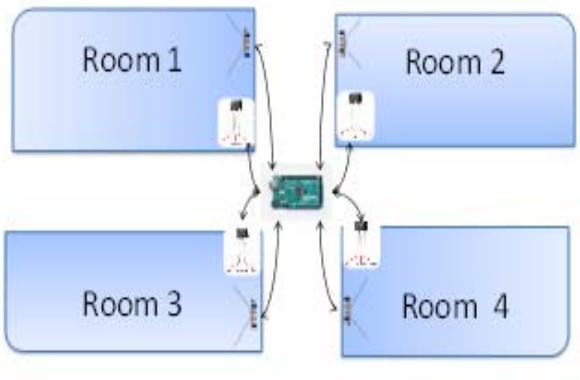


Fig.5. Four-room model

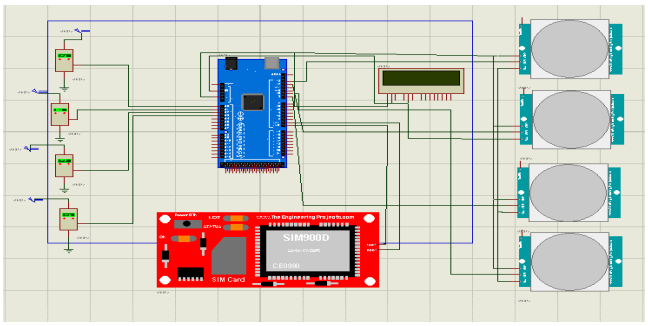


Fig. 6. Proteus simulation of the 4 Room Model.

4) Simulation: Using Proteus simulation software, we have simulated 4-room model which is shown in Fig.6. Also, fire warning to the mobile is simulated by using SIM900 GSM module in the simulation.

B. PLC based Exterior Automation

In our designed system, exterior electric machines such as water pump, generator, and elevator are controlled by using PLC. Though PLC is used especially in the industrial automation system, in the high rise building PLC can be used because of its reliability. We think about the automatic control of water pump, elevator, and generator such that our electricity cost minimizes.

a) Water Pump Control: For controlling water pump, two water level sensors (I1 and I2) are utilized. Ultrasonic sensors (HC-SR04) can be used to determine water level in the tank [23]. I1 is the upper water level indicating sensor and I2 is the lower water level indicating sensor. When the outputs of the two sensors are low, the water pump starts, and when the outputs of the two sensors are high, the water pump stops. In this case, when I1 is high, I2 becomes automatically high. Hence, when I1 becomes high, the water pump stops. Data transmission to control the water pump was provided by the PLC. A flow chart is shown in Fig.7. Simulation of this flow chart is shown in Fig. 8 where using Siemens LOGOsoft Comfort software is used.

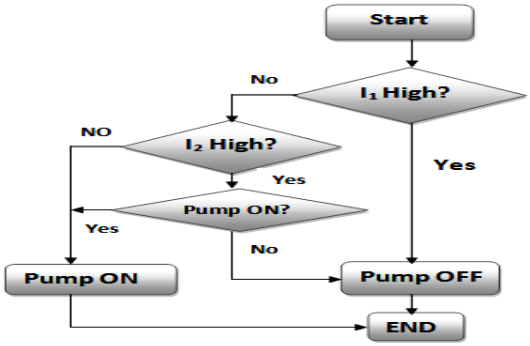


Fig.7. Flow chart of water pump control.

Table I . Truth Table Of Pump On/Off.

I1	I2	Q1	Q2	Pump (On/Off)
0	0	1	0	On
0	1	1	0	On
1	1	0	1	Off
0	1	0	1	Off
0	0	1	0	On

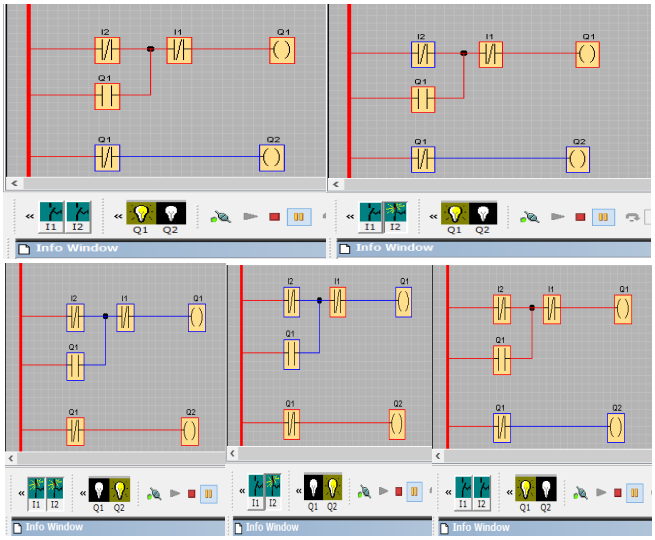


Fig. 8. The water pump on/off simulation.

b) Elevator and Generator controlling: The elevator and generator of a building will only start if people present in the elevator. To demonstrate this case, we only consider the on/off of a generator connected to the elevator during load-shedding. We have implemented our desired design using a load sensor connected to the elevator. The load sensor is used to detect the presence of a person in the elevator. When the load (weight) is applied to the base of the elevator the output of load cell will be high.

The elevating system is operated and controlled by stored programs in the PLC and is designed as an on-off controller. The operation of a power generator connected to an elevator is also controlled by PLC. Here we will design a controlling system of elevator and generator connected to PLC.

In our designed system, several cases are considered to control the generator connected to the elevator when grid electricity is not present.

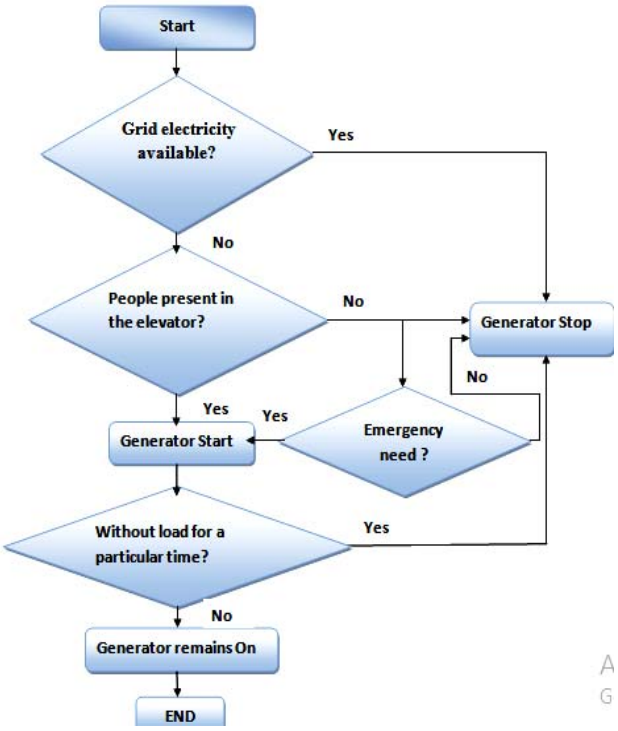


Fig.9. Flow chart of controlling elevator and generator connected to it by considering the presence of people.

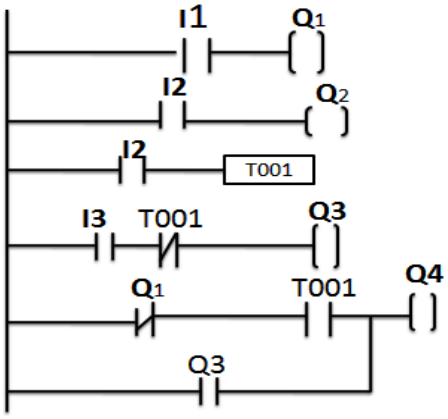


Fig.10. PLC ladder diagram for controlling generator connected to elevator according to the flow chart of Fig. 9.

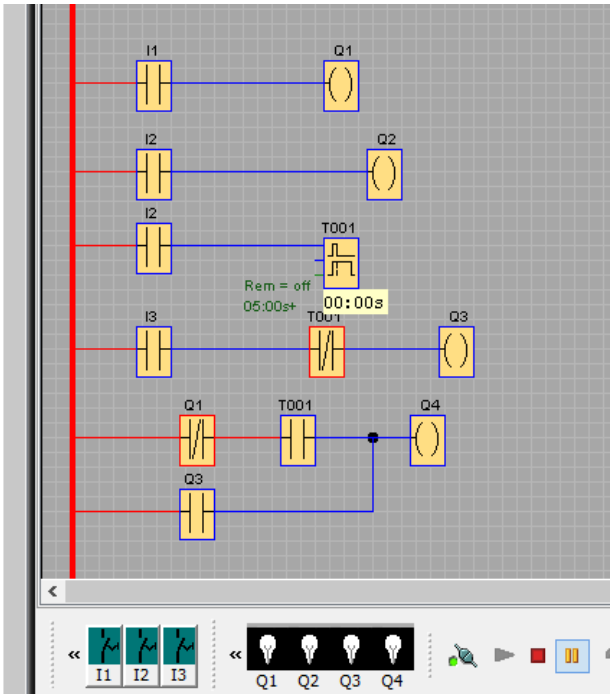


Fig.11. Simulation of controlling generator connected to elevator depending on the presence of person.

The ladder diagram using "Siemens LOGOsoft Comfort" software for the flow chart of Fig.9 is shown in Fig.12. In the ladder diagram,

- I1 indicates whether electricity supplying from the grid or not.
- I2 indicates whether people (load) present in the elevator or not.
- I3 indicates emergency needs (when anyone needs to run the elevator when both electricity and people (load) in the elevator are absent).
- T001 is an off-delay timer that keeps a generator running for 10 second after people leave the elevator (during load-shedding). As a result, another person can ride on the elevator within this time.
- Q1 output indicates the presence of electricity.
- Q2 output indicates the presence of people (load) in the elevator.
- Q3 indicates emergency need.
- Q4 output indicates the on /off of the generator connected to the elevator.

IV. Result and Discussion

Using this automation system, the utilities of the home automatically turn on/off considering the presence of people in a room and the condition of the environment. As a result, people need not concern about on/off of any utilities even when they out of a home. We have implemented a prototype of a smart home for 4-room model using microcontroller, sensors and SIM900 GSM module. Also, automatically opening system of the doors and windows have been interfaced with our designed prototype. It is possible to implement this system for a multi-room house even in the whole building. Moreover, the safety sensors will give important warnings by sending SMS to the mobile phone about any danger in the home via the GSM module. By using smoke and gas sensor and GSM module, we have implemented and verified safety system of the designed home automation. Also, controlling of an elevator is implemented using a Siemens PLC module and TIA PORTAL v12 software. In the future, we will investigate transactive control of demand response for flexible loads which can be used for energy efficiency programs [24], [25].

V. Conclusion

The main advantages of the proposed system are the low cost and user-friendliness. Another main advantage of this system is flexibility and automation. Though there are many smart home system using IoT, Bluetooth, Zigbee, cloud-based, smartphone-based, Tablet or PC based, among them this designed system is simple, low cost, energy saving and flexible for common people. People will feel relax and free from the manual switching system. Even people need not use any mobile apps to turn on/off of any utilities in normal situations. We have used PIR sensors instead of Grid-eye sensors in the Proteus simulation to detect people in a room since the library file for Grid-eye sensors is not available. We do not make an android app to control the utilities (turn on/off) when people want to change the pre-set condition of on/off of any utilities in any particular situation. Many papers have already published to control home appliance using android or wireless network, and people can use any convenient system like an android app or Bluetooth or IoT in this case.

References

[1] M. Alaa, A. A. Zaidan, B. B. Zaidan, M. Talal, and M. L. M. Kiah, "A review of smart home applications based on Internet of Things," *J. Netw. Comput. Appl.*, vol. 97, pp. 48–65, 2017.

[2] X. Ye and J. Huang, "A framework for cloud-based smart home," *IEEE Int. Conf. Comput. Sci. Netw. Technol. (ICCSNT), 2011*, vol. 2, pp. 894–897, 2011.

[3] R. Piyare and T. M, "Bluetooth based home automation system using cell phone," *IEEE15th Int. Symp. Consum. Electron. Singapore, 2011*, pp. 192–195, 2011.

[4] K. Moser, J. Harder, and S. G. M. Koo, "Internet of things in home automation and energy efficient smart home technologies," in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, 2014, vol. 2014–Janua, no. January, pp. 1260–

1265.

[5] Y. Yang, Z. Wei, D. Jia, Y. Cong, and R. Shan, "A cloud architecture based on smart home," in *2nd International Workshop on Education Technology and Computer Science, ETCS 2010*, 2010, vol. 2, pp. 440–443.

[6] S. Davidoff, M. K. Lee, C. Yiu, J. Zimmerman, and A. K. Dey, "Principles of Smart Home Control," in *Springer*, 2006, pp. 19–34.

[7] S. John A, "Research directions for the internet of things," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 3–9, 2014.

[8] A. ElShafee, K. H.-W. academy of science, undefined engineering, and undefined 2012, "Design and implementation of a WIFI based home automation system," *researchgate.net*.

[9] Y. Gao and D. Shang, "The Research of Home Intelligent Power System Based on ZigBee," *3rd Int. Conf. Consum. Electron. Commun. Networks (CECNet), IEEE*, pp. 703–706, 2013.

[10] R. A. Ramlee, M. A. Othman, M. H. Leong, M. M. Ismail, and S. S. S. Ranjit, "Smart home system using android application," in *2013 International Conference of Information and Communication Technology (ICoICT)*, 2013, pp. 277–280.

[11] Y. Jie, J. Y. Pei, L. Jun, G. Yun, and X. Wei, "Smart home system based on IOT technologies," in *Proceedings - 2013 International Conference on Computational and Information Sciences, ICCIS 2013*, 2013, pp. 1789–1791.

[12] H. Singh, V. Pallagani, V. Khandelwal, and U. Venkanna, "IoT based smart home automation system using sensor node," in *Proceedings of the 4th IEEE International Conference on Recent Advances in Information Technology, RAIT 2018*, 2018, pp. 1–5.

[13] R. Teymourzadeh, S. A. Ahmed, K. W. Chan, and M. V. Hoong, "Smart GSM based home automation system," in *Proceedings - 2013 IEEE Conference on Systems, Process and Control, ICSPC 2013*, 2013, no. January, pp. 306–309.

[14] B. Ghazal, K. Elkhatib, K. Chahine, and M. Kherfan, "Smart traffic light control system," *2016 3rd Int. Conf. Electr. Electron. Comput. Eng. their Appl. EECEA 2016*, pp. 140–145, 2016.

[15] H. Ghayvat, S. Mukhopadhyay, X. Gui, and N. Suryadevara, "WSN- and IOT-based smart homes and their extension to smart buildings," *Sensors (Switzerland)*, vol. 15, no. 5, pp. 10350–10379, 2015.

[16] S. Pirbhulal *et al.*, "A novel secure IoT-based smart home automation system using a wireless sensor network," *Sensors (Switzerland)*, vol. 17, no. 1, 2017.

[17] G. Kortuem, F. Kawsar, V. Sundramoorthy, and D. Fitton, "Smart objects as building blocks for the internet of things," *IEEE Internet Comput.*, vol. 14, no. 1, pp. 44–51, 2010.

[18] C. Barz, S. I. Deaconu, T. Latinovic, A. Berdie, A. Pop-Vadean, and M. Horgos, "PLCs used in smart home control," in *IOP Conference Series: Materials Science and Engineering*, 2016, vol. 106, no. 1, p. 012036.

[19] V. S. Gunge and P. S. Yalagi, "Smart Home Automation: A Literature Review," *Int. J. Comput. Appl.*, pp. 6–10, 2016.

[20] T. Hargreaves, C. Wilson, and R. Hauxwell-Baldwin, "Learning to live in a smart home," *Build. Res. Inf.*, vol. 46, no. 1, pp. 127–139, 2017.

[21] M. Nur and Y. Mishra, "The Role of Demand Response Aggregators and the Effect of GenCos Strategic Bidding on the Flexibility of Demand," *Energies*, vol. 11, no. 12, p. 3296, 2018.

[22] K. Ahmad, N. Mohammad, and M. Quamruzzaman, "Review on Frequency Adjustment for Power Systems with Grid Connected Wind Farm," *2019 Int. Conf. Robot. Signal Process. Tech.*, no. 1, pp. 617–621, 2019.

[23] M. Asadullah and K. Ullah, "Smart home automation system using Bluetooth technology," in *ICIEECT 2017 - International Conference on Innovations in Electrical Engineering and Computational Technologies 2017, Proceedings*, 2017, pp. 1–6.

[24] N. Mohammad and A. Rahman, "Transactive control of industrial heating–ventilation–air-conditioning units in cold-storage warehouses for demand response," *Sustain. Energy, Grids Networks*, vol. 18, p. 100201, 2019.

[25] N. Mohammad and Y. Mishra, "Transactive Market Clearing Model with Coordinated Integration of Large-Scale Solar PV Farms and Demand Response Capable Loads," *in Proceedings of the 2017 Australasian Universities Power Engineering Conference, AUPEC*, 2017, p. 6.

Subject Oriented Data Partitioning – A Proposed Data Warehousing Schema

[1] Abdul Moktadir, [2] N. M. Istiak Chowdhury
Department of Information & Communication Technology
Comilla University
Comilla-3506, Bangladesh
[1] am.moktadir@gmail.com, [2] swajibcse@yahoo.com

***Abstract*—Data warehouse is an inseparable part of modern business but the amount of data it stores is huge and growing so fast, the queries of the business leaders it processes have become very sophisticated as well. Data warehouse needs to response to those queries with high quality data having less response time but the major challenge for that is the scale of data and nature of the queries involved. It's very hard to access the required data set with minimum response time in data processing as the data warehouse needs to perform analytical operation on a large set of data analyzing sophisticated queries. This paper introduces a new data warehousing schema that constructs a multidimensional relation tree hierarchy, changing the traditional use of relational database, with partitioned database in its every level instead of any individual data. This schema allows data warehouse to partition and process a large data part by part in a small batch to minimize the response time and stores summarized data to increase data availability in query processing. Allowing easy and fast access to the required set of data, this schema improves the performance and the decision making ability of the data warehouse.**

Keywords—Data Warehouse, Data Warehouse Schemas, Data Partitioning, Data Warehouse Architecture, Data Quality, Response Time.

I. INTRODUCTION

Data warehouse technology has been a matter of interest and of paramount importance in today's business. The business communities and organizations are becoming more dependent on their ability to quickly access, easily use, efficiently maintain, quality and timely business information which they need to achieve their goals [1]. But the success of the data warehouse depends on how much the warehouse system is related to business intelligence. So, while data warehouse implementation, the system architecture should be designed and confirmed by the decision makers first [2]. Many data warehouse systems fail to fulfill its requirements to help in strategic decision making as the warehouses fail to understand the requirements of the business leaders.

The data warehouse is a repository of information collected from multiple operational sources, stored under a unified schema. Data in today's business are growing so fast making it very difficult to handle them and introducing many difficulties in data processing. The requirements of the business leaders have been complicated with the overgrowing business process and the larger data set [3]. So it is not just the requirements of the business leaders a warehouse has to deal with, it is also the size of the data stored in the warehouse that needs appropriate maintenance too. Efficient query processing is very important for any data warehouse system as the decision makers want the warehouse to answer complex queries with minimum response time. The warehouse stores a large amount of data and process much more complex queries, thus the response time is much higher. Most of the data warehouses are inefficient, consuming more resources and taking more query response time because of the volume of the data and given nature of the queries [4].

Data quality is another major concern for a data warehouse system and it plays an important role in business process [5] as the decision makers need appropriate data to take actions. The overgrowing data make it very difficult for a data warehouse to provide the expected data quality [6]. Data quality needs to be monitored, including timeliness and completeness of the data stored in the warehouse [1]. Poor data quality can have a severe effect on the effectiveness of the data processing operation but the problems of data quality are often being ignored in the process of data warehouse construction and utilization [7]. The quality of data is defined by its accessibility, availability, and acceptability to the users and should be retrievable with minimum response time. Many data warehouses fall apart just because they fail to maintain easy accessibility, enough availability, and acceptability of that over growing data to the users. Without easy access to enough data, a data warehouse is never going to be able to fulfill the needs of the business leaders.

After the introduction, the rest of the paper is organized as follows: Section II describes the aim of this work and section III describes the related work. In section IV, we propose a data warehousing schema where section IV(A) describes the methodology and section IV(B) point out a key challenge of the proposed schema followed by section IV(C) dedicated to the analysis of the schema. Finally, section V presents the conclusions and future work direction.

II. AIMS OF THE STUDY

This paper aims to introduce a data warehousing schema trying to overcome the difficulties of the data warehouse technologies using horizontal data partitioning in the warehouse design. Business organization needs the data warehouse to process certain set of data first to perform efficient analytical operations for fast and timely decision making [2] and certain queries in a data warehouse requires only a portion of data to be processed and the query response time can be minimized by using different portion of data for different type of queries [8].

This schema uses the data partitioning technique in data warehousing to reduce the query response time having fewer

data to access to process. We are going to apply the horizontal data partitioning to the fact tables where each horizontal partition will have a subset of tuples from the respective relation table. By reducing irrelevant data access and using parallel execution of OLAP query, horizontal partitioning will reduce the query execution time [8]. This schema will increase the accessibility and availability of data with additional attributes in relational tables to improve the data quality and to minimize the response time of the system considerably.

III. Related Work

There are three different schemas present in the warehousing technology to design a data warehouse: star schema, snowflake schema and fact constellation schema [9][10]. However, they are not that different from each other in designing a data warehouse. In star schema, one fact table shares multiple dimensions. Both Snowflake and fact constellation use the same concept as star schema except snowflake schema divide dimension into sub-dimension and fact constellation allows multiple fact tables to share the same dimension to reduce data redundancies. Therefore, technically speaking, those schemas are designed for one consideration only, fact table shares dimension tables to describe those facts. In these schemas, the size of the fact table is never considered but the over growing data is a major concern to the modern business. A huge amount of operational data are being stored within a very short time and these schemas are storing this huge data in a single fact table without considering the size of that table. Therefore, it has become very difficult for the existing data warehouse technologies to perform analytical operation on such a large data set and thus fail to provide the expected outcome.

Author(s) [8] described how data partitioning can be of use to maintain the large volume of data and, using partitioned data, the response time can be minimized. But they partitioned both the dimension tables and the fact tables, which results in a huge number of partitioned tables. But this schema will partition just the fact tables not the dimensions and those partitioned fact tables will be related to a specific dimension by some extra attributes. Those extra attributes will increase data availability in the dimension table and result in fast response to certain queries having quick access to require data.

IV. Proposed Schema: Subject Oriented Data Partitioning

This paper proposes a new data warehousing schema using the data partitioning technique on the large scale of fact data. Using horizontal partitioning, this schema stores the facts as small fragments of data and processes them separately. The partition is based on the records of a specific dimension and the choice of that dimension table used in partitioning the fact table plays an important role on the overall performance of the data warehouse system.

The facts are partitioned based on a specific dimension, not all dimension tables take part in the partitioning process. Some dimension tables just describe the facts as they are supposed to do. This schema doesn't just partitions the fact, it allows the warehouse system to hold a relational table, partitioned fact table, within an attribute of another relational table, the respected dimension. This schema also allows the inner relation to have its own structure and this way creates a relation tree hierarchy. In this schema, we are referring to those dimensions that partitions the facts as 'Parent table' and the partitioned fact tables as the 'Child table' of that hierarchy.

In this schema, the parent table has an additional attribute to its attribute set for every type of facts it partitions and the data of that attribute is an abstract value from its corresponding child table. The parent table may have several records and the number of parted fact tables will be as much as the records of that parent table. And as they all belong to the same attribute, they will have the same attribute set and same structure format for a specific type of fact.

A. Methodology

We have talked about how the data warehouse schemas design a data warehouse having dimensions describing the facts. Consider that the Sales, Ratings & Reviews are the operational fact tables and Product, Store, Location & Customer are dimension tables used to describe these

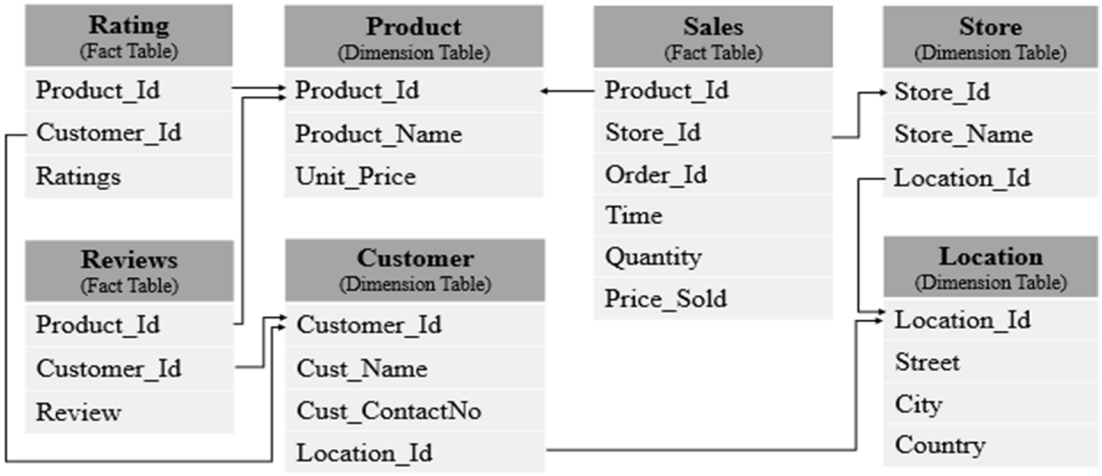


Fig. 1. General design schema of a data warehouse for Sales, Ratings & Reviews.

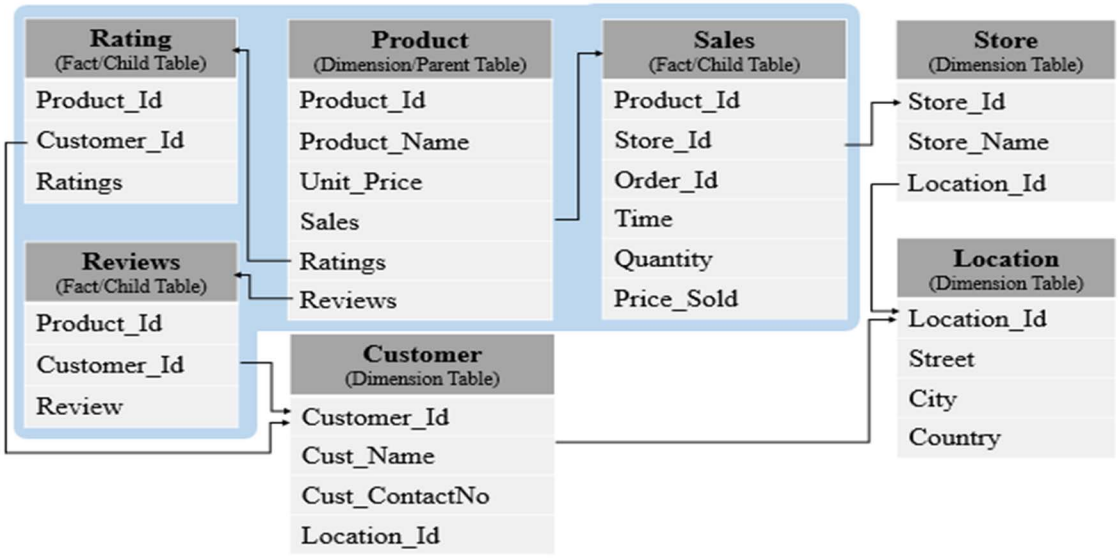


Fig. 2. Proposed schema of a data warehouse for Sales, Ratings & Reviews.

operational facts in a data warehouse and the attribute sets for these relations are-

Product: (Product_Id, Product_Name, Unit_Price)
Store: (Store_Id, Store_Name, Location_Id)
Location: (Location_Id, Street, City, Country)
Customer: (Customer_Id, Cust_Name, Cust_ContactNo, Location_Id)
Sales: (Product_Id, Store_Id, Order_Id, Time, Quantity, Price_Sold)
Ratings: (Product_Id, Customer_Id, Ratings)
Reviews: (Product_Id, Customer_Id, Reviews)

Fig. 1 describes the data warehouse for Sales, Ratings, and Reviews having the dimensions Product, Store, Location, and Customer to describe them using the existing data warehousing schemas. However, think differently, it would seem like these facts somehow belongs to a product. What it means is that sales, ratings or reviews should be part of a product, without a product there will be no sales, ratings, or reviews. Again, while retrieving the data record of a product, someone would like to get its sales, ratings or reviews as well. Because these facts are of a product rather than the product of these facts. Therefore, this method considers that the Product relation must have the sales, ratings & reviews record in its relational table and to do that, the Product relation must have additional attributes for these facts. Then the Product relation would be-

Product: (Product_Id, Product_Name, Unit_Price, Sales, Ratings, Reviews)

In this case, the Product relation does not describe these facts. These facts describe the corresponding attribute, with abstract data stored in their attribute field, of the Product table. Fig. 2 shows the data warehouse of these relations using the proposed schema. This design is the same as the general design schema except the relationship shown in the highlighted part of the figure. Here, the Product dimension table, the parent table in this structure, shares the facts to describe its Sales, Ratings, and Reviews attributes with appropriate data.

In this design, the parent table and the fact tables are not connected with the same attributes in both relations as they were in the previous design. Although they both have that common attribute, they are not bound to be connected. This method allows the parent table to contain the fact tables within its respected attribute field to describe the value stored in that field. So the facts will be partitioned into separated tables with the help of the ETL tools respect to the data stored in the parent table, e.g. in this design, the operational facts of sales, ratings or reviews will be partitioned into separate smaller in size fact tables as per the product that those facts describe. Therefore, the Sales, Ratings & Reviews attributes of the Product relation will have their own attribute set to describe their data recorded in the warehouse. To understand the relationship between a child table and its parent table, consider the two following relations-

Product: (Product_Id, Product_Name, Sales, Unit_Price)
Sales: (Order_Id, Quantity, Price_Sold)

Here, the Sales relation is dependent on the Product table as it describes the Sales attribute of the Product table. In the proposed schema, these two attribute sets are described by the following relation.

Product: (Product_Id, Product_Name, Sales: (Order_Id, Quantity, Price_Sold), Unit_Price))

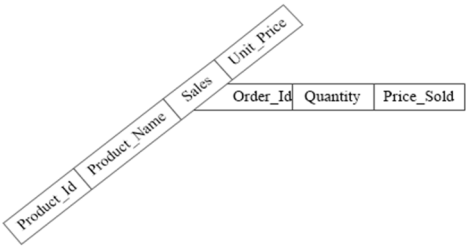


Fig. 3. Basic relationship of the proposed schema.

Fig. 3 shows the basic structure of the relationship used in the proposed schema but the inner attribute set does not actually belong to the respected attribute of the parent table. It actually belongs to the data field or record stored in that attribute, this is the main difference of this schema with the existing schemas. In existing schemas, all the relationship between different relational tables was record to record. This schema uses the relationship of record to table rather than record to record, a specific record of the parent table with the corresponding child table. The structure of this relationship is shown in Fig. 4.

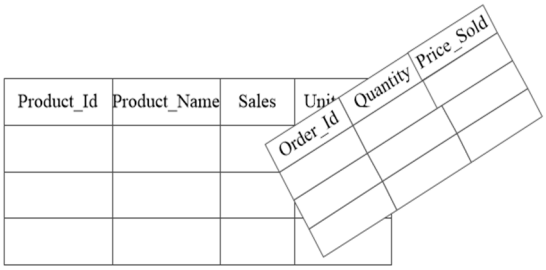


Fig. 4. Record to table relationship of the proposed schema

This schema designs the data warehouse a way that allows the parent table to store an abstract data to its attributes having an attribute set related to it. In this schema, the data stored in that attribute field is very important as this data is going to stand for its corresponding child table in data processing. So, the type of this data must be chosen carefully. To add additional attributes and selecting data for these attributes in the parent relation, this schema uses various data processing operations. Data integration and transformation operations allow to integrate different data sources and adding extra attributes to relations transforming data into appropriate form [11].

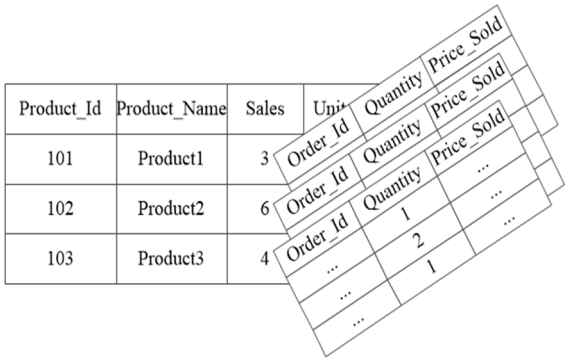


Fig. 5. Record to table relationships & retrieving the attribute value

Using the structure shown in Fig. 4, we are allowing the Sales attribute of the parent table to distribute the total sales to every product record and a partitioned Sales table in every record field of the Sales attribute. Fig. 5 shows that the sales attribute hold a relation table within its every record field and that record field itself stores a value. This value represents the whole table residing with it within that field. In this relationship, this abstract value is the total sales of that specific product. This structure ensures that if there is no record of a specific item in the parent table, there will be no facts stored in the warehouse corresponding to that item. Means no unnecessary facts to bother with.

B. Selecting the Parent (Dimension) Table

This schema partitions the facts according to the records of a specified dimension and the type of data stored in that dimension (parent table) have to be of great importance to the decision makers. Otherwise, the partitions will be of no use. So, the selection of the dimension table is very important in this schema. Author(s) [8] presents a greedy algorithm that select the best dimension table to partitions the fact table based on the queries asked most but, in this schema, the dimension could be selected directly by the choice of the decision makers too, as they might know what would be the best for them.

C. System Analysis

The Design of a data warehouse is an important part of warehousing technology. Success depends on that. But the most important thing that concerns most to the business leaders is the quality of the data and response time of the warehouse to different queries. To improve data quality and response time, a data warehouse needs to provide an easy and fast access to most required data and the data availability must be high but these both requirements depends on how the system is designed and how the data are interrelated. In general, a data warehouse stores data as fact and dimensions and relate them together with different relational key, with a similar type of attributes in both relations, mostly to reduce redundancies.

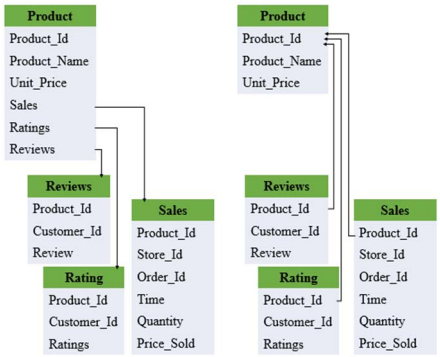


Fig. 6. Relation tree hierarchy of the proposed schema (left) & general relationships between relational tables (right).

Fig. 6 shows a general comparison of the relationships among different relational tables and how the proposed schema is different from the existing schemas. The relation tree hierarchy of the proposed schema uses the record to table relationships whereas the existing schemas use the relationship of record to record. The proposed schema partitions the facts into several child tables as per several records of the parent table and connects these child tables with their corresponding attribute data. So every attribute of the parent table containing an abstract data is related to a specific child table that the abstract data comes from.

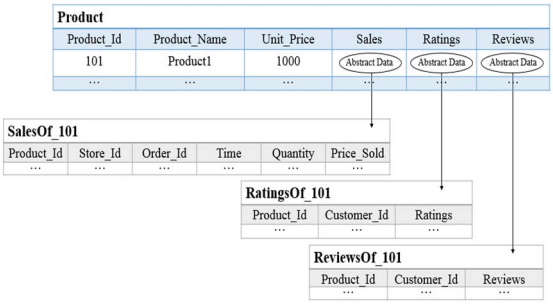


Fig. 7. Relation tree hierarchy of the Product relation.

Fig. 7 describes the relation tree hierarchy of the Product table having sales, ratings, and reviews as child tables in the hierarchy. As different child tables are related to different product record, they are distinct in a specific way, here every child table bear the 'Product_Id number' (could be anything that identifies distinct product) of the product they belongs to, and that allows the warehouse an easy and fast way to find the data set needed for specific query response. This figure describes the relationships of just one product with its child. The Product table will have the same relationship for every product of it.

Fig. 8 describes the general data distribution and the relationships between the fact tables and the dimension table. In this schema, every record of the fact tables is individually related to a corresponding product record of the dimension table using a similar attribute in both tables. In the proposed schema, operational facts (e.g. Sales, Ratings and Reviews) are partitioned as per every product in the dimension table and related to their corresponding product through a specific attribute. So, finding the data of a specific product would require less time as it would not need to bother with unnecessary data.

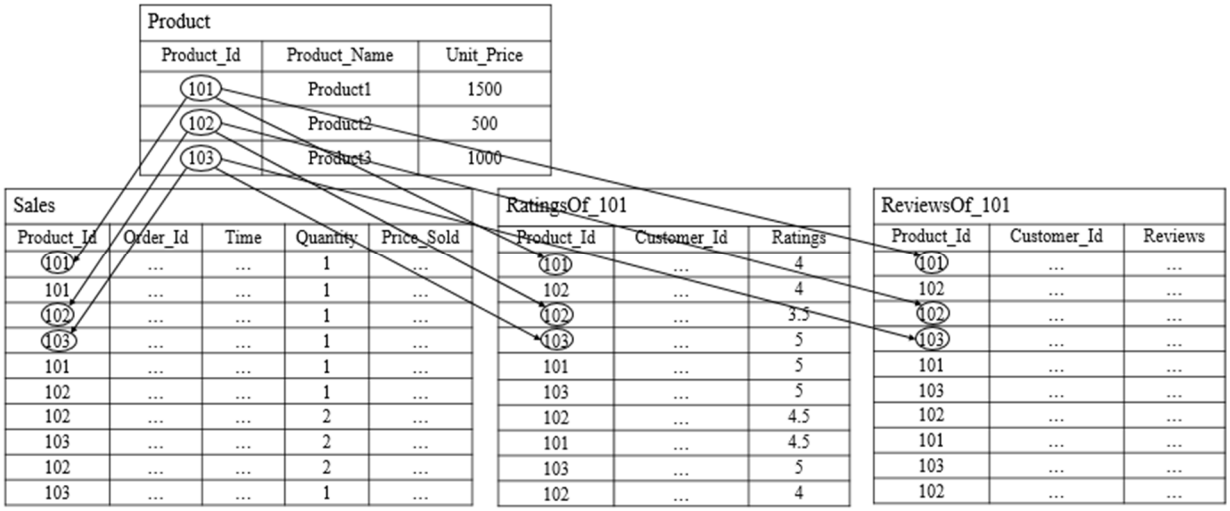


Fig. 8. General data distribution in a data warehouse.

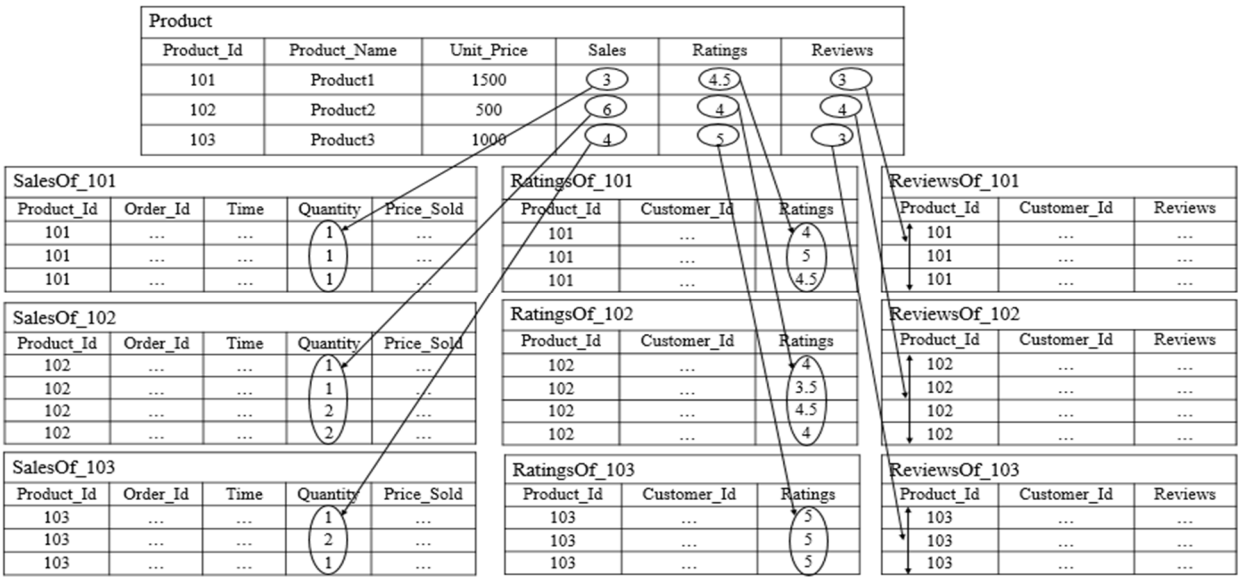


Fig. 9. Data distribution in a data warehouse according to the proposed schema.

Fig. 9 shows the distribution of operational facts among different child tables as per the records of the parent table. This figure shows the data distribution according to a major subject or entity. The Product table stores an abstract data to its specific attributes collected from its child tables. In this design, the Sales attribute gets the total sales quantity, average rating for Ratings and total review count for Reviews from their respected child tables.

This schema has additional attributes to a specific dimension table for every type of facts related to that dimension and these attributes record an abstract data that represents the child fact table related to it. So more data are available in the dimension table and they are easy to process to response certain queries. On the other hand, the child tables are related to different items of the parent table and are uniquely distinguishable, so these child tables are easy to access. Above all, as the facts are partitioned into small parts, they take less time to be processed.

In this schema, the additional attribute of the dimension table increases data availability and gives a complete view of data in the dimension table. As they store summarized data, they are easy to process as well. At the same time, by allowing the data warehouse to store fact tables within the dimension table, this schema increases data accessibility. As the facts are partitioned into small parts, the data warehouse is required to process less data to get the most required information for certain queries with minimum response time.

On the other hand, to design the relation tree hierarchy and the form of the abstract data, this schema is dependent on business intelligence along with other requirements of the decision makers. So, the data warehouse is more connected to decision makers than ever. And as this schema connects the warehouse and the decision makers more closely, the data warehouse is more able to meet the requirements and demands of the decision makers and helps better in strategic decision making with high quality data and improved response time.

V. Conclusion & Future Work

This paper proposed the foundation of a new data warehouse modeling schema. This schema overcomes the general thinking of designing a data warehouse where dimension describes facts. This schema uses the operational facts to describe certain attributes of a dimension and distributes those facts among the records of that dimension. This schema partitions the facts horizontally and results in several small fact tables. Those small fact tables are related to attributes, additional attributes for facts summarization, of a dimension table that increases the data availability in the dimension table. Again, the data warehouse requires less time in response to certain queries as it requires to process a small set of data. This schema also allows easy and fast access to data as they are uniquely distinguishable. Therefore, comes up with better performance in providing high data quality with faster response in various query processing.

However, we are working on some challenges we have come up with like the connection or relation between a child table and the specific record of its corresponding parent table, which is the main structure of this schema, and the common view of all the child tables of the same fact in the relation tree hierarchy. Keeping all these requirements and challenges in mind, we will carry on further work with this schema.

References

[1] Mohammad Rifaie, Keivan Kianmehr, Reda Alhajj and Mick J. Ridley, "Data warehouse architecture and design" in IEEE International Conference on Information Reuse and Integration, Las Vegas, NV, USA, pp. 58-63, August 2008.

[2] G. Blažić, P. Poščić and D. Jakšić, "Data warehouse architecture classification" in 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, pp. 1491-1495, May 2017.

[3] Vandana Bhagat and Dr. Arpita Gopal, "Comparative Study of Row and Column Oriented Database", in Fifth International Conference on Emerging Trends in Engineering and Technology, Himeji, Japan, pp. 196-201, November 2012.

[4] Ricardo Jorge Santos, Jorge Bernardino and Marco Vieira, "A survey on data security in data warehousing: Issues, challenges and opportunities", in IEEE EUROCON - International Conference on Computer as a Tool, Lisbon, Portugal, pp. , April 2011.

[5] Prayag Tiwari, Sachin Kumar, Avinash Chandra Mishra, Vivek Kumar and Bodena Terfa, "Improved Performance of Data Warehouse", in International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India, pp. 94-104, March 2017.

[6] Fatimah Sidi, Abdullah Ramli, Marzanah A. Jabar, Lilly Suriani Affendey, Aida Mustapha and Hamidah Ibrahim, "Data Quality Comparative Model for Data Warehouse", in International Conference on Information Retrieval & Knowledge Management, Kuala Lumpur, Malaysia, pp. 268-272, March 2012.

[7] Bing Chen, Xuchu Weng, Beizhan Wang and Xueqin Hu, "Analysis and solution of data quality in data warehouse of Chinese materia medica", in 4th International Conference on Computer Science & Education, Nanning, China, pp. 823-827, July 2009.

[8] L. Bellatreche, K. Karlapalem, M. Mohania and M. Schneider, "What can partitioning do for your data warehouses and data marts?", in Proceedings 2000 International Database Engineering and Applications Symposium, Yokohama, Japan, pp. 437-445, September 2000.

[9] Mandeep Kaur Sandhu, Amanjot Kaur and Ramandeep Kaur, "Data Warehouse Schemas", (IJIRAE) International Journal of Innovative Research in Advanced Engineering, Vol. 2, pp. 47-51, April 2015.

[10] Jiawei Han and Micheline Kamber, "Data Warehouse and OLAP Technology: An Overview" in Data Mining: Concepts and Techniques, 2nd Ed. San Francisco, USA: Diane Cerra, 2006, pp. 105-117.

[11] Jiawei Han and Micheline Kamber, "Data Preprocessing" in Data Mining: Concepts and Techniques, 2nd Ed. San Francisco, USA: Diane Cerra, 2006, pp. 61-86

Feature Selection and Classification of Spinal Abnormalities to Detect Low Back Pain Disorder using Machine Learning Approaches

Md. Shariful Islam, Md. Asaduzzaman, Mohammad Masudur Rahman
Department of Computer Sciencce and Engineering
Daffodil International University
Dhaka, Bangladesh
{shariful4312, asaduzzaman4285, masudur.cse}@diu.edu.bd

***Abstract*—With the advent of computational intelligence, the analysis of medical data using machine learning techniques benefits accurate classification of different health diseases and disorders. However, disorganization and variability of data make the job difficult. This paper has streamlined an ensemble learning approach for classification of low back pain disorder based on spinal abnormality data of 310 patients with 12 features. To overwhelm the misleading effect of inappropriate attributes, most influential features are identified using evolutionary feature elimination method. Experiments are performed in both way– with or without feature filtering. The basic machine learning algorithms used in the work: Logistic regression, Decision Tree, Naive Bayes, and in addition to the Random Forest ensemble learning method. Random Forest classifier, as expected, is recorded to exhibit the best accuracy of 94% over other classifiers.**

Keywords— Machine Learning, Feature Selection Random Forest, Low Back Pain, Lumber spine Disorder.

I. INTRODUCTION

Low Back Pain, abbreviated as LBP, is a very frequent health issue that affects people from all ages, younger to the elderly [1]. Almost 80% of adults population suffer from low back pain problem at some points in their lifespans [2]. It is reported as the single leading cause of job-related disability and a major contributor of preventing people from attending workplace as well as their daily activities. According to the Global Burden of Disease (GBD) Study report, low back pain is one of the top ten high burden diseases, which leads over HIV (AIDS), lung cancer, tuberculosis, chronic pulmonary disease and preterm birth complications [3,4]. The sufferings and magnitude of LBP disorder dispense a catastrophic effect and become a worldwide problem. In 1990, a study places low back pain in sixth position of burdensome maladies in the United States whilst in 2010, it ascended to third place, with heart disease and chronic pulmonary disease on the top. Meanwhile, low back pain prevalence in developed countries is reported to be up to 85% [2,4].

With the fast development of computational intelligence, researchers are paying attentions on prediction of different diseases from medical data. As low back pain, becoming global fatal problem, got concern of machine learning experts to identify abnormal patients analyzing theirs spinal region data. Physical test or examination and observation in historical patterns of signs and symptoms in a diagnostic view may assist the medical professionals in determining the abnormal patients [5]. In the work, we mainly focus on different angles and orientations of components of lumber spine. The novelty of this work investigates the following challenges: Do all the recorded features, that doctors suspects, are influential? How will the spinal abnormality in a certain region be responsible for low back pain? Which computerized learning techniques would be good for classification of low back pain? To resolve these queries, we first, build singleton machine learning models, then, build ensemble the models in the quest of higher accuracy. It is noted that ensemble models naturally perform better in comparison to single classifiers, but it is not true in all cases [17].

The rest of this paper is succinctly structured into various sections as follows. Section 2 probed some clinical and statistical studies regarding low back pain. The theory and mathematical computations of some essential machine learning approaches is illustrated concisely in section 3. Section 4 represents the spinal data set details and experiments followed by the performance analysis is provided in section 5. Finally, section 6 concludes the study with an interpretation of results and also pledges some recommendations for future research.

II. RELATED WORKS

A good number of studies have been conducted related to low back pain symptoms. Some of the studies find, very likely, to analyze medical aspects and global consequences of low back pain disorder. A few of them work are conducted on artificial intelligence traits.

Lin et al. proposed an advanced clinical decision support system for LBP diagnosis and gain overall performance of 71% in 2006 [6]. Alomari et al., in 2009, introduced a probabilistic classifier on the detection of abnormal intervertebral disc using three feature including appearance, location and context from clinical lower spine MRI [7]. In 2009, Neto and Barreto introduced a comparative analysis of singleton approaches (SVM, MLP, GRNN) and their combination in a ensemble on the detection of vertebral column for LBP [8].

In 2010, a study by Sari et al. proposed two systems for the estimation of LBP level those accept skin resistance and visual analog scale (VAS) values as input. The results from artificial neural network (ANN) and Adaptive neuro-fuzzy Inference (ANFIS) claim good performance in real application [9]. Gosh et al. experimented two approaches in 2011 in this area of detection LBP. The first one is the follower of the experiment of Alomari [7] where Gosh built a heterogeneous classifier as their novelty [10]. In the later experiment, Gosh combines composite features for automatic diagnosis of intervertebral disc herniation of lamber MR

Images. In 2012, Koh et al. designed a two level classification on the same purpose using a CAD framework [11].

A novel research by Mandal adduced a new hybrid machine learning ensembles consisting of Bayesian networks with Tabu search and Haar wavelet and then compare the method with existing models [12]. Later, in 2017, d'Hollosy evaluated three tree-based machine learning methods: decision tree, random tree, and boosted tree on a Clinical Decision Support System (CDSS) [13]. Abhijit et al. in 2017, applied basic machine learning methods- k-nearest neighbor, support vector machine, decision trees etc. They observed that decision tree algorithm yields an accuracy of 90% than other methods [14]. Later on., Kulkarni suggested a deep learning based techniques of classification for LBP disorder in 2017 [15].

III. Experimental Setup

In this section we will reveal the working flow of experiments. The workflow of the proposed methodology is shown in Fig. 1. The first step of the work, problem statement is defined and data collection and pre-processing are done, followed by the step of removing outliers from the data to improve the quality of data. Then the feature selection method is provided to find the importance of the features and select the significant factors. After that, the individual machine learning algorithms are employed along with Random Forest method.

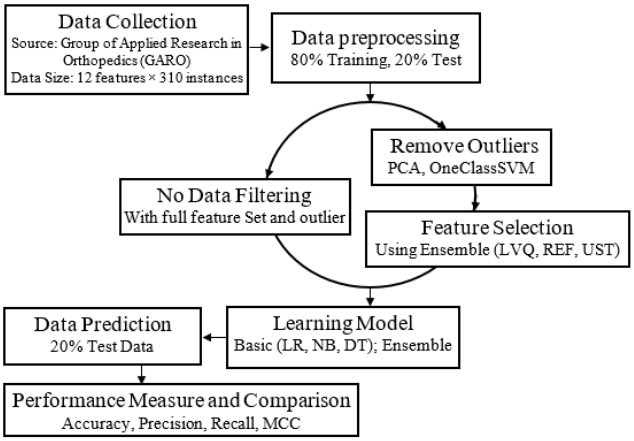


Fig. 1. System workflow of the proposed methodology.

A. Data Set and Data Pre-processing

The data set which is treated in this research is collected from the Group of Applied Research in Orthopaedics of the Centre Médico-Chirurgical de Réadaptation (CMCR) des Massues in Lyon, France during a residence period in spinal surgery. It contains the observations of radiogram of the spine of patients. A total of 310 spinal instances with 12 features are recorded. Each patient is represented with attributes derived from the shape and orientation of the pelvis and lumbar spine. The observations have been target labeled with binary class label either normal or abnormal. The number data items with normal class label is 100 whilst with abnormal class label it is 210. Thus the problem becomes an imbalanced problem. The data set does not contains any missing value. Table I enlists the full set of attributes.

TABLE I. Attributes

Attribute Name	*Type*	*Attribute Name*	*Type*
Pelvic incidence	Numeric	Pelvic slope	Numeric
Pelvic tilt	Numeric	Direct tilt	Numeric
Lumbar lordosis angle	Numeric	Thoracic slope	Numeric
Sacral slope	Numeric	Cervical tilt	Numeric
Pelvic radius	Numeric	Sacrum angle	Numeric
Degree spondylolisthesis	Numeric	Scoliosis slope	Numeric
Class Attribute/Label	Object		

B. Outliers Elimination

An outlier is an observation or data which has far distance from other observations. It deviates serious disturbances on building learning models. An outlier can caused by experimental error, variability in measurements. Whatever is happens, to yield good accuracy, we need to remove the outliers from the data set. There are several approaches to eliminate outliers such as extreme value analysis: z-score, interquantile, DBSCAN, regression model: PCA etc. To determine whether our data set contains outliers or not, we apply PCA in the data set. The output is plotted in the Fig. 2.

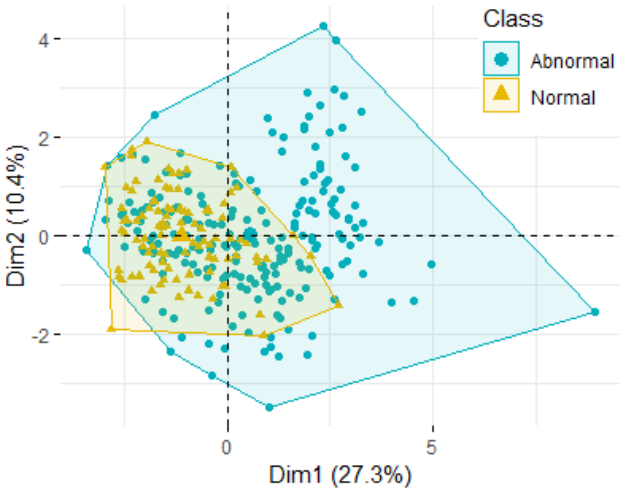


Fig. 2. Outlier projection using PCA.

For univariate outlier elimination, extreme value analysis methods work efficiently. As we are dealing with multivariate data, we apply support vector machine algorithm to detect outliers in the data. *OneClassSVM* technique allows us to identify isolated data points in the feature space considering with/without data distribution and expected percentage of outliers using the params:

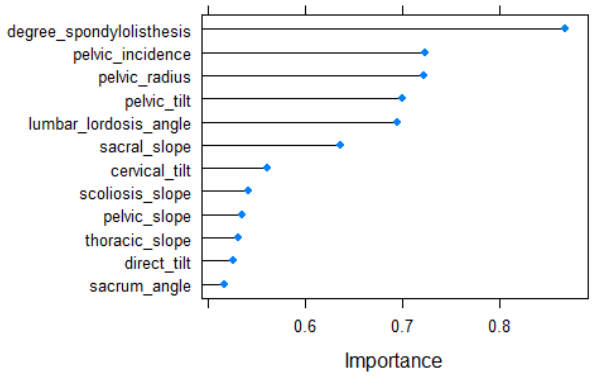
$$nu\ estimate = 0{:}95 * outliers\ percentage[0; 1] + 0{:}05$$

After the execution of *OneClassSVM*, the data set is clean and outlier free at our confidence level.

C. Feature Selection

Feature selection is an important ML task as excessive, irrelevant and redundant features may mislead the learning model. But feature selection procedures become crucial with large number of features. Moreover, feature selection is a combinatorial problem. There are two types of feature selection approaches: Filtering and Wrapper. Filtering processes find the importance of the features and rank them according to a weight. Meanwhile, wrapper methods search for subsets of features to maximize the computational time.

Wrapper methods start occasionally with empty or complete set of features while filter methods always initiates with full set features. We approach a setwise evolutionary based wrapper paradigm for feature selection. It combines multiple features selection algorithm such as- Learning Vector Quantization (LVQ), Recursive Feature Elimination (RFE), and Chi-Squared Univariate Statistical Test (UST), but these



algorithms worked in a incremental way by adding and removing features. While LVQ searched the best features, RFE worked on the LVQ selected features. UST worked on individual feature to find the best performer. Fig. 3 shows the best features selected that obtained from these algorithms. In the table figures indicates the individual contribution of the features.

Fig. 3. Feature Importance using REF.

Using the concept of voting Ensemble method, finally we have a set of features with majority votes. This is one of the novelties of our work of using ensemble in feature selection step.

D. Classification

We apply basic machine learning algorithm to build classifiers. We built five classifiers including Logistic Regression (LR), Naive Bayse (NB), Decision Tree (DT) and Random Forest (RF) algorithms incorporating feature reduction as our data may contain irrelevant attributes. These algorithms run individually to observe which performs best.

On the purpose of building classification model, first, we employ singleton machine learning algorithms: Logistic Regression, Decision Tree, Naive Bayes. We apply 10-fold cross validation on training data. Following that, we deploy Random Forest ensemble learning method consisting of best performing classifiers in the previous step of the experiment.

After modeling the system we test the performance. We evaluate our model in terms of accuracy, precision, and recall. As our work is supervised, the data set is divided into two set by splitting:

1. TRAIN Set – 248
2. TEST Set – 62

The models are trained using training set data and tested on test set after building the model. We acquired the output in terms of confusion matrix. Then, accuracy, precision, and recall are calculated with respect to the class predictions.

IV. Performance Analysis

In this section we will discuss about the results of the guided experiment. The results and performance have been evaluated using different measures such as accuracy, sensitivity, specificity, and precision [16]. These measures involves the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) (defined as in [16]) observations and predictions like:

	The number of
TP	correctly classified positive samples (Class=Normal)
TN	correctly classified negative observations (Class=Abnormal)
FP	incorrectly classified positive instances (Class=Normal)
FN	incorrectly classified negative attributes (Class=Abnormal)

Accuracy is a measure of overall *effectiveness* of the classification model. It indicates the ratio of correct classifications over total observations. It can be calculated as stated Eq. (1)

$$Accuracy = \frac{TP+FN}{TP+TN+FP+FN} \quad (1)$$

Precision is a measure of *exactness* of the classifier and it states the proportion of correct positive classifications over the positive predictions. It can be calculated as Eq. (2)

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Recall, also called Sensitivity or True Positive Rate or *completeness*, computes how well the model in identifying the positive instances (Class=Normal). It is calculated as Eq. (3)

$$Recall = \frac{TP}{TP+TN} \quad (3)$$

Table II and III depict the summary of results of the Random Forest classification method without and with Feature Selection, respectively. The confusion matrix states TP, TN, FP, and FN values for actual and predicted classes. In comparison, the classification rate in Table III is higher than Table II. This proves better performance of Random Forest classifier with selected features.

TABLE II. Confusion Matrix of Random Forest (Without Feature Selection)

	Predicted Normal	*Predicted Abnormal*
Actual Normal	17	7
Actual Abnormal	3	35

TABLE III. Confusion Matrix of Random Forest (With Feature Selection)

	Predicted Normal	*Predicted Abnormal*
Actual Normal	20	2
Actual Abnormal	2	38

TABLE IV. Comparison of Classification Methods (Without Feature Selection) (in %)

Measurements	***Logistic Regression***	***Naive Bayes***	***Decision Tree***	***Random Forest***
Accuracy	82	79	81	84
Precision	68	62	68	71
Recall	85	90	75	86

TABLE V. Comparison of Classification Methods (Without Feature Selection) (in %)

Measurements	***Logistic Regression***	***Naive Bayes***	***Decision Tree***	***Random Forest***
Accuracy	90	84	92	94
Precision	82	75	86	90
Recall	90	82	90	90

The similar patterns can be seen for Logistic Regression, Naive Bayes, Decision Tree as well. We can observe that from Table IV and V. The tables stated precision, recall, and F1 Score along with accuracy. We observe the highest accuracy of 94% gained by Random Forest ensemble.

It can be observed that the models perform better with feature selection. Feature Selection step of our experiment discards less important features and it may very useful for medical data processing with computational intelligence where huge bunch of data and features are recorded. This experiment answers our research questions as not all the features are useful to consider. However, a patient with low back pain may suffering from other lumber spine issues such as damage of spine due to age, radiculopathy etc. These factors may consider as less responsible for specific deceases or disorders or damages.

In the same experiments we focus on a particular research question which finds the supporting and contradicting trends of the attributes towards the abnormality of lumber spine. disorder. To answer the question, we use the contribution score of the attributes from feature selection step. Table VI summarizes the results. We projected that this information can help to more accurate diagnosis and prescribing the therapy for the patients.

TABLE VI. SUPPORTING AND CONTRADICTING ATTRIBUTES TOWARDS ABNORMALITY

Supporting Attributes	*Contradicting Attributes*
Degree spondylolisthesis	Scoliosis slope
Pelvic incidence	Pelvic radius
Sacrum angle	Sacral slope
Lumbar lordosis angle	Direct tilt

V. CONCLUSIONS AND FUTURE DIRECTIONS

The study shows that an automated system can be implemented in the diagnosis of low back pain. We conclude answering our research questions with experiments. More training data and more efficient algorithm this can be a real life implementation for clinical diagnosis in Bangladesh. Bangladesh medical system can accumulate more clinical data in an organized and structured way where each patient's medical data can be secured which can be later be used in many different studies for other diseases and also incorporate with the existing system for automated classification and detection of those diseases. The research intends an artificial intelligent, in the other, machine learning approach to identify responsible attributes of causing low back pain from the spine. Although the origin of low back pain remains unknown in most of the cases, some frequent occurring symptoms are identified to be useful in this study. This intuits medical professionals to early diagnosis of the low back pain disorder.

REFERENCES

[1] Devon I. Rubin, Epidemiology and Risk Factors for Spine Pain, Neurologic Clinics, Vol. 25, No. 2, 2007, Pages 353-371.

[2] Low Back Pain Fact Sheet, National Institute of Neurological Disorders and Stroke, USA. Available: https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/FactSheets/Low-Back-Pain-Fact-Sheet. Last accessed: 1 October, 2018.

[3] Hoy D, March L, Brooks P, et al. The global burden of low back pain 2010 study Annals of the Rheumatic Diseases. 2014;73:968-974.

[4] Janet K. Freburger et. el., The Rising Prevalence of Chronic Low Back Pain, US National Library of Medicine, National Institutes of Health, Arch Intern Med. 2009 Feb; 169(3): 251258.

[5] Tom Petersen, Mark Laslett and Carsten Juhl, Clinical classification in low back pain: best-evidence diagnostic rules based on systematic reviews. BMC Musculoskeletal Disorders; 2017. 18:188.

[6] Lin Lin, Paul Jen-Hwa Hu, Olivia R. Liu Sheng, A decision support system for lower back pain diagnosis: Uncertainty management and clinical evaluations, Elsevier Decision Support Systems, 42 (2006).

[7] Raja S. Alomari, Jason J. Corso, Vipin Chaudhary, Gurmeet Dhillon, Computeraided diagnosis of lumbar disc pathology from clinical lower spine MRI, Intl. J. of Computer Assisted Radiology and Surgery (2010) 5: pp.287-293.

[8] A.R. Da Rocha Neto, G.D.A. Barreto, On the application of ensembles of classifiers to the diagnosis of pathologies of the vertebral column: a comparative analysis, IEEE Latin Am. Trans. 7 (4) (2009) 487-496.

[9] Sari, Murat, et el., Prediction of Low Back Pain with Two Expert Systems, Journal of Medical System (2012) 36: pp. 1523-1527.

[10] S. Ghosh, R. S. Alomari, V. Chaudhary and G. Dhillon, "Computer-aided diagnosis for lumbar mri using heterogeneous classifiers," IEEE Intl. Symp. on Biomedical Imaging, Chicago, IL, 2011, pp. 1179-1182.

[11] S. Ghosh, R. S. Alomari, V. Chaudhary and G. Dhillon, "Composite features for automatic diagnosis of intervertebral disc herniation from lumbar MRI," Annual Intl. Conf. of the IEEE Engineering in Medicine and Biology Society, Boston, 2011, pp. 5068-5071.

[12] Indrajit Mandal, Developing new machine learning ensembles for quality spine diagnosis, Knowledge-Based Systems, Volume 73, 2015, pp. 298-310.

[13] W. O. Nijeweme-dHollosy, L. van Velsen, M. Poel, C. G.M. Groothui, Remko Soer, Hermie Hermens, Evaluation of three machine learning models for self-referral decision support on low back pain in primary care, Intl. J. of Medical Informatics, Vol. 110, 2018, pp. 31-41.

[14] Abhijitsingh Putu Gaonkar, Ritwik Kulkarni,Ronnie D. Caytiles and N.Ch.S.N Iyengar, "Classification of Lower Back Pain Disorder Using Multiple Machine Learning Techniques and Identifying Degree of Importance of Each Parameter," International Journal of Advanced Science and Technology, Vol.105 (2017), pp.11-24.

[15] K Ritwik Kulkarni, Abhijitsingh Gaonkar, V Vijayarajan and K Manikandan, Analysis of lower back pain disorder using deep learning, IOP Conf. Series: Materials Science and Engineering 263 (2017).

[16] Farid, Dewan Md, Mohammad Masudur Rahman, and M. A. Al-Mamuny. "Efficient and scalable multi-class classification using nave Bayes tree." Intl. Conf. on Informatics, Electronics & Vision (ICIEV). IEEE, 2014.

[17] Zhang,G., Patuwo, B.E., and Hu, M.Y., (1998). Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, 14 (1998), pp. 35–62.

An Efficient Ensemble Method for Cancer Detection

Md. Sirajum Munir Prince*, Asib Hasan†, Faisal Muhammad Shah‡
Department of Computer Science and Engineering, Ahsanullah University of Science and Technology
Dhaka, Bangladesh
Email: *sirajummunirprince@gmail.com, †asibhasan.cse@gmail.com, ‡faisal505@hotmail.com

***Abstract*—Very few diseases are as deadly as Breast cancer and Lung cancer that causes a high percentage of deaths every year. Data mining approach works quite effectively in diagnosing cancer. This has been possible mainly because of the potential of classification systems that has been improved in a great arrangement in helping medical experts. In this research work, data mining has been applied for the diagnosis of cancer. For the purpose of that, an ensemble method has been used to extract the most important features in detecting cancer and classify the pattern of cancer according to the most influential features. For feature extraction the advantages of Principal Component Analysis (PCA), Pearson Correlation Coefficient (PCC) and Chi Square (Chi2) have been considered using set union operation. By the help of the hybrid model of feature extraction, and classification, Naive Bayes, K Nearest Neighbors, Decision Tree (J48, ID3) classifiers were used to classify the model. Results from individual classifiers have been voted by majority voting to obtain the best possible outcome. k- fold cross validation (10 and 5 folds) and train (80% and 70%) test (20% and 30%) split have been applied to examine the performance of the model(AEEMCD).**

***Index Terms*—Data Mining, Cancer Detection, Feature Selection, Set Union Operation, Classification, Ensemble, Majority Voting, PCA, PCC, Chi2, Naive Bayes, KNN, Decision Tree.**

I. Introduction

Cancer has been one of the most lethal diseases in the recent years. Among women in this world the most common cancer is breast cancer. Men are highly affected by lung cancer. Even though the incidence of breast cancer in developing countries is lower than in Western countries, it is rising rapidly in the particular region. Significant rates of death due to cancer could be reduced in the long term if the cancer is detected in early stages. For best prognosis, identifying cancer cells in early stages is the most critical point. Early stage cancer diagnostic and high sensitive method is a must. Classification and Data Mining methods are effective ways to categorize data especially in the medical field. Computing, and machine learning tools can significantly be used to help solve health care problems by assisting physicians in diagnosing, and predicting diseases in early stage by developing expert systems. Computer Aided Diagnosis (CAD), and medical experts systems, and tools have become one of the foremost research areas in the field of medical diagnosis [1]. In this research work, a few of classifiers are used to detect cancer on their samplings. The precision of these classifiers depends both on their algorithms and how they are trained. In sequence, train them properly using a dataset, the most applicable features of the feature vector have to be taken. In this way, we can increase the efficiency of the classifiers as well as reduce the learning runtime. For this logic, we work to find out the important features of our training datasets and then we use set theorem method on that features to find out the most common features of our datasets.

There are 10 features in the Wisconsin Breast Cancer (Original) Datasets, 32 features in the Wisconsin Breast Cancer (Diagnostic) Datasets and 57 features in Wisconsin Lung Cancer Dataset. All the experiments were performed on WEKA 3.6 [2]. This research work is assembled in these ways: Section II covers the related works on these classifiers and datasets. Section III describes proposed model and followed by the result analysis, and conclusion and future works in section IV, and V.

II. Literature Review

There exist innumerable papers and books demonstrating the convenience of the feature selection process [3], [4].

On the Wisconsin Breast Cancer (original) datasets, the performance of four algorithms of machine learning: Naive Bayes (NB), Decision Tree (C4.5), k-Nearest Neighbors (k-NN) and Support Vector Machine (SVM) where SVM gives the highest accuracy (97.13%) was proposed [5]. All experiments were executed in WEKA data mining tool. The potency of the classifier was evaluated in terms of Kappa statistics (KS), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Relative Absolute Error (RAE) and Root Relative Squared Error (RRSE). KS, MAE and RMSE were represented in numeric values. RAE and RRSE were represented in percentage.

A method to extract classification rules from trained Naive Bayes and discussion of its application to breast cancer diagnosis is stated in [6]. The pre-processing of data set that can raise the accuracy of the neural Nave Bayes and the correctness of the rules as some rules may be extracted from human experience, and may be inaccurate were also explained. Laplacian Correction was used when their method got zero probability values. The experiment was carried out in MATLAB (matrix laboratory). The Nave Bayes classification algorithm showed the success rate as around 85-95 percent and the error rate was around 10-15 percent.

A bio inspired model named BATELM was suggested in [7] which combined Bat algorithm (BAT) and Extreme Learning Machines (ELM) on Wisconsin Breast Cancer Prognostic (WBCP) dataset which is a numeric dataset. The concept of BAT and ELM which had many advantages when constrasted to the existing algorithms of their genre had inspired them

to build a model that could predict the medical data with high accuracy and minimal error. The objective of the concept was to predict whether the breast cancer is recurrent or non-recurrent. The accuracy was found to be 95.7% with an RMSE of 0.32. Another objective was to predict the time of recurrence, the result was found to be 93.75% accurate with an RMSE of 0.30.

The collation of Support Vector Machine (SVM), Random Forrest (RF), and Bayesian Network (BN) with and without using k-fold validation techniques was shown in [8]. WEKA machine learning tool was used. Results showed that Bayesian Network (BN) was the best in terms of performance, recall and precision. On the other hand, Random Forest (RF) technique had the optimum ROC performance when compared to the two other techniques. This implied that RF had a higher chance of discriminating between malignant and benign cases. Results had showed that SVMs have the highest performance in terms of accuracy, specificity and precision. However, RFs had the highest probability in correctly classifying tumor.

The accuracy was 97.90% in WDBC when using feature selection technique (Pearson Correlation Coefficient) and Naive Bayes, Logistic Regression, Decision Tree classifier [9]. In this 70% of overall data was used training and the remainder was used for testing the accuracy of classification technique. The simulation result owe support to Logistic Regression with the minimum time taken to build the model is 0.14 seconds.

According to [10], deleterious lumps in an early stage may raise the number of cancer healing patients. The specialists had confirmed that prior detection of cancer can be a great attainment for medical science. In this model the false positive rates were very little. There were some conflicts about the accuracy of the screening.

In this research work, an ensemble method is proposed that relies on voting mechanism for result rather than depending on single classifier. Multiple feature selection techniques find the repetition of features from the individual ones to select final feature sets.

III. Methodology

The main purpose of this study is to design an ensemble method that can detect cancer efficiently. In the process, cancer dataset is given as input. Feature selection techniques are used to select the most influential features with the help of set union operation. Sequentially, classification algorithms are the evaluation process for the precise dataset. It is to be noted that both cross validation and train-test split are chosen for the purpose of comparing the outcomes. In the end, the results are voted using majority voting to achieve the final outcome. The flow chart of the whole model is depicted in figure 1.

For the greater purpose of this research the whole work shown in the flow chart of figure 1 is divided into three phases. The phases are denoted as phase 1, 2, and 3.

A. Phase 1

In this phase the most important features of the three cancer datasets are selected. All the features that are selected using PCA, PCC, and Chi2, are minimized after performing set theorem operation to obtain the most common and influential ones.

Feature selection is primarily an essential process in data mining. Feature selection is a process which explains attribute selection, variable selection and variable subset selection. It is the procedure of choosing a subset of relevant features (variables, predictors) for use in model assembly. A feature selection algorithm defines the combination of a technique of search in reason of proposing new feature subsets. It includes an evaluation measure that make a score of the different feature subsets [11]. The final feature sets are gained using set theorem. To be precise, union operation is performed to get the common features from the feature sets that are found from individual feature selection techniques. After performing this operation, only the common features from the cancer datasets are gained, and eliminated feature sets are considered to be not so influential as they are not common in all the feature sets after performing individual feature selection methods. A brief idea of the feature selection techniques that have been used in this research work, is given is this section.

1) Principle Components Analysis(PCA): For the purpose of discovering the decoration of data that is of high dimension, PCA is a very handy method. [12]. It basically highlights the similarities and the dissimilarities among the data. Larger number of variables are shrink to a smaller number, which have all the necessary information required. This process turns the larger number of interconnected variables into a few variables that are not interconnected. All of the variables are bound to the variability in the data for every remaining variable. It is to be noted that this process only performs for square matrices.

$$\omega_1 = argmax\sum_{i=1}^{n}(t_1)^2 = argmax\sum_{i=1}^{n}(X_1.\omega)^2 \qquad (1)$$

2) Pearson Correlation Coefficient(PCC): Pearson Correlation Coefficient also known as PCC is a feature selection procedure which is used to measure the strength of a linear connection between two variables. The range of the returned value by the formula is -1 to 1, where 1 mentions a strong positive connection and -1 mentions a strong negative connection [13]. The returned value of zero mentions, there is no relationship at all.

$$R = \frac{n\sum xy - \sum x \sum y}{\sqrt{(n\sum x^2) - (\sum x)^2 (n\sum y^2) - (\sum y)^2}} \qquad (2)$$

3) Chi Squared(Chi2): The Chi-Square test of independence is a statistical test to determine if there is a significant relationship between two categorical variables [14]. The Chi-Square test is used to find which variables have an association with the Survived variable. The Chi-Square statistic will also test whether there is a significant difference in the observed vs the expected frequencies of both variables.

$$X^2 = \frac{N(AN - MP)^2}{PM(N-P)(N-M)} \qquad (3)$$

To get the final feature set, we have used three kinds of feature selection methods and have done set union operation on the basis of intersecting the selected features. We have minimized the size of the feature set and it has guided to a simple and faster cancer detection classification model. Let's assume we get feature set F_1, F_2, and F_3 from PCA, PCC, and Chi2 feature selection approach. Then the primary feature sets that have come from extracting the feature between F_1, F_2 and F_3 and then combined to evaluate the feature set to get the most common feature set F. This process is as follows:

$$F = (F_1 \cap F_2) \cup (F_2 \cap F_3) \cup (F_3 \cap F_1) \tag{4}$$

The idea of doing union operation on the results of three kinds of feature selection methods is easy and it also avoids complexity.

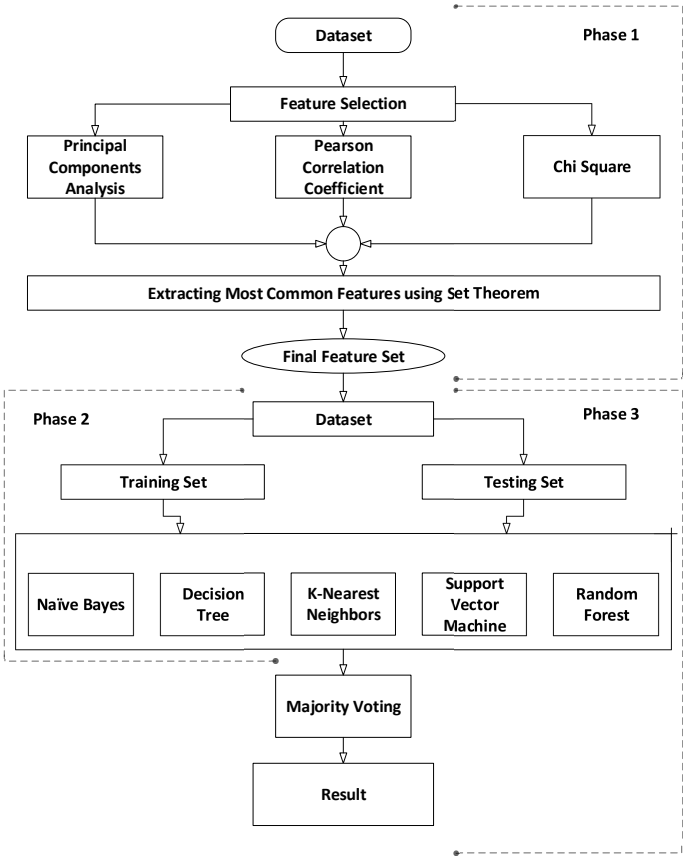


Fig. 1. Methodology

B. Phase 2

The model is trained in this phase. NB, DT, KNN, SVM, and DT classifiers are used to train the model. In this context, the cancer datasets are cross validated and train-test split for the values of cross validation k=5 and 3 and train and test split ratio of 80% and 70%.

Classifiers that play the most important role in this research work are described in this phase. A generalized knowledge of the classifiers are summarized in this part.

1) Naive Bayes(NB): Naive Bayes is one of the most effective, simple, fast and probabilistic based machine learning classifier. The classifier makes classifications using the Maximum Posteriori decision rule. A simple Bayesian network can represent it. It is popular for text classification. The Classifier is also popular for the solution for problems such as spam detection [15].

$$Posterior = \frac{Prior \times Likelihood}{Evidence} \tag{5}$$

Naive Bayes needs less training data. The classifier is highly scalable.

2) Decision Tree(DT): Decision Tree is a simple and widely used non-parametric supervised learning method for both classification and regression tasks [16]. A straightforward idea is applied to solve the classification problem. In this classifier the data is continuously split according to a certain parameter. In our problem we used two types of decision tree classifier: ID3 and J48.ID3 or Iterative Dichotomiser 3 builds decision tree by using a top-down, greedy search through the given sets of training data to test each attribute at every node [17].

3) K-Nearest Neighbors(KNN): K Nearest Neighbors (KNN) is an algorithm that keeps all available cases and classifies new cases based on a similarity action [18]. In both classification and regression predictive problems KNN can be used for solution. The boundary becomes smoother with different value of K. The training error rate and the validation error rate are two parameters to access on different K-value. KNN is very much helpful in the field of classification techniques because of: robust to noisy training data, effective if the training data is large, it has large search problem to find nearest neighbors, it also need storage of data, computation cost is high of KNN.

4) Support Vector Machine(SVM): A Support Vector Machine (SVM) is a discriminative classifier [19]. Decision planes that define decision boundaries, is the basic concept of this classifier. A separating hyperplane describes this classifier. For supervised learning, an optimal hyperplane is returned by this algorithm which categorizes new examples. A line dividing a plane in two parts by a line is the hyperplane in two dimensional space where in each class lays in either side. Both regression and classification tasks are supported and multiple continuous and categorical variables can be handled by this classifier. A dummy variable is created with case values as either 0 or 1 for categorical variables. This classifier has a regularization parameter which can avoid over-fitting. The kernel trick is used by it. Kernel models can be quite sensitive to over-fitting the model selection criterion.

5) Random Forest(RF): Random Forest Classifier is an ensemble algorithm. A set of decision trees are created by random forest classifier from randomly selected subset of training set [20]. The votes from different decision trees are summed by it to determine the final class of the test object. From any decision tree, weight concept can be applied for considering the impact of result. Tree with low error rate are

given high weight value and vice versa. The decision impact of trees would be increased with low error rate. Total number of trees to be generated and decision tree related parameters like minimum split, split criteria are the basic parameters of this classifier.

These different classification algorithms are used to get the test accuracy of the datasets. At first the model is trained with 80% and 70% of the data, and then tested with the left data accordingly. Secondly, the model is trained with 10 and 5 folds. Each time the model is tested with the remaining data in the combination.

C. Phase 3

The final and the resulting process is this phase. After training in the previous phase, the model is tested, given the test data. The results are then voted by majority voting to obtain the final outcome.

The majority voting [21] is one of the simplest and most instinctive ensemble compound techniques.

If a prediction is made by a model for each test case and final output prediction is that one which collects more than half of the votes then it is called majority voting. The ensemble method could not make a stable prediction if no predictions get more than half of the votes. In general, the majority vote classifier is defined as a series of votes. In other word, a classification of an unlabeled instance is applied on a class that gets the maximum number of votes.

Majority voting is used to achieve the final result. To get the final result, each, and every accuracy is passed through voting which decides the final output based on majority, and the accuracy gained through this process is considered to be the decided result coming at the end of the all simulation processes.

IV. Result Analysis

For this research purpose three datasets have been used. The datasets are known as Wisconsin Breast Cancer(Original) Dataset [22], Wisconsin Breast Cancer(Diagnostic) Dataset [23], and Wisconsin Lung Cancer Dataset [24]. All of the mentioned datasets are obtained from UCI Machine Learning Repository.

A. Wisconsin Breast Cancer(Original) Dataset

Wisconsin Breast Cancer(Original) Dataset consists of 699 instances having 10 attributes. There are two classes named Benign and Malignant. The class distribution of this dataset is 458(65.5%) for Benign and 241(34.5%) for Malignant.

1) Wisconsin Breast Cancer(Diagnostic) Dataset: Wisconsin Breast Cancer(Diagnostic) Dataset consists of 569 instances having 32 attributes. There are two classes named Benign and Malignant. The class distribution of this dataset is 357(62.7%) for Benign and 212(37.26%) for Malignant.

2) Wisconsin Lung Cancer Dataset: Wisconsin Lung Cancer Dataset consists of 32 instances having 57 attributes. Attribute 1 is the class label, All predictive attributes are nominal, taking on integer values 0-3. The class distribution of this dataset is 9, 13, and 10.

Here, The simulation results of our An Efficient Ensemble Method for Cancer Detection model, abbreviated as AEEMCD are demonstrated, and a comparison is drawn according to the obtained results. The results achieved from the cancer detection system under study are mounted side by side using table to give a clear view of the classifier performance differences. Results shown here are categorized according to datasets. Tables given in this section are spread over Wisconsin Breast Cancer(Original), Wisconsin Breast Cancer(Diagnostic), and Wisconsin Lung Cancer datasets respectively, Accordingly final feature sets are: 1-8, 1-11, and 1-21

TABLE I
Simulation Result for Train Test Split on Wisconsin Breast Cancer(Original) Dataset

Model	Split	Accuracy	KS	RMSE	Precision	Recall	F-Measure
NB	80-20	0.792857	0.4627	0.3519	0.905	0.904	0.904
	70-30	0.809524	0.4826	0.3323	0.803	0.81	0.806
KNN	80-20	0.892857	0.9098	0.209	0.957	0.956	0.956
	70-30	0.909524	0.7558	0.2617	0.908	0.91	0.908
SVM	80-20	0.828571	0.964	0.1325	0.983	0.982	0.982
	70-30	0.847619	0.5861	0.3904	0.843	0.848	0.845
RF	80-20	0.95	0.8757	0.1824	0.94	0.939	0.939
	70-30	0.966667	0.9111	0.2006	0.967	0.967	0.966
J48	80-20	0.935714	0.8568	0.2617	0.93	0.934	0.93
	70-30	0.952381	0.869	0.2081	0.954	0.952	0.952
ID3	80-20	0.935714	0.8193	0.2555	0.936	0.93	0.931
	70-30	0.952381	0.8738	0.2182	0.952	0.952	0.952
AEEMCD	80-20	0.935717	0.8272	0.2535	0.936	0.936	0.87
	70-30	0.952381	0.8707	0.2182	0.953	0.952	0.951

The performance measurement of different classifiers simulated on Wisconsin Breast Cancer (Original) dataset are embedded in Table I. In this simulation Random Forrest performs best in all categories and Naïve Bayes performs worst in accuracy and Kappa Statistics. We found RMSE the least in Random Forrest (80-20) simulation. The best accuracy we found after simulation is 0.967 by Random Forrest (70-30) model. The accuracy of our proposed model AEEMCD performs best in (70-30) simulation which is 0.951 and in (80-20) simulation gives only 0.870.

TABLE II
Simulation Result for Cross Validation on Wisconsin Breast Cancer(Original) Dataset

Model	Fold	Accuracy	KS	RMSE	Precision	Recall	F-Measure
NB	10	0.835479	0.5564	0.3243	0.83	0.835	0.832
	5	0.83834	0.5625	0.3256	0.833	0.838	0.834
KNN	10	0.938484	0.8347	0.2153	0.938	0.938	0.937
	5	0.937053	0.8312	0.2244	0.937	0.937	0.936
SVM	10	0.856938	0.6177	0.3782	0.853	0.853	0.854
	5	0.856938	0.6177	0.3782	0.853	0.853	0.854
RF	10	0.972818	0.9296	0.1732	0.973	0.973	0.973
	5	0.974249	0.9334	0.177	0.974	0.974	0.974
J48	10	0.967096	0.9147	0.1731	0.967	0.967	0.967
	5	0.951359	0.8755	0.2027	0.952	0.951	0.952
ID3	10	0.969957	0.9221	0.1731	0.97	0.97	0.97
	5	0.971388	0.927	0.1692	0.972	0.971	0.972
AEEMCD	10	0.966996	0.9219	0.1733	0.970	0.970	0.970
	5	0.971388	0.9254	0.1692	0.971	0.971	0.971

In table II we have cross validated the dataset, clarifies the performances of different classifiers simulated on Wisconsin Breast (original) Cancer dataset. In this simulation the accuracy is between 0.835 and 0.974. We have found RMSE the least in Decision Tree (ID3) 10 fold simulation. The best accuracy we have found after simulation is 0.974 by Random Forrest 5 fold simulation. Our proposed model AEEMCD performs best in 5 fold simulation which is 0.971 and in 10 fold simulation gives 0.967 which is almost similar. Our Proposed model also gives the highest kappa statistics which is 0.952. From the previous two tables, we have come to a decision that Random Forrest classifier works best on Wisconsin Breast Cancer (Original) dataset.

TABLE III
SIMULATION RESULT FOR TRAIN TEST SPLIT ON WISCONSIN BREAST CANCER(DIAGNOSTIC) DATASET

Model	Split	Accuracy	KS	RMSE	Precision	Recall	F-Measure
NB	80-20	0.903509	0.8046	0.3105	0.905	0.904	0.904
	70-30	0.918129	0.8273	0.2860	0.919	0.918	0.918
KNN	80-20	0.956140	0.9098	0.2090	0.957	0.956	0.956
	70-30	0.959064	0.9124	0.2018	0.959	0.959	0.959
SVM	80-20	0.982456	0.9640	0.1325	0.983	0.982	0.982
	70-30	0.976608	0.9501	0.1529	0.977	0.977	0.977
RF	80-20	0.938596	0.8757	0.1824	0.940	0.939	0.939
	70-30	0.964912	0.9246	0.1507	0.966	0.965	0.965
J48	80-20	0.929825	0.8568	0.2617	0.930	0.930	0.930
	70-30	0.929825	0.8502	0.2624	0.930	0.930	0.930
ID3	80-20	0.929825	0.8193	0.2555	0.936	0.930	0.931
	70-30	0.928525	0.8240	0.2587	0.931	0.930	0.930
AEEMCD	80-20	0.956140	0.8831	0.2094	0.957	0.956	0.956
	70-30	0.953216	0.8777	0.2163	0.953	0.953	0.953

In table III where we split the dataset, clarifies the performances of different classifiers simulated on Wisconsin Breast (Diagnostics) Cancer dataset. The best accuracy we have found after simulation is 0.974 by Support Vector Machine (80-20) simulation. We found highest Kappa Statistics in Support Vector Machine simulation. The accuracy of our proposed model AEEMCD in (80-20) simulation which is the best resulting 0.956 and in 10 fold simulation gives 0.953 which are almost similar.

TABLE IV
SIMULATION RESULT FOR CROSS VALIDATION ON WISCONSIN BREAST CANCER(DIAGNOSTIC) DATASET

Model	Fold	Accuracy	KS	RMSE	Precision	Recall	F-Measure
NB	10	0.926186	0.8418	0.2648	0.926	0.926	0.926
	5	0.922671	0.8346	0.2708	0.923	0.923	0.923
KNN	10	0.961336	0.9171	0.1963	0.961	0.961	0.961
	5	0.959578	0.9133	0.2006	0.960	0.960	0.960
SVM	10	0.978910	0.9545	0.1452	0.979	0.979	0.979
	5	0.978910	0.9545	0.1452	0.979	0.979	0.979
RF	10	0.959578	0.9128	0.1731	0.960	0.960	0.959
	5	0.961336	0.9171	0.1763	0.961	0.961	0.961
J48	10	0.929701	0.8502	0.2608	0.930	0.930	0.930
	5	0.938489	0.8691	0.2422	0.939	0.938	0.939
ID3	10	0.947276	0.8665	0.2296	0.948	0.947	0.947
	5	0.926186	0.8123	0.2717	0.956	0.981	0.968
AEEMCD	10	0.952548	0.8765	0.2178	0.952	0.953	0.952
	5	0.943761	0.8765	0.2178	0.952	0.953	0.952

In table IV where we cross validate the dataset, clarifies the performances of different classifiers simulated on Wisconsin Breast (Diagnostic) Cancer dataset In this simulation Support Vector Machine performs best in all categories. The best accuracy we found after simulation is 0.978 by Support Vector Machine both 5 & 10 fold simulation. The accuracy of our proposed model AEEMCD performs best in 10 fold simulation which is 0.952 and in 10 fold simulation gives 0.943 which are almost similar. Our Proposed model also gives the highest Kappa Statistics which is 0.875.

From the previous two tables, it can be concluded to a decision that Support Vector Machine classifier works best on Wisconsin Breast Cancer (Diagnostic) dataset.

TABLE V
SIMULATION RESULT FOR TRAIN TEST SPLIT ON WISCONSIN LUNG CANCER DATASET

Model	Split	Accuracy	KS	RMSE	Precision	Recall	F-Measure
NB	80-20	0.3333	0.2000	0.6634	0	0.333	0
	70-30	0.50000	0.2754	0.5461	0.700	0.500	0.493
KNN	80-20	0.66667	0.4286	0.4490	0.667	0.667	0.683
	70-30	0.700000	0.5522	0.4233	0.0.730	0.700	0.688
SVM	80-20	0.166667	0.0	0.5443	0.083	0.167	0.111
	70-30	0.300000	0.0145	0.5018	0.550	0.300	0.327
RF	80-20	0.333333	0.1111	0.4722	0.750	0.333	0.378
	70-30	0.50000	0.2647	0.459	0.650	0.500	0.520
J48	80-20	0.50000	0.3333	0.5800	0.875	0.500	0.500
	70-30	0.400000	0.0769	0.6325	0	0.4	0
AEEMCD	80-20	0.333333	0.1111	0.6667	0.750	0.333	0.378
	70-30	0.500000	0.2647	0.5774	0.650	0.500	0.520

In table V where we have split the dataset, clarifies the performances of different classifiers simulated on Wisconsin Lung Cancer dataset. In this simulation all data come scrambled. Best accuracy is given by KNN and worst accuracy is given by Support Vector Machine. The accuracy of our proposed model AEEMCD performs best in (70-30) simulation which is 0.500 and in (80-20) simulation the model gives 0.333 accuracy.

TABLE VI
SIMULATION RESULT FOR CROSS VALIDATION ON WISCONSIN LUNG CANCER DATASET

Model	Fold	Accuracy	KS	RMSE	Precision	Recall	F-Measure
NB	10	0.625000	0.4311	0.4823	0.631	0.625	0.627
	5	0.593750	0.3791	0.4918	0.598	0.594	0.595
KNN	10	0.531250	0.3033	0.5346	0.507	0.531	0.504
	5	0.593750	0.3962	0.4957	0.589	0.954	0.574
SVM	10	0.500000	0.2493	0.4564	0.489	0.500	0.492
	5	0.562500	0.3353	0.4539	0.563	0.563	0.563
RF	10	0.531250	0.2931	0.4277	0.540	0.531	0.533
	5	0.531250	0.2760	0.4220	0.547	0.531	0.534
J48	10	0.406250	0.0979	0.5707	0.406	0.406	0.406
	5	0.468750	0.1893	0.5865	0.465	0.469	0.6459
AEEMCD	10	0.531250	0.2910	0.5590	0.524	0.531	0.527
	5	0.531250	0.2982	0.5590	0.514	0.531	0.518

In table VI where we have cross validated the dataset, clarifies the performances of different classifiers simulated on Wisconsin Lung Cancer dataset. In this simulation all of the data also appear to be scrambled but not alike the previous one. Best accuracy is given by Naïve Bayes classifier and worst accuracy is given by Decision Tree (J48) Classifier. The accuracy of our proposed model AEEMCD performs best in both simulation which is 0.531.

From the previous two tables, it can be seen that the resulting accuracy from the tables are not satisfactory. We need to work more on this dataset.

V. Conclusion And Future Work

From table I, and II, it is obvious that the maximum accuracy is obtained by Random Forrest considering Wisconsin Breast Cancer (Original) Dataset. As per table III, and IV, we can notice that the maximum accuracy is achieved by Support Vector Machine considering Wisconsin Breast Cancer (Diagnostic) Dataset. From table V, and VI, it is visualized that the highest accuracy is obtained by K-Nearest Neighbor considering Wisconsin Lung Cancer Dataset. From above it can be concluded to a decision that, any individual classifier can perform quite better for any kind of dataset, which is not certain. For that reason, we propose a model that gives much better results than average. To analyze table I to VI, numbers show that our designed model performs better than individuals. In this experiment, we used two very common indicators which are mean of the two types of errors named Root Mean Square Error(RMSE), and Kappa statistics. Kappa statistic is utilized to calculate the classification accuracy for discriminating between the reliability of the data collected and their credibility.

The most common technique for evaluating the true accuracy of a given model is to select the most common features performing set union operation, and then apply cross validation which is based on splitting the available dataset into a training set, and a validation set and finally get the result after applying majority voting on classifiers.

The main objective of this paper is to explore the impact of the most common features of the dataset into the result by reducing the features applying set union operation in the proposed hybrid feature selection method. We have also used majority voting to find the most perfect result from the classifier. The false positive rate (FP) of the proposed model is 2.11% with a true positive rate (TP) of 97.89%. This model focuses on the misuse detection only, so there is a great scope for further research. For future work, we would like to use a great quantity of classifiers, as we should try to minimize the complexity of the classification model. Again simulation results for Wisconsin Lung Cancer dataset are not so satisfactory. That would be a place where we are very interested to work on.

References

[1] Z. Pang, D. Zhu, D. Chen, L. Li, and Y. Shao, "A computer-aided diagnosis system for dynamic contrast-enhanced mr images based on level set segmentation and relieff feature selection," *Computational and mathematical methods in medicine*, vol. 2015, no. 2015, 2015.

[2] M. A. H. C. J. P. Ian H Written, Eibe Frank, *Data Mining Concepts*. Elsevier, 2016.

[3] V. Bolón-Canedo, N. Sánchez-Maroño, and A. Alonso-Betanzos, "A review of feature selection methods on synthetic data," *Knowledge and information systems*, vol. 34, no. 3, pp. 483–519, 2013.

[4] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157–1182, 2003.

[5] H. Asri, H. Mousannif, H. Al Moatassime, and T. Noel, "Using machine learning algorithms for breast cancer risk prediction and diagnosis," *Procedia Computer Science*, vol. 83, pp. 1064–1069, 2016.

[6] G. Rashmi, A. Lekha, and N. Bawane, "Analysis of efficiency of classification and prediction algorithms (naïve bayes) for breast cancer dataset," in *Emerging Research in Electronics, Computer Science and Technology (ICERECT), 2015 International Conference on*. IEEE, 2015, pp. 108–113.

[7] M. U. Salma *et al.*, "Bat-elm: a bio inspired model for prediction of breast cancer data," in *Applied and Theoretical Computing and Communication Technology (iCATccT), 2015 International Conference on*. IEEE, 2015, pp. 501–506.

[8] D. Bazazeh and R. Shubair, "Comparative study of machine learning algorithms for breast cancer detection and diagnosis," in *Electronic Devices, Systems and Applications (ICEDSA), 2016 5th International Conference on*. IEEE, 2016, pp. 1–4.

[9] S. K. Mandal, "Performance analysis of data mining algorithms for breast cancer cell detection using naïve bayes, logistic regression and decision tree," *International Journal Of Engineering And Computer Science*, vol. 6, no. 2, 2017.

[10] R. B. Rao, J. Bi, G. Fung, M. Salganicoff, N. Obuchowski, and D. Naidich, "Lungcad: a clinically approved, machine learning system for lung cancer detection," in *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2007, pp. 1033–1037.

[11] A. G. Karegowda, M. Jayaram, and A. Manjunath, "Feature subset selection problem using wrapper approach in supervised learning," *International journal of Computer applications*, vol. 1, no. 7, pp. 13–17, 2010.

[12] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, vol. 374, no. 2065, 2016.

[13] M. M. Mukaka, "A guide to appropriate use of correlation coefficient in medical research," *Malawi Medical Journal*, vol. 24, no. 3, pp. 69–71, 2012.

[14] M. L. McHugh, "The chi-square test of independence," *Biochemia medica: Biochemia medica*, vol. 23, no. 2, pp. 143–149, 2013.

[15] H. Zhang and D. Li, "Naïve bayes text classifier," in *Granular Computing, 2007. GRC 2007. IEEE International Conference on*. IEEE, 2007, pp. 708–708.

[16] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*. Elsevier, 2011.

[17] P. Chandrasekar, K. Qian, H. Shahriar, and P. Bhattacharya, "Improving the prediction accuracy of decision tree mining with data preprocessing," in *Computer Software and Applications Conference (COMPSAC), 2017 IEEE 41st Annual*, vol. 2. IEEE, 2017, pp. 481–484.

[18] W. Shang, H. Huang, H. Zhu, Y. Lin, Z. Wang, e. Y. Qu, Youli", J. Liu, Y. Wang, Y.-m. Cheung, H. Yin, L. Jiao, J. Ma, and Y.-C. Jiao, "An improved knn algorithm – fuzzy knn." Berlin, Heidelberg: Springer Berlin Heidelberg, 12 2005, pp. 741–746.

[19] Y. Yang, J. Li, and Y. Yang, "The research of the fast svm classifier method," in *Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2015 12th International Computer Conference on*. IEEE, 2015, pp. 121–124.

[20] A. Pretorius, S. Bierman, and S. J. Steel, "A meta-analysis of research in random forests for classification," in *Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech), 2016*. IEEE, 2016, pp. 1–6.

[21] S. ABD MANAF, N. MUSTAPHA, M. N. SULAIMAN, N. A. HUSIN, M. SHAH ZAINUDDIN, M. SHAFRI, and H. ZULHAIDI, "Majority voting of ensemble classifiers to improve shoreline extraction of medium resolution satellite images." *Journal of Theoretical & Applied Information Technology*, vol. 95, no. 18, 2017.

[22] W. H. Wolberg and O. Mangasarian, "Breast cancer wisconsin (original) data set," *UCI Machine Learning Repository, University of California*, 1989.

[23] W. H. Wolberg, W. N. Street, and O. L. Mangasarian, "Breast cancer wisconsin (diagnostic) data set," *UCI Machine Learning Repository [http://archive. ics. uci. edu/ml/]*, 1992.

[24] Z. Q. Hong and J. Y. Yang, "Optimal discriminant plane for a small number of samples and design method of classifier on the plane," *Pattern Recognition*, 1991.

Big Data Mining in the Presence of Concept Drifting

Tabassum Siraj, Efrana Jannat, Warafta Rasul, Meher Afroz Chowdhury,
Swakkhar Shatabda, and Dewan Md. Farid*
Department of Computer Science & Engineering, United International University
United City, Madani Avenue, Badda, Dhaka 1212, Bangladesh
Email: tabassum.siraj13@gmail.com, efrana.jannat@gmail.com, warafta@gmail.com,
meherafrozmac@gmail.com, swakkhar@cse.uiu.ac.bd, dewanfarid@cse.uiu.ac.bd

***Abstract*—Concept drift in big data mining is an absolute, highly demanding research issue in this digital era. A concept in "concept drift" involved in the field of data mining (DM) and machine learning (ML) studies is referred to the relationship between input features and class variables. In real-life classification problems, a concept can be changed or new concepts may appear over the time. Innumerable data mining classification models, new methods are propounded every day to enhance the solution to this raised issue. In this paper, we have extended our vision to attempt an investigation in resolving the issue of concept drift detection, and carried out analysis for creating a better method to adapt to the newly arriving concepts. We have addressed the issue of concept drifting for mining big data, and presented an evolutionary, concept-adaptive, rule-based approach that classifies novel class instances in "concept drifting". The proposed method clusters the big data and extracts decision rules from each cluster for classifying instances with existing classes. For classifying new unlabelled instances, the proposed method first examines and ensures if the new instances belong to the clusters in existence or not. If the new instance does not appear to belong to the existing clusters, we have considered this instance as a novel class instance. Then, we have extracted new classification rules from the new instances data and added these new rules with existing rules. Performance evaluation tests of the proposed method have been conducted using a number of datasets provided by UCI (University of California, Irvine) machine learning repository. The calculated result proves proposed approach as an effective and efficient means of detecting novel class, which implies concept drift identification at the same time.**

Keywords*—Concept Drift; Big Data; Data Mining; Machine Learning; Novel Class Instance;

I. Introduction

Big Data analysis becomes an attractive research area in ML(machine learning) and DM(data mining) at present time [1], [2], [3]. Detection of concept drift in streaming big data captured so much attention of intelligence computational researchers [4], [5]. In big data a large volume of new data generated continuously with time that cannot be restricted to any pre-defined order [6]. Big data are not always accurately stationary. A new data, in a non-stationary environment, could have different knowledge pattern and can probably change the data distribution and makes the current prediction model unsuitable. This circumstance is recognized as concept drift [7]. Concept drift occurs due to difference in the assignment of data in disparate time phases or change the pattern enciphers [8], [9]. Generally, big data should have the following four properties: (1) volume, (2) variety, (3) velocity, and (4) veracity. Volume refers to the amount of generated and stored data. Variety refers to the number of types and nature of data. Velocity is at which pace the data is being processed. Veracity is the quality of data that affects the accurate analysis.

The inauguration of big data applications with huge volumes of high-dimensional instances, often become accessible in stream mode, and impose major obstacles to all learning tasks, among which classification is significant. [10]. On account of numerous reasons, where absent features or acquisition failures are included, the data vectors that are available might often have missing entries, which due to improper treatment, can result in classification. High-dimensional data normally exhibit recognizable patterns that can be leveraged to effectively embed them into low-dimensional spaces. Classification is the function of learning such embedding maps from a finite number of training data with their correct class labels. Big Data applications are rising high and being researched by the region of computer science, in which applications require distributed and present classification function and pattern recognition of huge data sets accordingly which are collected from many different sensor networks, medical data, online form registrations, video and imaging systems, etc. In our paper, we have proposed a new approach for big data mining in the presence of concept drift. Our proposed approach applies clustering technique with adaptive rule-based classifier to classify the unlabelled instances in big data and also predict the novel class instances in big data. The suggested method clusters the big data initially and then extracts classification rules from each of the cluster. To classify new unlabelled instances, it checks if the instances belong to the existing clusters or not. If the new instance does not belong to any of the existing clusters, it indicates that concept drift happened. Afterwards, it extracted new classification rules from the new instances and added these new rules with existing rules. We have assessed the execution of propound formula on conventional datasets from University of California, Irvine (UCI) machine learning repository [11] and the results that are originated from experiments proved the methodology we suggested as effective and efficient in classifying data with identifying novel class instances in data.

The remaining portions of this manuscript is designed as Section II elaborately examines the related works. Section III propounds the proposed method. Section IV reviews the experiment results. Finally, Section V explores the conclusions and future work plans.

II. Related Work

In 2017, Liu et al. [12] proposed a concept drift detection method on the basis of subspace learning that mostly depends on angle optimized global embedding (AOGE) and principal component analysis (PCA). AOGE and PCA algorithms analyze the projection angle and the projection variance for concept drift identification. With observation, the change of objective functions for data streams, concept drift is distinguished. If concept drift is present, it is re-organized, based upon subspace learning which utilizes the new data stream patch. After that, the data streams are processed with dimension reduction for classification procedure. Lu et al. [13] proposed a chunk-based incremental learning method called Dynamic Weighted Majority for Imbalanced Learning (DWMIL), which analyzes concept drift in data streams and class imbalance problem. DWMIL avails an ensemble framework that assigns the weight of the base classifiers dynamically with respect to their performance on current data chunk. DWMIL is an incremental learning model and achieves high performance by keeping limited number of classifiers. In 2016, Haque et al. [14] proposed an active (semi-supervised) framework to classify data streams. It can detect concept drift and ensures chunk limit in a robust way by identifying crucial changes in classifier confidence. Additionally, it utilizes confidence scores to acutely choose specified number of data instances that arrives from the newest chunk as flags to update the classifiers later. They hypothesized the use of confidence estimators and the impact of concept drift on classifier confidence.

In 2015, Yang and Fong [15] used three representative singletree induction algorithms (VFDT, ADWIN, iOVFDT) to explore the occurrence of concept drift. iOVFDT balances accuracy, tree size, and learning speed with a compact size of model and less usage of memory. The results conclude that ioVFDT displayed better performance compared to the other two aforementioned algorithms, due to its adaptive tie threshold, making it suitable for both artificial and real world concept drift data streams. In 2015, Ortiz et al. [16] created an approach named Fast Adapting Ensemble (FAE), which is an ensemble method designed to adapt to concept drifts rapidly and specifically deals with recurring concepts with expertise. FAE harbors a set of inactive base classifiers that change their state to being active within a very short span as the concept they represent re-emerges. FAE uses a drift detector (this helps to process the abrupt concept drifts) to conclude when to construct and include a new base classifier. FAE also uses a weighted majority vote to acquire the global ensemble decision and establishes a formula for adjusting the base classifiers weights, which permits the algorithm to expand or reduce the weights according to their performances. Krawczyk, and Woźniak [17] proposed a novel modification of weighted one-class support vector machine transformed to the relentless streaming data analysis. Their hypothesis prioritized the improved alterations of weights used by WOSVM. They implemented the incremental training of weights, which was followed by forgetting mechanism. They finalized with the theory that the efficient one-class classification model for data streams along with concept drift would be reprized from combination of incremental learning assigned with highest weights to object that arrive from new data chunk with forgetting by sigmoid based function.

In 2015, Wang and Abraham [18] concentrated on concept drift detection that affects binary classification models and demonstrates a framework (Linear Four Rates, LFR), which can identify the concept drift and distinguishes the updated data points of the new concept. The wide-range vision of LFR makes it possible to function for both batch and stream datasets, which are imbalanced datasets. It can also use user-specified parameters that are automatically easy to understand, different from other methods of detecting concept drift. LFR stands out with better performance than the existing popular approaches in respect of early concept drift detection, low false alarm rate and high detection rate across the various sorts of concept drifts. Mirza et al. [19] presented an Ensemble of Subset Online Sequential Learning Machine (ESOS-ELM) to learn from class imbalanced data that occurs from drifting data streams. ESOS-ELM is an ensemble approach for classification in imbalanced environment. It's also an ELM-Store module, which harbours information of old concepts and a change detector that can instantly detect concept drifts. In ESOS-ELM, the ensemble is processed with the equitable subsets of streaming data. In 2015, Kanoun and Schaar [20] addressed the limitations of streaming big data by suggesting an online energy-efficient scheduler that can promptly learn the environmental changes by reinforcement learning techniques to increase the QoS (i.e. throughput and output quality) even with restrictions for resource use. They identified the scheduling problem as a Stochastic Shortest Path problem (SSP) and introduced a learning algorithm that reinforces itself to adapt to environmental changes and resolve this barrier even with recurrence of concept drift. L. Rutkowski et al. (2015) [21] proffered a new criteria that apply the theory of splitting nodes of decision trees for classifying data streams. The nodes are split according to the count of impurity, otherwise addressed popularly as misclassification error. The proposed measure or criteria observes if the selected ideal one among all the resident properties of current series of elements in node is equally the best in the entire stream or not. Also, this criterion in question was blended along with the criterion developed from Gini index. The synthesized algorithm reproduced quite efficient results while processing data streams in different environments.

In 2013, Farid et al. [22] introduced an adaptive ensemble learning classifier to handle deficiencies that particularize concept drift, infinitesimal length, minimal labeled data, and concept evolution for detecting novel class in data streams with concept drifting. The suggested model frequently upgrades itself with freshly arrived data points to keep up with most up-to-date concepts in data stream. The majority of weighted predictions among the classifiers in model classifies unlabeled observations. The model, M, labels the instances present in each sub-stream and a new classifier are adapted with the most contemporary dataset.

III. Big Data Mining with Concept Drifting

A. Concept Drift

Big Data is a vast source of instances that appear in an organized order. So, it enforces certain restraints on the learning system, which are unable in accomplishing the destined result with the help of the orthodox algorithms from this realm. Suppose, a stream of data made up of some states which are grouped, $G = \{G_1, G_2, \cdots, G_n\}$, where G_i is originated by a

diffusion D_i. If we consider streaming data which is stationary where the order of instances are symbolized as $D_j = D_{j+1}$, the transition can be marked as $G_j \rightarrow G_{j+1}$. Meanwhile, the behaviour of data may change in time due to many reasons in our modern and realistic problems. This event is addressed as concept drift. The fundamental problem is the alteration in the input and output(I/O) with time. A "drift" or change from a standard concept occurs because of the "subtle changes" in attribute values of big data. Concept drift possibly appears where data is collected over time, with or without predictions on supervised learning problems. Typically, these are known as online learning problems, where change is expected in the data with time. Concept drift is basically bifurcated into (1) real concept drift, and (2) virtual concept drift. The type of concept drift that can alter the decision boundaries also known as posterior probabilities and can also have an effect on unconditional probability density function is called real concept drift which imposes a threat on learning system. Another type of concept drift is virtual that does not affect the decision boundaries, but transform conditional probability density functions, thus it cannot overpower the recently used learning methods. But, it is mandatory for us to identify concept drift. The four prior applications that are popularly used to deal with concept drift are: concept drift detectors, online learners, sliding windows, and ensemble learners. We might need different concept drift detection and handling procedures depending on different situations. Several ways to handle concept drift are: periodically update, detect and choose model, learn the change, weight data, and data preparation.

B. Proposed Approach

If concept drift is existent in big data, classifying it will be quite an arduous task. Because, machine learning models are trained on fix number of class labels and learning models misclassify the novel class instance as existing classes in classification process. It's very common in real-life classification problem that observations with latest class may manifest itself over the time. A rule-based learning model has been propounded to classify large data with clustering technique that can classify unknown instances and detect novel class instances. The proposed method divides the big data into small sub-datasets, so that each of the sub-data can fit into the computer memory and easy to manage. Then we cluster the each sub-data into several clusters. Data clustering is an approach of segmenting data observations that are similar to each other [23]. Then classification rules are extracted from clusters using decision tree (C4.5) algorithm [24], as extracting classification rules from trees are very well known approach [25]. C4.5 decision tree algorithm is a top-down divide and conquer recursive process, which classify the unseen instances based on the several seen instances [26], [27]. C4.5 decision tree algorithm uses Gain Ratio to select the best splitting attribute. Generation of a classification rule from individual leaf of the decision tree. From the root to leaf defines a path that corresponds to a rule-set. The perk of extracting rules from tree is that it's easy to marge and can be executed in any order. Latest rules can be merged to previously extracted rules. Where as decision tree merging is quite complex process. For classifying new unlabelled instances, the proposed method first checks the new instances belong to the existing clusters or not. If the contemporary observations does not fit in to the existing clusters, then we have considered this observations as a novel class instance. Then, we have extracted new classification rules from the new instances and added these new rules with existing rules. Table I presents the frequently used notations and terminologies in this paper. Fig. 1 describes the proposed method for mining big data in concept drifting employing clustering with rule-based classifier, which is illustrated in algorithm 1.

TABLE I: Frequently used notations and terminologies.

Notation	Terminology
D_{Big}	Big Data
D_i	Small data that into the memory
C_j	Cluster
DT_j	Decision tree
$rule$	Classification rule
x_{New}	New data point/ instance
x_{Novel}	New instance with novel class

Algorithm 1 Big Data Mining with Concept Drift

Input: *Big Data, D_{Big}; C4.5 Classifier;*
Output: *Rule-set*; // Classification rules.
Method:

Rule-set = $\emptyset$;
divide D_{Big} into small sub-data sets $D_1, D_2, \cdots, D_N$;
for i = 1 to N **do**
  cluster D_i into several clusters $C_1, C_2, \cdots, C_M$;
  for j = 1 to M **do**
    build decision tree, $DT_j \leftarrow C_j$ using C4.5 classifier;
    if $error(DT_j) \leq 0.5$ **then**
      Rules $\leftarrow DT_j$; // extracting the rules from DT_j.
    end if
  end for
  Rule-set = *Rule-set* $\cup$ *Rules*;
end for

To classify new instance, x_{New}:

if x_{New} belongs to any C_j **then**
  classify x_{New} with existing rules;
else
  extract *Rule* $\leftarrow x_{New}$;
end if
add new rule with existing rules;

IV. Experimental Analysis

This section constitutes details of datasets, experimental setup and performances.

A. Datasets

Datasets are assembled from UCI machine learning repository [11]. Table II provides a brief interpretation of the datasets. NSL-KDD dataset is used quite frequently as a standard for assessment of concept drift detection approaches. This dataset is the updated form that has solved the shortcomings present in the KDD99 dataset. In Table III, we presented the variety instructed samples of different genre used in the NSL-KDD dataset. There are 683 data instances and 35 categorical attributes with 19 class values in the large soybean dataset. Another dataset employed in our work is image segmentation data

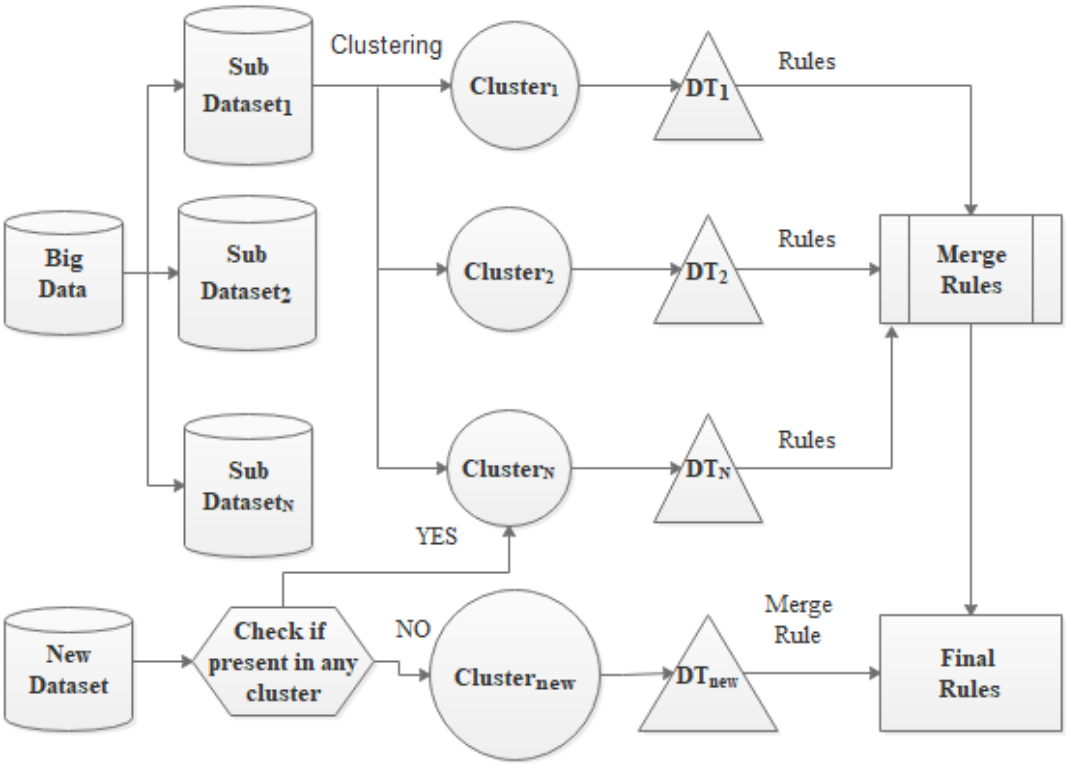


Fig. 1: Propound method for mining big data in presence of concept drift.

of 7 outdoor images, where instances were drawn randomly. Categorization with respect to each pixel was generated from handpicked images. The 3X3 region was assigned for every instances in the employed dataset. The basic target for using this specific dataset is accomplishing a factual reasoning for image segmentation research. This database has 19 features and 7 categories and 2310 data instances.

TABLE II: Dataset details.

Dataset Name	Features	Category	Observation	Target
NSL-KDD	41	Real & Categorical	25192	23
Soybean-Dataset	35	Categorical	683	19
Segment-Test	19	Numeric	2310	7

TABLE III: The NSL-KDD dataset.

Category	Target	Observation
Normal	normal	13449
Denial_of_Service (DoS)	neptune, land, back tenddrop, smurf, pod	9234
Remote_to_User (R_2_U)	warezmaster, guess_passwd, impa, spy, ftp_write, warezclient, multihop, phf	209
User_to_Root (U_2_R)	rootkit, buffe_overflow, loadmodule, perl	11
Probing	protsweep, nmap ipsweep, satan	2289
	Total = 23	Total = 25192

B. Experimental Setup

We have enacted the proposed hypothesis in Java using NetBeans IDE 8.2. NetBeans is an open-source integrated development environment that gives assistance to the development for all types of Java application. The code for decision tree (C4.5) is adopted from Weka 3.8, which is a data mining software in Java [28].

C. Experimental Results

Tables IV to VI describe how the presence of novel classes have an impact on the performance of the datasets that we have used. The outcomes represent that our proposed method, which is taught with a stable amount of sets, displays satisfactory fitness in learning and classifying new concepts, and overshadows the conventional C4.5 decision tree induction algorithm. Table VII shows the summary of some useful metrics of success of prediction when the classes are very imbalanced. It presented the values of the metrics called Precision, Recall and F-score respectively on each of the datasets that are shown in Eq. 1 to Eq. 3. Fig. 2 to Fig. 4 show the classification accuracy between proposed method and C4.5 classifier in the appearance in novel classes on the NSL-KDD, Soybean, Segment-test datasets respectively.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F - score = \frac{2 \times precision \times recall}{precision + recall} \tag{3}$$

TABLE IV: Performance of propounded method in the novel class manifestation using the NSL-KDD dataset.

Training Observations	Testing Observations	Enduring Classes	Novel Classes	Accuracy
13449	9234	1	6	85.68 %
22683	220	7	10	75.64 %
22903	2289	19	4	66.85 %

TABLE V: Performance of propounded method in the novel class manifestation using the Soybean dataset.

Training Observations	Testing Observations	Enduring Classes	Novel Classes	Accuracy
192	90	5	2	86.97 %
364	172	10	3	85.16 %
630	53	15	4	83.03 %

TABLE VI: Performance of propounded method in the novel class manifestation using the Segment-test dataset.

Training Observations	Testing Observations	Enduring Classes	Novel Classes	Accuracy
180	50	6	1	90.85 %
987	315	5	2	78.12 %
1343	417	4	3	65.57 %

TABLE VII: The standard optimum of the precision, recall and f-measure of proposed method on datasets derived from calculations.

Dataset	Precision	Recall	F-measure
The NSL-KDD	0.88	0.72	0.79
The Soybean	0.78	0.69	0.73
The Segment-test	0.82	0.95	0.88

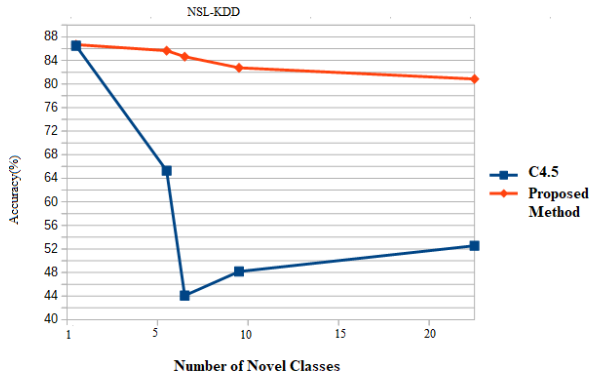


Fig. 2: Performance of propounded method and C4.5 classifier in the novel class manifestation using the NSL-KDD dataset.

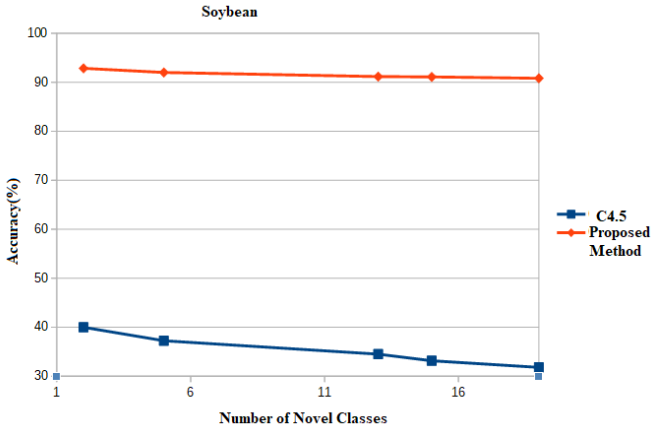


Fig. 3: Performance of propounded method and C4.5 classifier in the novel class manifestation using the Soybean dataset.

V. Conclusions & Future Work

In this paper, we successfully attempted in introducing a standard approach for mining big data in the presence of concept drift employing clustering with rule-based classifier. A major problem of dealing with big data is managing and storing it into machine's computer memory. The most popular technique is the partitioning of big data into small sub-datasets, so that each sub-data can easily be alloted to the memory of a computer. The proposed system clusters each of the sub-datasets to group the instances that are similar to each other. We then extracted classification rules from each cluster using decision tree induction technique. Merging decision trees or any other learning classifiers is a quite difficult task, whereas merging rules are quite easy owing to the fact that rules do not require any definite order of execution. We can add new rules to previously extant rules in our system without disrupting them. To find the concept drifting in data we have considered clustering technique. If any new data point does not belong to any clusters existing in our learning method, we have considered this instance as a novel class instance. Novel class instances come with new class labels in the data, generally in streaming data environment. Our conducted experiments demonstrated the performance of the proposed approach as efficacious in manner of generating better and accurate results with minimized error. We plan on extending our work in future under dynamic feature environment.

References

[1] C. P. Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on big data," *Information Sciences*, vol. 275, pp. 314–347, August 2014.

[2] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *Journal of Big Data*, vol. 2, no. 1, pp. 1–21, February 2015.

[3] X. Jin, B. W. Wah, X. Cheng, and Y. Wang, "Significance and challenges of big data research," *Big Data Research*, vol. 2, no. 2, pp. 59–64, June 2015.

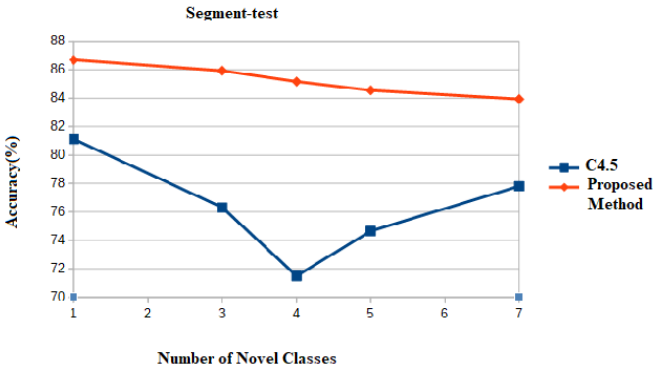


Fig. 4: Performance of propounded method and C4.5 classifier in the novel class manifestation using the Segment test dataset.

[4] D. Brzezinski and J. Stefanowski, "Prequential AUC: properties of the area under the roc curve for data streams with concept drift," *Knowledge and Information Systems*, vol. 52, no. 2, pp. 531–562, 2017.

[5] J. Gao, W. Fan, J. Han, and P. S. Yu, "A general framework for mining concept-drifting data streams with skewed distributions," in *Proceedings of the 2007 SIAM International Conference on Data Mining*. SIAM, 2007, pp. 3–14.

[6] G. I. Webb, R. Hyde, H. Cao, H. L. Nguyen, and F. Petitjean, "Characterizing concept drift," *Data Mining and Knowledge Discovery*, vol. 30, no. 4, pp. 964–994, 2016.

[7] I. liobait, M. Pechenizkiy, and J. Gama, "An overview of concept drift applications," in *Big data analysis: new algorithms for a new society*. Springer, 2016, pp. 91–114.

[8] B. Ramakrishna and S. K. M. Rao, "Concept drift detection in data stream mining: The review of contemporary literature," *Global Journal of Computer Science and Technology*, 2017.

[9] G. Song, Y. Ye, H. Zhang, X. Xu, R. Y. Lau, and F. Liub, "Dynamic clustering forest: an ensemble framework to efficiently classify textual data stream with concept drift," *Information Sciences*, vol. 357, pp. 125–143, 2016.

[10] M. Hammoodi, F. Stahl, and M. Tennant, "Towards online concept drift detection with feature selection for data stream classification," 2016.

[11] M. Lichman, "UCI machine learning repository," University of California, Irvine, School of Information and Computer Sciences, 2013. [Online]. Available: http://archive.ics.uci.edu/ml

[12] S. Liu, L. Feng, J. Wu, G. Hou, and G. Han, "Concept drift detection for data stream learning based on angle optimised global embedding and principal component analysis in sensor networks," *Computers & Electrical Engineering*, vol. 58, pp. 327–336, 2017.

[13] Y. Lu, Y. ming Cheung, and Y. Y. Tang, "Dynamic weighted majority for incremental learning of imbalanced data streams with concept drift," in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*. AAAI Press, 2017, pp. 2393–2399.

[14] A. Haque, L. Khan, M. Baron, B. Thuraisingham, and C. Aggarwal, "Efficient handling of concept drift and concept evolution over stream data," in *IEEE 32nd International Conference on Data Engineering (ICDE)*. IEEE, 2016, pp. 481–492.

[15] H. Yang and S. Fong, "Countering the concept-drift problems in big data by an incrementally optimised stream mining model," *Journal of Systems and Software*, vol. 102, pp. 158–166, 2015.

[16] A. O. Diaz, J. del Campo-Avila, G. Ramos-Jimenez, I. F. Blanco, Y. caballero Mota, A. M. Hechavarria, and R. Morales-Bueno, "Fast adapting ensemble: A new algorithm for mining data streams with concept drift," *The Scientific World Journal*, vol. 2015, 2015.

[17] B. Krawczyk and M. Wozniak, "One-class classifiers with incremental learning and forgetting for data streams with concept drift," *Soft Computing*, vol. 19, no. 12, pp. 3387–3400, 2015.

[18] H. Wang and Z. Abraham, "Concept drift detection for streaming data," in *International Joint Conference on Neural Networks*. IEEE, 2015, pp. 1–9.

[19] B. Mirza, Z. Lin, and N. Liu, "Ensemble of subset online sequential extreme learning machine for class imbalance and concept drift," *Neurocomputing*, vol. 149, pp. 316–329, 2015.

[20] K. Kanoun and M. van der Schaar, "Big-data streaming applications scheduling with online learning and concept drift detection," in *Proceedings of the 2015 Design, Automation & Test in Europe Conference & Exhibition*. EDA Consortium, 2015, pp. 1547–1550.

[21] L. Rutkowski, M. Jaworski, L. Pietruczuk, and P. Duda, "A new method for data stream mining based on the misclassification error," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 5, pp. 1048–1059, 2015.

[22] D. M. Farid, L. Zhang, A. Hossain, C. M. Rahman, R. Strachan, G. Sexton, and K. Dahal, "An adaptive ensemble classifier for mining concept drifting data streams," *Expert Systems with Applications*, vol. 40, no. 15, pp. 5895–5906, 2013.

[23] D. M. Farid, A. Nowé, and B. Manderick, "A feature grouping method for ensemble clustering of high-dimensional genomic big data," in *Future Technologies Conference*, San Francisco, United States, December 2016, pp. 260–268.

[24] J. R. Quinlan, "Induction of decision tree," *Machine Learning*, vol. 1, no. 1, pp. 81–106, August 1986.

[25] C. Apté and S. Weiss, "Data mining with decision trees and decision rules," *Future Generation Computer Systems*, vol. 13, no. 2-3, pp. 197–210, November 1997.

[26] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Chapman and Hall/CRC, 1984.

[27] J. R. Quinlan, *C4.5: Programs for Machine Learning*. Morgan Kaufmann, 1993.

[28] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: An update," *ACM SIGKDD Explorations*, vol. 11, no. 1, pp. 10–18, June 2009.

Efficient Implementation of Dynamic Array System Using MapReduce Framework

K. M. Azharul Hasan, Mehnuma Tabassum Omar, Rahat Haider, Sk. Md. Masudul Ahsan
Department of Computer Science and Engineering
Khulna University of Engineering & Technology (KUET)
Khulna, Bangladesh
Email: azhasan@gmail.com, misty2409@gmail.com, rahat.tech@gmail.com, masudul.ahsan@gmail.com

Abstract— **In recent years managing Big Data is a big challenge. Big data has huge volume, variety of format and its volume increases at a very high velocity. Traditional data structures fail to handle these types of data. Moreover Big Data storage scheme is very expensive. So some efficient scheme is needed. In this paper we show the application of dynamic Extendible Array for Big data storage. It is an efficient scheme which has better performance over other approaches. We used Map Reduce concept to distribute the data in heterogeneous environment. The data that belongs to different dimensions are distributed to different machines to do the operations efficiently. The basic operations including insertion, deletion, update and retrieval of various types namely point key query, single key query and range key query are performed. Moreover the dynamic extension of the structure without reorganizing the existing data is performed. Experimental results are well studied. We used cheap commodity machines for our implementation.**

Keywords—MapReduce, Extendible Array, Dynamic Array storage, Big Data Storage, Future data maintanance.

I. INTRODUCTION

The large volume with diverse dimensionality attribute of arrays make it popular in various researches, like medical imaging, geographic information system, environmental and astronomical surveillances, or high precision prototypes of physical consequence [1][2]. These applications force an array to demand a scalable system for array operations. But the Traditional Multidimensional Array (TMA) does not provide dynamic scalability as it requires reorganization of the previous data in case of a new data set entry. Hence, the TMA structure is unable to fulfill the dynamic data velocity of Big Data. Much IT investments are going towards managing, maintaining and analyzing big data. Sometimes big data is represented by three Vs. - Velocity, Variety and Volume [3]. Big data comes with some big problems. One of them is big data storage. Huge data Volume which is increasing with high Velocity (termed as future data) with Variety in format can't be handled using tradition storage systems. Sometimes big data is unstructured. So storing huge amount of unstructured data in relational database is not a very good idea. Moreover analysis of huge amount of data is time consuming in relational database systems. So some efficient and dynamic data storage techniques should be adopted which will speed-up storing and analyzing of big data especially future data.

In this paper, we propose a big data storage system using commodity machines in heterogeneous environment. The scheme employs a distributed storage model using extendible multidimensional arrays[2][9][15]. For storing high dimension data arrays are used extensively[3][11]. This paper presents a dynamic array storage namely Extendible Karnough Array[2][12][16] based map reduce storage technique. The EKA is a dynamic array system that can store the future data. In this paper we proposed map reduce algorithm for fast retrieval of array elements in distributed environment for future data maintenance. Map Reduce is now mostly used for distributed analysis system. Map reduce algorithms are used extensively for example, Google implements the map reduce paradigm internally for the last decade and jobs of one hundred thousand are executed per day on Google's clusters which is nearly twenty petabytes of data per day. Map Reduce offers scalability, distributed analysis and reliability which are crying need for an efficient big data storage system [3][17].

II. RELATED WORKS

Some of the dynamic array models are summarized in this section. [16] employs an extendible array that extends the array boundary dynamically. It on ly treats the history tables in an efficient way. The extendible array model presented in [21] introduces axis vectors that stores necessary information like the auxiliary tables.[4][5] mentions the Extended Karnaugh Map Representation (EKMR) model that signifies an n dimensional data by a set of two dimensional arrays. But EKMR is also not dynamically extendible. The Extendible Array (EA) model [6]-[12] can meet the expended data velocity by solving the limitation of a TMA using dynamic memory allocation during run time as form of subarrays. [9] represents an EA based model which can convert the n dimensions of an array by only 2 dimension. The subarrays of EA based models are n-1 dimensional and hence it can utilize better storage compared to TMA. However, the EA based models cannot meet the expected huge storage requirements like Big Data applications as they cause address space overflow. [10] represents an EA model based on [9] which can delay the address space overflow and can meet the expected storage requirements of large volume applications efficiently compared to other EA models. [2]. [11][12] represents another EA solutions to address space overflow based on EKMR. None of

the above model used either the idea map reduce retrieval or distributed array storage. We implemented distributed array storage and applied ma preduce technique for retrieval. The map reduce architecture is a way to originate and activate large data sets that enlarges the scalability of parallel applications efficiently which attains high performance [13]. [14] improves [13] named Map-Reduce-Merge which introduce the Map-Reduce with a Merge phase that can competently combines the partitioned and sorted results of the previous map and reduce units. [15] employs Map-Reduce on Big Data applications. In this paper, we have used the Map Reduce architecture to utilize the data sets of an EKA based model. The introduction of map reduce model in dynamic arrays is employed in this paper to speed up the data retrieval and increase the storage capacity.

III. THE EXTENDIBLE KARNOUGH ARRAY (EKA)

The EKA is illustrated in this section. The details can be found in [2][12]. The idea of date distribution is possible by EKA because the dimensions and subarrays introduced in EKA are completely independent to each other. Hence the subarrays are send to different machines and merged by another machine for cumulative results.

A. Structure of an EKA

The EKA is a multidimensional Extendible Array model that utilizes the concepts of Karnaugh Map (K-map). The K-map is awell-knowndepiction employed inBoolean expression minimization typically assisted by mapping values for eachpotential combinations.Fig. 1 (a) depicts a 4 variable K-map representation of a Boolean function (2^4 combinations). The row is denoted bythe pair (w, x) and the column is denoted bythe pair variables (y, z). The row and the column indicate the potential combinations of a Boolean function in a form of two dimensional array.

Adjacent Dimension. The row pair and column pair of a K-map are re-expressed as row dimension and column dimension respectively of an EKA. Here, the indices of the row dimension are adjacent to each other and the indices of the column dimension are adjacent to each other. The adjacent dimensions of row is denoted as $adj(z) = y$ or $adj(y) = z$ and adjacent dimensions of column is denoted as $adj(w) = x$ or $adj(x) = w$ and.The EKA representation of Fig. 1(a) is shown in Fig. 1(b).

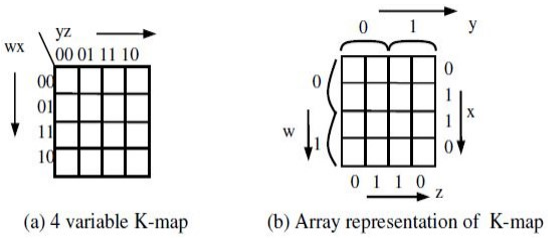
Auxiliary Table. The EKA model requires some auxiliary tables for managing scalability of its structure and also for managing faster data retrieval from the structure as follows:

- *History Table* which is required monitoring extension information.
- *Co-efficient Table* which is required for storing the co-efficients of addressing function. The addressing function can be defined as follows:

$$f(x_n, x_{n-1}, x_{n-2},\ldots,x_2, x_1) = l_1l_2\ldots l_{n-1}x_n + l_1l_2l_3\ldots l_{n-2}x_{n-1} + \ldots + l_1x_2 + x_1. \quad (1)$$

where, $< l_1l_2\ldots l_{n-1}, l_1l_2l_3\ldots l_{n-2}, \ldots, l_1>$ are theco-efficients of addressing function and l_i indicates the length of dimension d_i.

- *Address Table* which is required for monitoring the extended subarays of an EKA.As the subarrays of EKA are divided into segments, hence this table will store the first address of each segments by storing first address.



(a) 4 variable K-map (b) Array representation of K-map

Figure 1: Realization of an EKA Using K-map

B. Operations on EKA

Extending Dimension. Consider the array in Fig. 2(a), the dimensions are d_1, d_2, d_3 and d_4 and thesize of the array is [l_1, l_2, l_3, l_4] andsubscripts varies from 0 to l_i-1. In the current example $l_i = 2$. The dimension (d_1,d_3) and (d_2,d_4) are adjacent dimensions respectively. The extension of $d_1$that can be done logically isshown in Fig. 2(b). The extension is done by allocating subarrays and is determined by [l_2, l_3, l_4] (i.e. 3 other dimensions). The number of segments is thelength of the adjacent dimension, $adj(d_1) = d_3$; In this case it is $l_3 = 2$. The size ofeach segmented subarray extended along dimension d_1 is determined by [l_2, l_4]. Afterextending along dimension d_1, the length of that dimension is incremented by 1. Foreach extension the corresponding auxiliary tables are maintained accordingly. Fig.2(c), 2(d) and 2(e) shows the extension realization along dimension d_2, d_3 and $d_4$respectively.

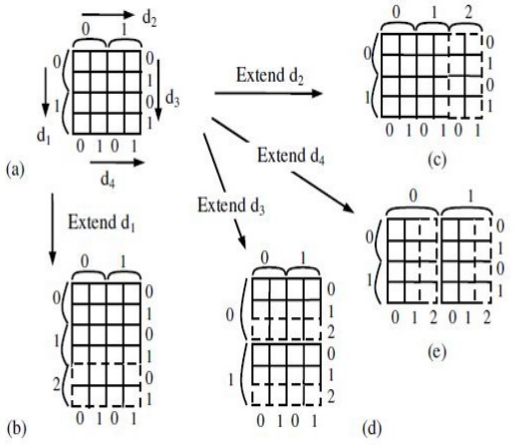


Figure 2: Logical extension of a 4-dimensional EKA.

The EKA(4) shown in Fig.3 is taken as an example in this section. This EKA(4) is going to be extended for the corresponding 4 dimensions. The initial situation is shown in Figure 3(a). The value of history counter is initially 0, the address table has the first address and coefficient table is initialized to 1. During the extension of $d_1$and d_3 dimension size of the segment is $l_2 \times l_4$ which is a two dimensional array and so coefficient vector is one dimensional. Hence for our example

we use l_2 as coefficient vector for d_1 and d_3 dimensions. Similarly $l_3$is used as coefficient for $d_2$and d_4 dimensions.

Retrieval. Let the value to be retrieved is indicated by the subscript $\langle x_1, x_2, x_3, x_4\rangle$. The maximum history value among the subscripts h_{max}= max($H_{d1}[x_1]$, $H_{d2}[x_2]$, $H_{d3}[x_3]$, $H_{d4}[x_4]$) and the dimension (say d_{max}) that corresponds to history value h_{max} is determined. h_{max} is the subarray that contains the desired element. This is because;the history values are linear numbers, therefore subarray having maximum history value was constructed at last. Hence the desired element remains in the segment that makes the subarray having history value h_{max}. Now the first address and offset from the first address is to be found out. The adjacent dimension $adj(d_{max}) = d_{adj}$(say) and its subscript x_{adj} is found. Now the first address of the segment is

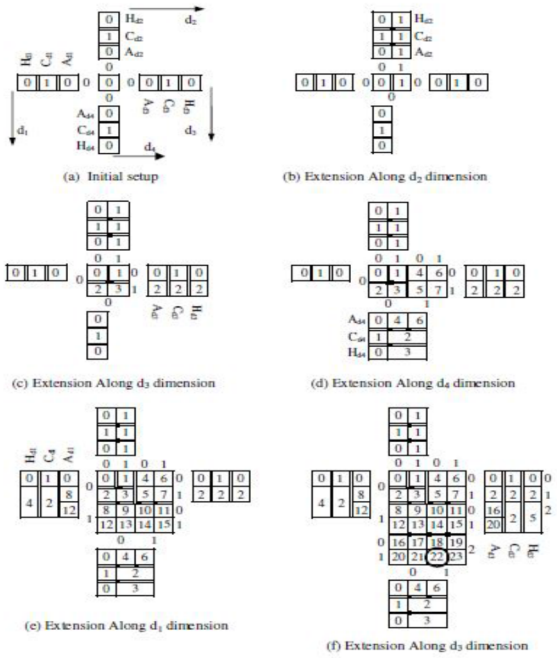


Figure 3: An Extension Realization of EKA(4)

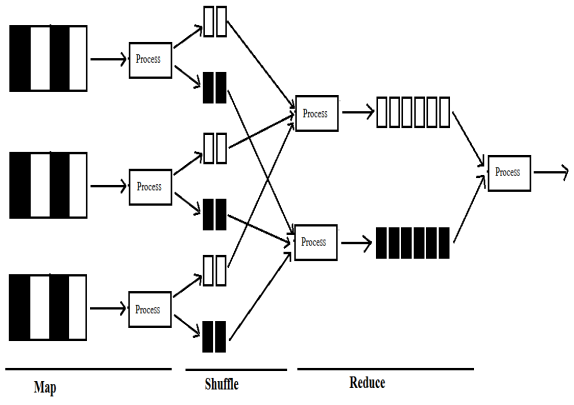


Figure 4: The MapReduce implementation model

found from$A_{dmax}[x_{max}][x_{adj}]$. The *offset* from the first address is computed using the addressing function (see Eq. 1); the coefficient vectors are stored in C_{dmax}. Then adding the *offset* with the first address, the desired array cell (x_1, x_2, x_3, x_4) is found.Let four subscripts ⟨1, 0, 2, 1⟩ for dimension d_1, d_2, d_3, and d_4 is given (see Figure 3.3(f)). Here h_{max}= max($H_{d1}[1]$, $H_{d2}[0]$, $H_{d3}[2]$, $H_{d4}[1]$)= max(4, 0, 5, 3) =5, and dimension corresponding to h_{max}i.e. d_{max} = d_3 whose subscript x_{max}= 2 and $adj(d_{max}) = adj(d_3) = d_1 = d_{adj}$*and* $x_{adj} = 1$. So the *firstAddress*=$A_{d3}[2][1]$ = 20, and offset is calculated using the coefficient vector stored in coefficient table C_{d3} which is 2. Here, *offset*=$C_{d3}[2]*x_4 + x_2 = 2*1 + 0 = 2$. Finally adding the offset with the first address the desired location 20 + 2 = 22 is found and circled in Fig.3.

IV. The Map-Reduce Model for Distributed Array Storage

The mapping phase is implemented by invocatingmachines automatically partitioning the input data into a number of splits. These splits are processed according to parallel scheme by a number of machines. The reduce phase merges the distributed partitions into the intermediate key using a partitioning function namely hash function. Fig.4 shows the overall flow of a MapReduce operation in our implementation [13].

We have used EKA[2] as basic storage structure. In its initial format has 4-dimensions. We will use this basic format and extend along the each dimension. Each of the dimensions can be used for storing specific type of data. For example dimension one and two may store string type data where dimension three and four may store integer type of data.

A. Distribution of Dimensions

The EKA has initially four dimensions. It has also internal mechanism to store and retrieve data. This internal mechanism and storage of dimensional data introduce extra overhead to single machine. This burden of single machine is distributed among many machines. In its simple form each dimension is distributed with single node machine. So for four dimensions we need four single node machines. Where internal mechanism for storing and retrieving data is maintained by the master machine. Fig.5 shows a 4D EKA. Fig. 6 shows the distribution scheme for master and slave machines.

Master Machine: Master machine performs all the initializations for single node machines. The main task of master machine is to maintained internal structure and mechanism of EKA which is complex. So master machine is responsible for storing EKA's History table, Co-efficient table and Address table. Operations are given through this machine.

The machine modify or access these auxiliary tables with respect to desire operations

Slave Machine: Slave machines are single node machines simply store the dimension data. They are responsible for storing data in their corresponding format. In response to the

master machine the single node machines simply allocate memory and retrieve data according to the type of the query.

V. Implementing Map Reduce Concept

In this section EKA we describe the implementation in heterogeneous environment. Heterogeneous environment in the sense that the single node machine may have different configurations, and most importantly they store data of different format or type. This heterogeneity is maintained and merged by implementing Map-Reduce concepts in extending dimension and data retrieving.

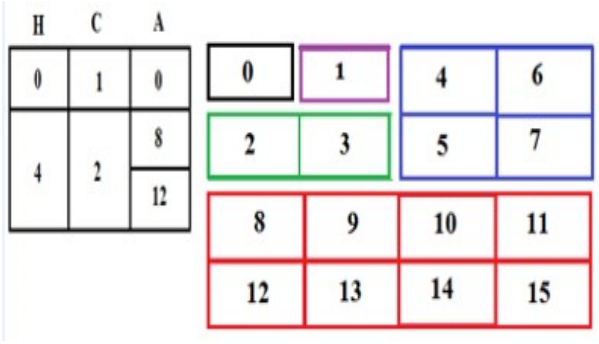


Figure 5: Main structure of EKA.

Map Reduce Concepts in Extending Dimensions

***Map-Phase*:** Extending along any dimension is done by first determining the specific node machine and then delivering the task of extension to the corresponding node.

***Reduce-Phase*:**The corresponding node performs all the tasks for extension such as memory allocation, data storing and sending feedback message to master node. Then the master node modify the auxiliary tables with respect to the feedback from the corresponding single node.

Map Reduce Concepts in Retrieving Data

Retrieving data may include communicating with all the single node machines. Master node first takes the instructions and then separates the indices that corresponds to the data in specific dimension that resides in specific single node machine. Then data is retrieved.

Map Phase: Gets instruction, separates indices and send requests to the specific single node with indices that belongs to the corresponding dimension.

Reduce Phase: Each single node machine collects the corresponding data for given indices and sends feedback to master node with retrieved data. Thenthe master node merges and show all the retrieved data. Fig. 7 shows a schematic diagram for map reduce technique for retrieval of data.

VI. Experimental Results

We measured the required time to retrieve data for different extension length along each dimension. For retrieve we accessed the middle square of the whole data space. We implemented a range query as shown in black part of the Fig. 8. For a multidimensional we implemented the different part in each dimension. For example if the length of dimension one is 20 then 5 to 15 indices will be selected for dimension one. This will help us to get average result.The experiment is repeated up to each dimension length is 60. The result is compared with single machine implementation of EKA (4).

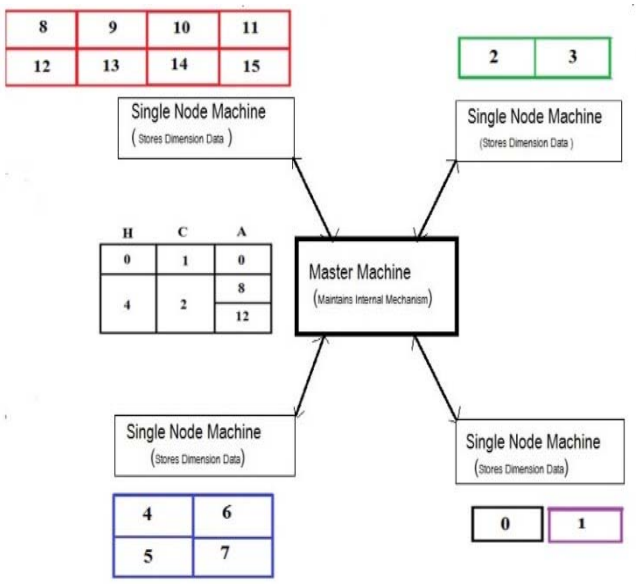


Figure 6: Distribution of dimension

We measured the required time to retrieve data for different extension length along each dimension. For retrieve we accessed the middle square of the whole data space. We implemented a range query as shown in black part of the Fig. 8. For example if the length of dimension one is 20 then 5 to 15 indices will be selected for dimension one. This will help us to get average result.The experiment is repeated up to each dimension length is 60. The result is compared with single machine implementation of EKA. The performance is shown in Fig.9. When we use multiple machines the length of dimension of the EKAis increased from 75 to 110 this means the volume of the array increases $(75)^4$ to $(110)^4$. This is huge improvement for array storage. If the average data size is 4 unit and M is the size needed to maintain the structure then

In single machine we get capacity, $c1 = (75)^4 \times 4 + M_1$

In multi machine we get capacity, $c2 = (110)^4 \times 4 + M_2$

Increased capacity = (c1/c2), which is approximately 4.63.So using multiple machine we able to increase the capacity 4.63 times larger than single machine.

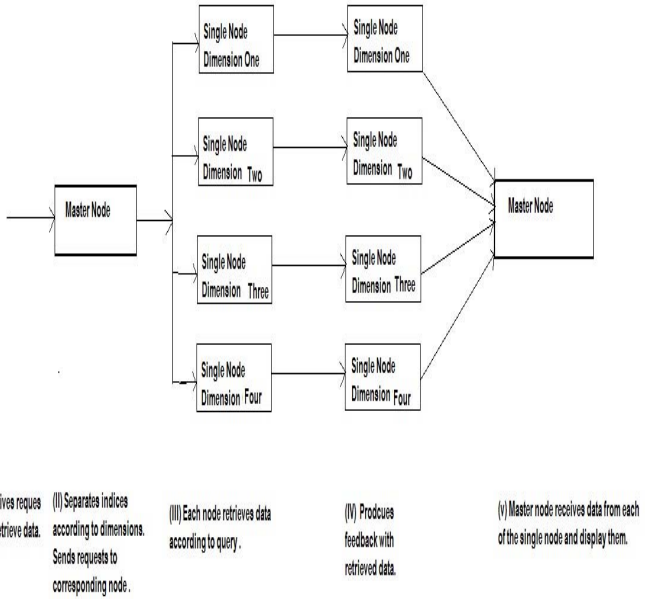


Figure 7: Map-Reduce in retrieving data

From Fig. 10 we can also see the retrieval performance is also improved for map reduce technique. This is because 4 processor will be more powerful than the single one. Figure 9 shows the improvement i.e speed up of the two schemes. The speed up is not so good because we used heterogeneous environment where the machines are connected through optical fiber and core switch for networking among them.

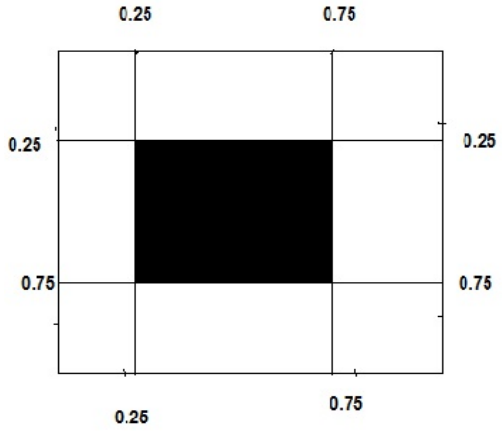


Figure 8: Map-Reduce retrieving in EKA

VII. CONCLUSION

In this paper we analyze and implement new array storage using EKA and Map Reduce concepts. Distributed technique is compared with the single machine technique. As in distributed technique many computations are done in parallel, from the analysis we can simply show that our technique is better than single machine concept. We all the dimension data and structure of EKA in main memory. We can now recommend our scheme to use it for big data storage where multiple machines are available and huge amount of data need to be stored. The scheme will be efficient in for big data storage in actual situation because the subarrays are independent to each other and makes easy to implement in parallel environment. We believe the scheme can be more effective for actual parallel environment for huge data d management especially for future data.

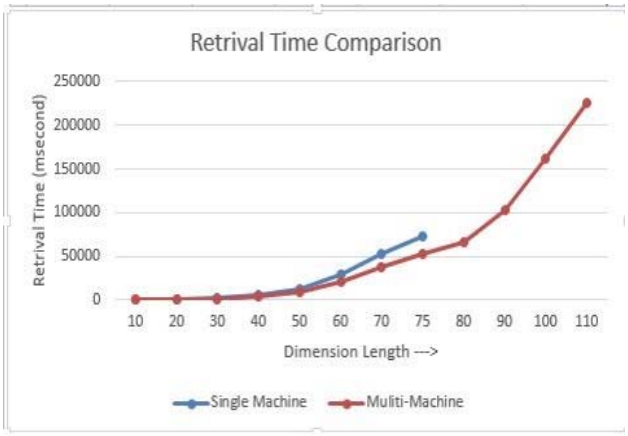


Figure 9: Map-Reduce retrieval performance data

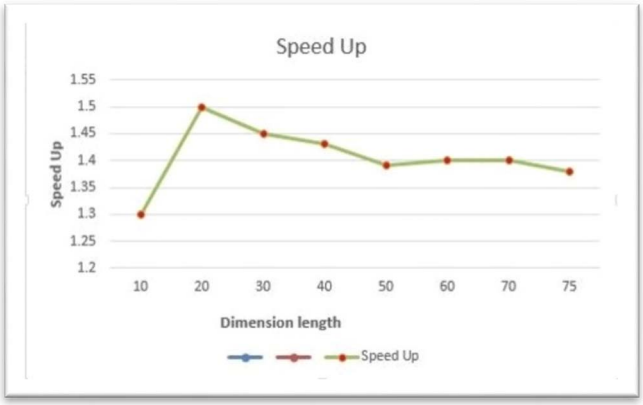


Figure 10: Speed up for Map-Reduce implementation.

REFERENCES

[1] Fay Chang, Jeffrey Dean. "Bigtable: A Distributed Storage System for Structured Data". ACM Transactions on Computer Systems (TOCS) TOCS Homepage archive Volume 26 Issue 2. Article No. 4. June 2008.

[2] Sk. Md. Masudul Ahsan, K. M. Azharul Hasan. "An Implementation SchemeforMultidimensional Extendable Array Operations and Its Evaluation" ICIEIS 2011, Part III, Springer-Verlag Berlin Heidelberg, pp 136-150, 2011.

[3] Jeffrey Dean, Sanjay Ghemawat. "MapReduce: a flexible data processing tool".Magazine.Communications of the ACM - Amir Pnueli: Ahead of His Time CACM Homepage archive.Volume 53 Issue 1. pp 72-77. January 2010.

[4] L. Chun, C. C. Yeh, and S. L. Jen, "Efficient representation scheme for multidimensional array operations," IEEE Computer, 51(3), pp. 327−345, 2002.

[5] Y. L. Chun, C. C. Yeh, and S. L. Jen, "Efficient data parallel algorithms for multidimensional array operations based on the EKMR scheme for distributed memory ulticomputer," IEEE Parallel and Distributed Systems, 14(7), pp. 625−639, 2003.

[6] K M Azharul Hasan, Tatsuo Tsuji, Ken Higuchi, "A parallelimplementation scheme of relational tables based on

multidimensional extendible array", International Journal of Data warehousing and Mining, vol. 2(4), pp. 66-85, 2006.

[7] D. Rotem, E. J. Otoo, and S. Seshadri, "Optimal chunking of large multidimensional arrays for data warehousing", In Proceedings of the ACM tenth International Workshop on Data Warehousing and OLAP, pp. 25-32,November 2007.

[8] K. M. Azharul Hasan, Masayuki Kuroda, Naoki Azuma, Tatsuo Tsuji, Ken Higuchi: "An extendible array based implementation of relational tables for multi dimensional databases". DaWaK, pp. 233-242, 2005.

[9] MehnumaTabassum Omar, K. M. Azharul Hasan (2016) A scalable storage system for structured data based on higher order index array. The 3rd IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT) 2016.

[10] MehnumaTabassum Omar, K. M. Azharul Hasan (2016) Towards an Efficient Maintenance of Address Space Overflow for Array Based Storage System. 17th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT), 2016.

[11] Sk. Md. Masudul Ahsan, K. M. Azharul Hasan: Segment Oriented Compression Scheme for MOLAP Based on Extendible Multidimensional Arrays. CIT 23(2): 111-121 (2015).

[12] K M Azharul Hasan, Kamrul Islam,Mojahidul Islam, Tatsuo Tsuji, An Extendible Data Structure for Handling Large Multidimensional Data Sets, In: Proceedings of the 12th International Conference on Computer and Information Technology (ICCIT 2009), Dhaka, Bangladesh, pp. 669-674, 2009,

[13] Jeffrey Dean and Sanjay Ghemawat. "MapReduce: Simplified Data Processing on Large Clusters". Magazine. Communications of the ACM - 50th anniversary issue: 1958 - 2008 CACM Homepage archive.Volume 51 Issue 1.pp 107-113.January 2008.

[14] Hung-chih Yang, Sunnyvale." Map-reduce-merge: simplified relational data processing on large clusters". Proceeding. SIGMOD '07 Proceedings of the 2007 ACM SIGMOD international conference on Management of data. pp 1029-1040.

[15] Kyuseok Shim. "MapReduce algorithms for big data analysis". Journal.Proceedings of the VLDB Endowment VLDB Endowment Hompagearchive. Volume 5. Issue12.pp 2016-2017 .August 2012.

[16] Otoo, E.J., and Rotem, D.Efficient Storage Allocation of Large-Scale Extendible Multi-dimensional Scientific Datasets. *In proceedings of the 18th International Conference on Scientific and Statistical Database Management*, 179-183, 2006.

[17] K. M. Azharul Hasan, Tatsuo Tsuji, Ken Higuchi, " An Efficient Implementation for MOLAP Basic Data Structure and Its Evaluation, In : Proceedings of the 12th International Conference on Database Systems for Advanced Applications, {DASFAA}, Bangkok, Thailand, April 9-12, PP.288-299, 2007.

Develop a Model to Secure and Optimize Distributed File Systems for ISP Log Management

Mohammad Shazzad Hossain
Department of ICT
Bangladesh University of Professionals
Mirpur, Bangladesh
shazzad.hossain@southeastbank.com.bd

Pratik Kumar Bhowmik
Department of ICT
Bangladesh University of Professionals
Mirpur, Bangladesh
pratik@southeastbank.com.bd

Md. Abidur Rahman
Department of ICT
Bangladesh University of Professionals
Mirpur, Bangladesh
abidur@southeastbank.com.bd

Mahbub Uddin
Department of ICT
Bangladesh University of Professionals
Mirpur, Bangladesh
mahbub.uddin@synesisit.com.bd

Mohammad Abu Yousuf
Institute of Information Technology
Jahangirnagar University
Savar, Bangladesh
yousuf@juniv.edu

Kazi Abu Taher
Department of ICT
Bangladesh University of Professionals
Mirpur, Bangladesh
ktaher@bup.edu.bd

***Abstract*— Distributed file system (DFS) is a transformation of traditional file systems to perform file read, write and execution of petabyte or larger-sized datasets with high-velocity and different structures. In order to process these large amounts of data in an inexpensive and efficient way, Hadoop Distributed File System (HDFS) is used and designed to scale up from a single server to hundreds of servers, with a very high degree of fault tolerance. Secured and optimized distributed file system for ISP user log management is a centralized log analytics platform that supports collaboration in a distributed environment. DFS for ISP log provides a uniform name space for all ISP along with location independence and location transparency. Any user's internet access log can be accessed and analyzed using apache sparks data brick file system on stored log in DFS. In this paper, at first, limitation of popular DFS like NFS and CIFS is explored. ISP log management, DFS file architecture and security is considered for designing proposed ISP log management system. Finally, simulation on proposed distributed file system and query processing is performed to obtain the desired result related to optimization and security of the proposed distributed log analysis system.**

Keywords—Hadoop, DFS Security, NFS, ISP log, Optimize

I. INTRODUCTION

From a security point of view, the purpose of a log is to act as an alert or red flag to detect suspicious activity. Reviewing access logs could help identify malicious activities on any system. The primary objectives for log management implementations are concerns about security, system and network operations and regulatory compliance. Given the large amount of log data generated by systems, it is impractical to preserve and review all of these logs manually each day. Effectively analyzing large volumes of diverse logs can pose many challenges, such as volume, normalization, velocity and veracity. Commercial organizations may procure complete solutions or build their own log management tools, assembling the features from various open-source components, or acquire systems from commercial vendors. Log management is a complicated system and organizations often make mistakes while approaching it.

A distributed file system (DFS) is a file system with data stored on multiple servers [1]. The data is accessed and processed as if it was stored on the local client machine. The DFS makes it convenient to store and process data files of any relational databases or share files among users on a network in a controlled and authorized way [2]. The server allows the client or users to share files and store data just like they are processing the information locally.

An NFS server can mount directories that are exported by other NFS servers [3]. However, it cannot export share to its clients. NFS supports location transparency but not location independence. This is due to its decision not to enforce a common name space. This makes it harder to share files. CIFS cannot really be described as a legitimate distributed system. It merely provides a way to access resources over a network. It does not provide either location transparency or location independence. AFS does not use unix file semantics. AFS only supports directory permissions which may confuse those who are used to the standard unix chmod permission sets. The need to authenticate everything and everyone may sometimes cause problems for people who have no idea how to handle authentication. XFS is a high-performance 64-bit journaling file system created by Silicon Graphics and ported to the Linux kernel. XFS is supported by most Linux distributions i.e., Debian and REDHAT. XFS supports a maximum file system size of 500 TB and a minimum file size of 16 TB. An XFS file system can be shrunk only by backing up, re-installing, and restoring, e.g. using xfsdump and xfsrestore. Samba uses netbios which is not the world's most efficient network protocol.

De-centralized ISP log management and lack of standard for ISP log format and retention has motivated us to design and implement a centralized log management system where hadoop distributed file system will be implemented due to limitation of other popular distributed file systems namely NFS and CIFS. Apache spark performs faster than map reduce on executing no-sql query and ranger security suite provides best access control across distributed file system. Considering the volume and necessity of ISP user's log, the research issues of the paper are set as follows:

- Review the features of all DFS. Choose highly reliable, resilient DFS for big data.
- Review and analysis of log mandate, format and propose a common log format.
- Security of proposed DFS. Ensure confidentiality of stored log, integrity and availability of proposed DFS.
- Implementation of proposed DFS with 3 nodes and optimization of DFS using open source software resources.

- Perform log analysis and simulation to obtain desired results like best performance of read, write latency and security features. Optimization of DFS using open source software resources.

II. LITERATURE REVIEW

Deepa et al. state that the components of HDFS file system, Big Data and processing of this terabytes to multiple petabytes which will help to manage and process large volume of data [4]. And also addresses the fundamental of HDFS file system with architecture and challenges. It is also mentioned how storing less number of larger files rather than huge number of smaller files. V. Sajwan et al. state about HDFS architecture, name node and Data Node configuration [5]. HDFS is designed and developed to scale up from a one server to thousands of machines device with extremely high degree of ability.

BTRC has described about ISP license and internet user statistics. This is about the total number of internet subscribers which has already reached 92.466 Million at the end of October, 2018 [6]. This survey provides a clear view of necessity of centralized log server for keeping the track of internet usage. CoreStack addresses the initiative to improve transparency of ISPs in India [7]. This paper will describe how to preserve access log for all users and share log files with central regulatory body using log mandate.

Internet has significant contributions to our life. Most people are connected to internet for 24 hours using BWA and WIFI device which in turn connects to local ISP and WIMAX/3G/LTE/4G internet service provider. Local ISP's are maintaining their own log servers where customers' internet session is logged using different format [8]. The Indian ISPs have to comply with the directions issued by the Department of Telecommunications (DoT). In its initiative to improve transparency, the DoT had mandated ISPs to maintain logs of Internet Protocol Detail Record (IPDR) and Network Address Translation (NAT).The directive also instructed storage of log data using format for one year by every ISP owner itself.

HDFS is a distributed file system designed to run on x86 processor based server. HDFS is highly fault-tolerant and is designed to be deployed on low-cost hardware [9]. HDFS provides high read, write throughput access to application data and is suitable for applications that have large data sets. Features of HDFS includes hardware failure, streaming data access using POSIX semantics [10], large data set, data replication, the communication protocol layered on top of the TCP/IP [11] and remote procedure call (RPC) [12].

Apache Sentry, an open source project by Cloudera is an authorization module for Hadoop that offers the granular, role-based authorization required to provide precise levels of access to the right users and applications [13]. It supports for role-based authorization, fine-grained authorization and multi-tenant administration [14].

Apache Knox is a security gateway that provides a single point of authentication and access gateway for various Hadoop services in a cluster. It provides a perimeter security solution for HDFS. It supports various authentication and token verification scenarios. Knox manages security across multiple clusters and versions of Hadoop. It also provides SSO solutions, and allows integrating other identity management solutions such as LDAP, Active Directory (AD), and SAML based SSO and other SSO systems [15].

Project Rhino provides an integrated complete data security solution to the Hadoop ecosystem [16]. It provides a token based and single sign-on solution. It offers Hadoop crypto codec framework and implementation to provide block level encryption for the data stored in Hadoop. It supports key distribution and management so that Map-Reduce can decrypt data block and execute the program as per requirement. It also enhances the security of HBase by offering cell level authentication and transparent encryption for table stored in HDFS. It supports audit logging framework for easy audit trails. Spark is a no-sql query processing tool developed by Apache to process programs written in a high-level concise language and distributes the execution of its tasks onto DFS.

It achieves this through an API such as Data Frames and Datasets built atop Resilient Distributed Datasets (RDDs) – a distributed dataset abstraction that performs calculations on large clusters in a fault-tolerant manner. Spark's architecture improves RDD's performance significantly:

- Spark allows users to take advantage of memory-centric computing architectures by persisting DataFrames, Datasets, and RDDs in-memory, enabling fast iterative processing use case such as interactive querying or machine learning.
- Spark's high level DataFrames and Datasets APIs also enable further intelligent optimization of user programs.
- Project Tungsten and Catalyst Optimizer as part of the Spark SQL engine significantly boosts spark's execution speed in many cases by 5-10 times.

III. PROPOSED SYSTEM ARCHITECTURE

Addressing the problem of existing ISP log management process and studying the relevant work of other countries a system architecture is proposed in this paper. It contains 1 master node and 3 slave nodes for DFS to interconnect ISP log servers. Proposed DFS, log analysis process, access control policy, security features and data protection process is discussed. A common log format for all ISP is also shown in this section.

A. Proposed DFS Model

HDFS exposes a file system namespace and allows user data to be stored in files. Internally, a file is split into one or more blocks and these blocks are stored in a set of Data Nodes. The Master Node executes file system namespace operations like reading, closing and modifying files and directories. It also determines the mapping of blocks to Slave Nodes. The Slave Nodes are responsible for serving read and write requests from the file system's clients. The Slave Nodes also perform block creation, deletion and replication upon instruction from the Master Node. The details of nodes and functionalities are described in Table I.

The Name node makes all decisions regarding replication of blocks by receiving a heartbeat and a block report from each of the Data Nodes in the cluster. Reply message of a heartbeat implies that the Data Node is functioning properly. A block report contains a list of all block's status on a Data Node.

TABLE I. ISP LOG ANALYTICS NAME NODE AND DATA NODE INFORMATION

SL	*Node Name*	*Disk Space*	*RAM*	*IP Address*	*Role*
1	Name Node	50 GB	12 GB	103.103.35.137	Name Node
2	Data Node 1	50 GB	12 GB	103.103.35.137	Data Node, Spark Server
3	Data Node 2	50 GB	7 GB	103.103.35.138	Data Node
4	Data Node 3	50 GB	7 GB	103.103.35.139	Data Node, Apache Ranger

A Data Node stores data in the distributed file systems. A functional file system has more than one Data Node, with data replicated across them. On startup, a Data Node connects to the Name node and spinning until that service comes up. It then responds to requests from the Name Node for file system operations. Our proposed architecture will include all ISP's as Data Node. Here Data Nodes are represented by debian, node1 and node2 respectively.

B. Log analysis over Proposed DFS

We've deployed spark at master node 'debian' to process query over ISP internet access log in CSV format. Apache Spark has had two core contexts that are available to the user. The sparkContext made available as sc and the SQLContext made available as sqlContext, these contexts make a variety of functions and information available to the user. The sqlContext makes a lot of DataFrame functionality available while the sparkContext focuses more on the Apache Spark engine itself. In our proposed deployment, Apache Spark 2 contains just one context - the SparkSession.

Databricks is a platform for running Apache Spark which provides a host of features to help its users be more productive with Spark. It's a point and click platform for those that prefer a user interface like data scientists or data analysts. However, this UI is accompanied by a sophisticated API for those that want to automate aspects of their data workloads with automated jobs. To meet the needs of enterprises, Databricks also includes features such as role-based access control and other intelligent optimizations that not only improve usability for users but also reduce costs and complexity for administrators. In Spark environment, a DataFrame is a distributed collection of data organized into named columns. It is equivalent to a table in a relational database or a data frame but with richer optimizations under the hood. Data Frames can be built from a wide range of sources such as: structured data files, tables in Hive, external databases, or existing RDDs.

C. Access Control of Proposed DFS

Apache Ranger is a security solution to enable, monitor and manage comprehensive data security across the HDFS file system. The vision with Ranger is to provide centralized and comprehensive security across the Apache Hadoop ecosystem using attribute based access control (ABAC). With the advent of Apache YARN, the Hadoop platform can now support true data lake architecture. Organizations can run multiple workloads, in a multi-host environment. Data protection within HDFS needs to evolve to support multiple use cases for data access, while also providing a framework for central administration of security policies and monitoring of user access.

Policies are set centrally and enforced by plugins. Plugins are simple libraries that run within data processing application and enforce policies enforces set in the policy manager. Apache Ranger goes beyond access control, with the ability to dynamically masking data when data is accessed in Hive. Centralized policy management and distributed enforcement around various data processing and storage engines using ranger is illustrated in Fig. 1.

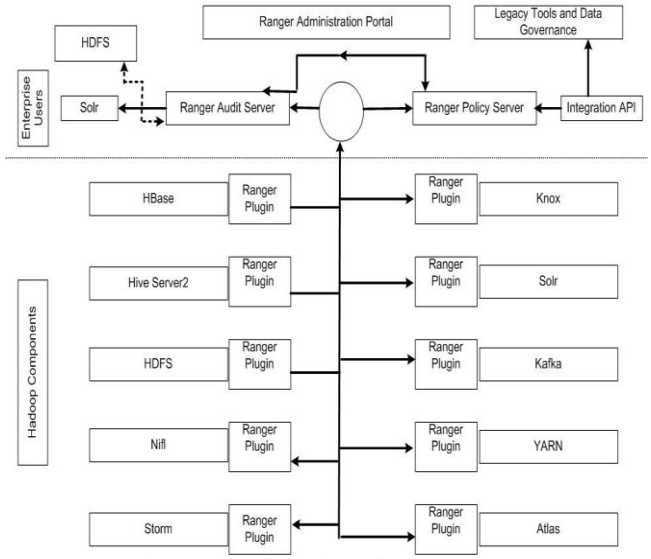


Fig. 1: Proposed security policy management

D. Security Features of Proposed DFS

Initial Hadoop distribution, Data Nodes do not enforce control. Due to this reason, the malicious user can read arbitrary Data Node blocks from Data Nodes and bypass the access control restrictions or write the garbage data in Data Nodes. Therefore, any user can submit a job to Job Tracker and could arbitrarily execute the job. The Yahoo Hadoop community realized the security issues and enforced robust security controls through Kerberos as the authentication mechanism for Hadoop. The latest version of HDFS accomplished the goal using following security goals.

- Kerberos remote procedure call (RPC) for mutual authentication of users, their processes, and services.
- Internet clients implement their pluggable authentication called SPNEGO (Simple and Protected GSSAPI).
- Generic Security Service Application Program Interface and Negotiation Mechanism.
- Access control lists of users and groups on HDFS.
- Delegation tokens and subsequent authentication checks.
- Master Node takes access control decisions for Data blocks access.
- Job tokens ensure task authentication to verify that tasks could do only assigned tasks.
- Enforce network level encryption to ensure security.

E. Proposed Log Format for ISP

In order to store all ISP log and merge into a single file, we have proposed a common log format. This format will include unique ISP id having at least 4 digits and a subscriber id to be provided by his/her related ISP. The individual ISP

log is merged and rotated once a week or a month according to the total data volume. The common log format is provided in Table II.

TABLE II. PROPOSED LOG FORMAT

SN	*Field attribute*	*Description*
1	ISP ID	ISP identification number
2	Subscriber ID	Subscriber identification number
3	Name of Person / Organization	Full name of person or organization
4	Address	Customer address
5	Contact No	Customer phone number
6	Alternate Contact no	
7	Email Address	Customer email address
8	User ID for Internet	Identification number
9	IP Address Assigned	Customer's assigned real IP address
10	Static / Dynamic IP Address Allocation	Natted IP address
11	Source port of the public IP address	Source port for real IP address
12	BST Start Time (hh:mm:ss)	Timestamp in Bangladesh time
13	BST End Time (hh:mm:ss)	Timestamp in Bangladesh time

Log rotation policy is determined according to retention policy and duration i.e., preserve log for six, twelve months or two years. For long duration, log rotation's frequency will be executed frequently to maintain minimum space of any log file and merge all ISP's log file to produce a single file called 'NAT.csv' for processing using below script as illustrated in Fig. 2.

```
#!/bin/bash
OutFileName="NAT.csv"                    # Fix the output name
i=0                                      # Reset a counter
for filename in ./*.csv; do
 if [ "$filename"  != "$OutFileName" ] ;  # Avoid recursion
  then
   if [[ $i -eq 0 ]] ; then
     head -1  $filename >  $OutFileName   # Copy header if it is the first file
   fi
   tail -n +2  $filename >> $OutFileName  # Append from the 2nd line each file
   i=$(( $i + 1 ))                       # Increase the counter
fi  done
```

Fig. 2: ISP Log merge script

IV. IMPLEMENTATION

A. DFS Cluster Environment

Master and Data Nodes must have similar version of debian linux kernel and java or jre version for ideal hadoop's distributed file system. The master node in clusters host the various storage and processing management services. For each Data Node to communicate with its Master Node, we have configured '/etc/hosts' file at all three nodes to add the IP address of the three servers. The master node will use an ssh-connection to connect to other nodes with key-pair authentication to manage the cluster without password. The key based password less SSH connectivity among three nodes will be configured from master node [17]. After completion of key based SSH communication among nodes, from master node 'debian' below commands is executed to synchronize node 1 and 2 hadoop binary installation directory.

```
# rsync -avzh /home/hadoop/ hadoop@node1:/home/hadoop/
# rsync -avzh /home/hadoop/ hadoop@node2:/home/hadoop/
```

After completion of HDFS Data Block replication, Yarn, configuration of Data Nodes, cluster is started. Apache spark is extracted and hdfs user profile is configured to perform sql jobs. Spark's job execution, completed SQL query's duration can be monitored from web gui using url http://103.103.35.137:4040 as shown in Fig. 3.

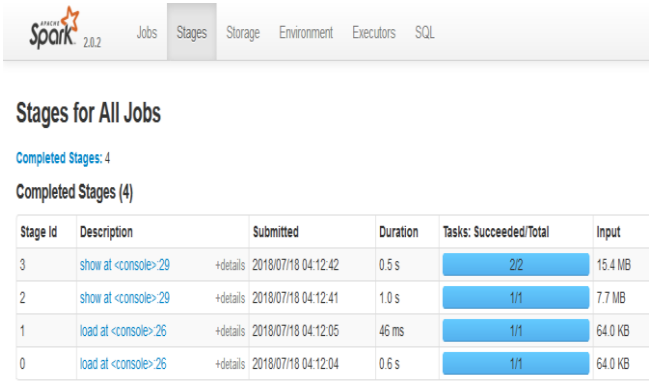


Fig. 3: Apache Spark Web GUI

B. Implementation of Access Control

For access control, Apache Ranger has been installed and configured in third node of distributed file system and integrated with Log analytics. The Ranger administration portal can be accessed using url: 103.103.35.139:6080 where user named as Shazzad and Mahfuz has been created with admin and user role is shown in Fig. 4.

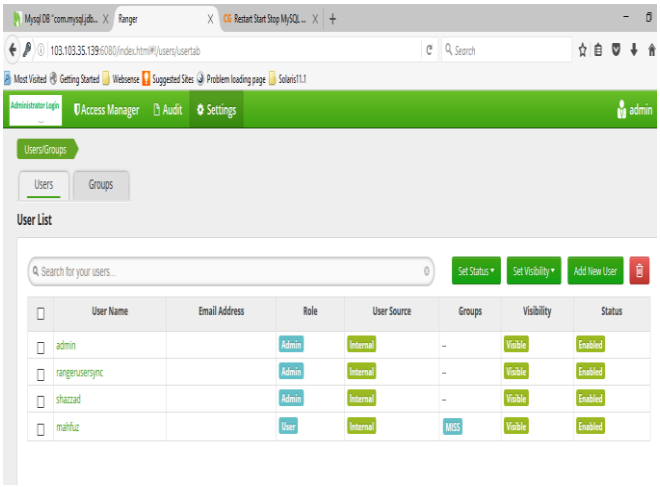


Fig. 4: Log analytics User and Group information

From Fig. 4, user name, role and status can be verified and add, delete, modification of privileges can be done. Audit of administrator and user activity can be accessed from the mentioned web page.

C. Secured Connectivity among Nodes

In order to ensure secured communication among name node and all Data Nodes, we apply following best practices for configuration of VPN and IPTABLE in participating hosts [18].

- Enable Host Based Firewall.

- Allow required Ports i.e. 50070, 8088, 6080 and restrict or deny access to all other ports,
- Configure IPSec tunnel between Master and Slave Nodes in shown in Fig. 5.

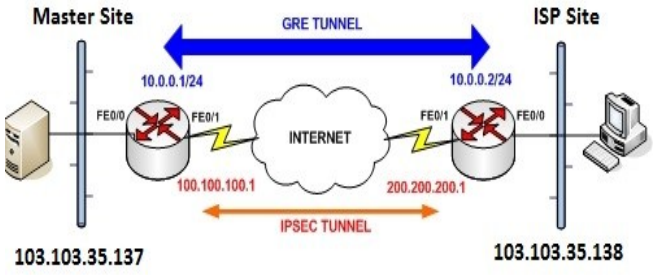


Fig. 5: IPSec over GRE tunnel configuration between Master and one ISP site

- Define access control list (ACL) to allow only master and slave nodes.
- Use strong cryptography key and passphrase for tunnel configuration.
- Allow authorized user or administrator to modify or update network configuration and log activities.
- Arrange for secondary or backup link connectivity between Name node and each Data Node.

D. Setup and Configure NFS for benchmarking

To setup NFS mounts for performance comparison with proposed DFS, one NFS server with three clients are configured. For sharing a directory with NFS, an entry is required in "/etc/exports" configuration file. A new directory named "isplog" in "/" partition is created to share with client server using below commands:

```
# vi /etc/exports
/isplog 103.103.35.137 (rw,sync,no_root_squash)     // NFS server share name
```

At the NFS client end i.e, from Data Node 2, check the nfsshare using below commands:

```
# showmount -e 103.103.35.137
Export list for 103.103.35.137:
/isplog 103.103.35.138
```

Above command shows that a directory named "isplog" is available at server having IP address "103.103.35.137" to share with Data Node 2. To mount that shared NFS directory at the client's node, following command will be used:

```
# mount -t nfs 103.103.35.137:/isplog /mnt/isplog
```

CIFS mount point is configured in the same way for performance comparison.

V. SIMULATION AND RESULTS

The architecture and relevant configuration of the proposed DFS is used to simulate and evaluate the performance of the file systems, query processing and security features. To evaluate the performance of log processing, a comparison between proposed query processing method and map reduce is performed. ISP log is used by Spark utility to perform search based on specific attribute and the latency of read, write operations shell be recorded in proposed DFS, NFS and CIFS shares respectively to draw performance comparison.

A. Data Collection

Our proposed solution processes user's internet activity log in real-time and for that purpose, we have collected ISP user's log from two ISP. User access log contains required information is shown Table III.

TABLE III. SAMPLE ISP USER INTERNET ACTIVITY LOG

Date Time	*Client IP*	*Src Port*	*Dst IP*	*Dst Port*	*Pop Name*
00:59.0	54.255.166.250	80	10.254.27.206	39389	104.30
00:59.0	10.254.31.76	51906	31.13.79.63	443	104.30
00:59.0	10.254.27.19	42978	47.88.65.224	6798	104.30
00:59.0	172.217.166.106	443	10.254.27.237	60589	104.30
00:59.0	43.245.142.146	443	10.254.31.202	38215	104.30

B. Experiments with Proposed Method

From below query using Spark's scala command shell, we can identify the client communicating with destination POP address in Table IV where connected clients source, source port, destination IP address, destination port no. and POP id through which the clients connect to internet are displayed.

```
scala>df.select("ClientIP","Src_Port","Dst_IP","Dst_Port","Pop_Name").show()
```

TABLE IV. SCALA OUTPUT

Client IP	*Src_Port*	*Dst_IP*	*Dst_ Port*	*Pop_Name*
54.255.166.250	80	10.254.27.206	39389	Cyber-POP-104.30
10.254.31.76	51906	31.13.79.63	443	Cyber-POP-104.31
10.254.27.19	42978	47.88.65.224	6798	Cyber-POP-104.32
172.217.166.106	443	10.254.27.237	60589	Cyber-POP-104.33

C. Comparison between Proposed Method and Map Reduce

Proposed Spark method can do processing in memory, while Hadoop MapReduce has to read from and write to a disk . As a result, the speed of processing differs significantly, Spark may be up to 100 times faster. However, the volume of data processed also differs. Spark's performance over map reduce may be obtained with work count programs performance which was tested and result is provided in Fig. 6. In Fig. 6, the analysis used three Data Nodes and processed up to 48 GB data and comparison of required time by hadoop's map reduce and spark are shown in the graph as blue (Hadoop) and red or orange (Spark). Fig. 7 and 8 shows that proposed DFS provides much better read and write latency respectively over NFS and CIFS. For 100 GB data read and write operation, latency of proposed DFS is lower than that of CIFS and NFS.

D. Security Comparison of Proposed DFS

Security comparison among NFS, CIFS and proposed DFS is described in below Table V.

TABLE V. SECURITY COMPARISON AMONG NFS, CIFS AND PROPOSED DFS

File System	*NFS*	*CIFS*	*Proposed DFS*
Architecture	File-based	File-based	Cluster based
Processes	Stateless	Stateful	Stateful
Naming	No central metadata server	No central metadata server	Central metadata server
Synchronization	No	No	Write-one-read-many
Session Requirement	No	Yes	Yes
Communication speed	Fast	Slow	Fast
Fault Tolerance	No	No	Yes
Security mechanism	No	Kerberos	Security in the form of AAA

From Table V it is found that the architecture of NFS and CIFS is file based protocol but proposed DFS is cluster based. NFS has stateless process, CIFS and proposed DFS have stateful process. There is no central metadata server for NFS and CIFS except proposed DFS. There is no synchronization in NFS and CIFS. Communication speed is fast in both NFS and proposed DFS. There is no security mechanism in NFS, kerberos in CIFS and AAA security in proposed DFS



Fig. 6: Comparison between proposed method and map reduce (3 Instances)

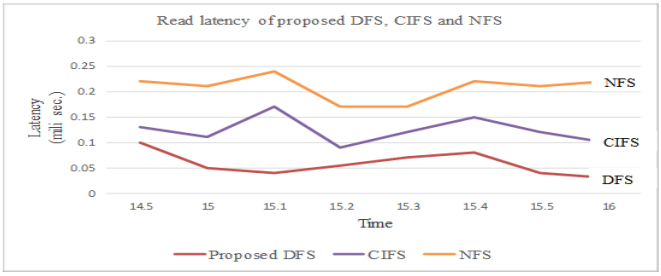


Fig. 7: Read latency of proposed DFS, CIFS and NFS

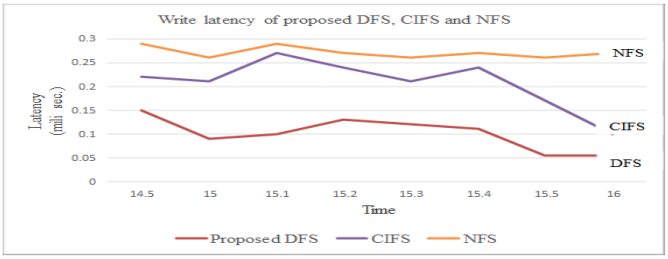


Fig. 8: Write latency of proposed DFS, CIFS and NFS

VI. CONCLUSION

In Perspective of Bangladesh a centralized log analytics system has been designed to access and monitor user's activity in internet using commodity hardware and open source software for cost optimization and rapid deployment. All experiment results regarding log analytics distributed file system implementation, map reduce algorithm, no-sql query using spark and access control features have been included. The log analytics system offers flexibility and ability to process large datasets. Our proposed solution will also preserve user activity log for a long duration to detect and investigate suspicious activity that occurred a long time ago.

REFERENCES

[1] A. S. Tanenbaum and M. Steen, Principles and Paradigms, 3rd ed., Prentice Hall, 2003.

[2] H. Darwen, An Introduction to Relational Database Theory, 3rd ed., Ventus Press, 2009.

[3] B. Callaghan, NFS Illustrated. 1st ed., Addison-Wesley Professional Press, 1999.

[4] L. Jiang, L. Bing, and M. Song, "The optimization of HDFS based on small file", pp. 912-915, Sep. 2010.

[5] G. Saranya, V. Janardhan babu, and G. H. Kumar, "High Performance Analytics of Bigdata Using Hadoop Framework," Journals of Advanced Research in Computer Science and Software Engineering, vol. 8, pp. 200-203, May 2018.

[6] Bangladesh Telecommunication Regulatory Commission (BTRC), "BTRC License Summary", btrc.gov.bd, pdf file, Feb. 13, 2019. [Online]. Available: http://www.btrc.gov.bd/license-summary-0. [Accessed Oct. 12, 2018].

[7] CoreStack, "The Log Mandate Enabling Indian ISPs to adhere to DoT compliance rules", corestack.io, para. 1-2, June 1, 2017. [Online]. Available: https://www.corestack.io/blog/the-log-mandate-enabling-indian-isps-to-adhere-to-dot-compliance-rules/ [Accessed Oct. 15, 2018].

[8] A. Chowdhury, M. Chowdhury, and M. Imran, "Branding strategies for service firms- a study on the selected Internet Service Providers (ISPs) in Bangladesh", Journals of Asian Business Review, vol. 2, no. 1, 2015. [Accessed Jun. 2018].

[9] T. White, The Definitive Guide, 4th ed. O'Reilly Media, 2015.

[10] A. D. Lewine, POSIX Programmer's Guide, 1st ed., O'Reilly Media, ,1994.

[11] Pete Loshin, TCP/IP Clearly Explained. San Francisco, California: Morgan Kaufmann Publishers, 1998.

[12] N. B. Brian, E. T. Anderson, D. E. Lazowska and M. H. Levy, "Lightweight Remote Procedure Call," ACM Transactions on Computer Systems, vol. 8, no. 1, pp. 37-55, Feb. 1989.

[13] P. Priya Sharma and P. Chandrakant Navdeti, "Securing Big Data Hadoop: A Review of Security Issues, Threats and Solution," Journal of Computer Science and Information Technologies, vol. 5(2), pp. 2126-2131, May 2014.

[14] H. Lin, S. Seh, W. Tzeng, and B. P. Lin, "Data Confidentiality via Integrating Hybrid Encryption Schemes and Hadoop Distributed File System" In Proc. 26th IEEE International Conference on Advanced Information Networking and Applications, '03,2012, pp.745-747.

[15] M. Chithik Raja, and M. Munir Ahamed Rabbani, "Comprehensive and Coordinated Security of Knox gateway in Big Data," Journal of Advance Research in Science and Engineering, vol.4, issue 3, pp. 61-69, Mar. 2015.

[16] A. Aleksiyants, O. Borisenko, D. Turdakov, A. Sher and S. Kuznetsov, "Implementing Apache Spark jobs execution and Apache Spark cluster creation 3 for Openstack Sahara", *Proceedings of the Institute for System Programming of RAS*, pp.35-48, 2015..

[17] D. S. Mansfield, "Interview: TatuYlonen, SSH Communications Security", Computer Fraud & Security, vol.2012, no.5, pp.13-16,2012.

[18] S. Jung and K. Sung,"A study on the Iptables Ruleset Against Dos Attacks", The Journal of Korea Navigation Institite, vol. 19, no.3, pp.257-263,2015.

Feature Selection and Comparative Analysis of the Supervised Learning Model for Hyperspectral Image Classification

Abu Sayeed[1], Md. Ali Hossain[2] and Md. Rabiul Islam [3]
[1,2,3] Dept. of Computer Science & Engineering, Rajshahi University of Engineering and Technology, Rajshahi-6204, Bangladesh
Email: abusayeed.cse@gmail.com[1], ali.ruet@gmail.com[2] and rabiul_cse@yahoo.com[3]

Abstract*— In remote sensing image classification, really it is an intimidating when kernel supervised learning approaches stands in need of adequate amount of training samples. Often there is a vital problem for definition and acquisition of reference data. For Hyperspectral image classification, improved spectral information is required to make it suitable for ground object identification. In this paper, Support Vector Machine with RBF kernel (KSVM) and the spectral angle mapper (SAM) are used for performance comparison in terms of classification accuracy in Hyperspectral image classification. Kenel support vector machine is more preferable for the mastery to generalize better hyperplane when limited availability of training samples and separate the classes competently in a new dimension feature space. Experiments are performed on NASA Airborne Visible Infrared Spectrometer (AVIRIS) image and it shows KSVM outperforms SAM and obtains the highest accuracy. Due to more well-conditioned against the outliers, KSVM significantly reduced the classification complexities than SAM.

Keywords*— Spectral Information, Object Identification, Hyperspectral Images Classification, Support Vector Machine, Spectral Angle Mapper.

I. INTRODUCTION

Earlier scientists and geologists obtained land cover information through manual surveying by observing places on movable cars, bus or foot. In many circumstances, it was found that the above methods are not efficient and not practical in some location of the world. For example, suppose you have to implement this for large and remote areas. So now you might be thinking how much time and money should be required and also thinking if there are other way to do so. So manual surveying approach will practically require more expensive and consume more time.

The geographical land information is captured by mounting camera on aircraft and using the power of this technique, the researchers need to obtain land cover information in much shorter time. However, they found such method dependent on the weather to some extent for the increasing number of air accidents. In accordance with the development of modern space technologies, remote sensing satellite took the place of aerial photography technology where satellites captured land cover information by satellite's sensor. The cost slope is relatively slow down when this opportunity are created by satellite sensor than the aerial photography that makes the whole operation of monitoring land covers efficient and several places can be covered and observed at a time [1-3]. As a result, in a particular slot of time, the time to time changed land cover information detect effectively. By the change of land cover natural conditions and human activities related to environmental, agriculture and urbanization, it is essential to make accurate decision [4-7]. Previously different digital classification algorithms have been used to predict the nature and classes existent in those images such as Maximum Likelihood Classification (MLC), Vertex component analysis (VCA), Kernel Support Vector Machine (SVM) and Spectral Angle Mapper (SAM) etc. To overcome the lacking of MLC, SAM was got introduced. The Performance of these classifiers depends on the accurate selection of training samples. When the input dimension is high and the limited availability of training sample are presented, then the hyperspectral image classification is seriously daunted with poorly classified labeled training samples. To solve this problem, the alternative standard multiclass classifiers are developed. Spectral Angle Mapper (SAM) classifier is technique that matches each pixel in the input image with every end member for each class instead of matching it to the input training pixels [7]. The problem of this SAM classifier is that is does not considered the sub-pixels when the end member is selected and creates more problem for multiclass or heterogeneous earth surface. Again the development of the accuracy of classification can be obtained by the improved the kernel support vector machine which makes a good job of classification for Hyperspectral image [7-8]. By using SVM with RBF kernel, the optimal separation is obtained by generating a hyperplane between the classes considered. The support vector machine selects in the race of thousands of decision boundaries if pattern recognition has two classes that are linearly separable and select the one that has a minimized generalization error. By this, the decision boundary that has been taken into consideration has to be the one that has biggest margin between two class levels, where margin is equal to sum of the length to the hyper-plane of the two classes' closest point. By means of standard quadratic programming, the problem of maximizing the margin can be solved. The margin is measured by using closest distance between the data points and the hyper-plane and these data points are called support vectors [9-10]. Therefore, in term of counting, small numbers of support vectors are present. As detecting environmental and global ecology alter, main forwardness has to be taken about land cover and land use. Land use or land cover classification is an essential element; in there that element is made of different parameters. Different parameters are jointed together on the rule of necessities. Agricultural land, water, buildings, soil, fallow bare land, vegetations, and artificial cover are these that indicate the land cover and others seen on the land [11].

II. CLASSIFICATION TECHNIQUES

Two classification techniques are adopted in this study to classify the hyperspectral image classes.

A. Spectral Angle Mapper (SAM)

Spectral Angle Mapper (SAM) approach is a supervised technique that uses spectral angular information for the task of image classification. The SAM works well since it utilizes the spectral similarity between the reference spectra

and the image spectra [13-15]. SAM measures spectral similarity by calculating the angle between two spectras. This spectrum is a vector in a space with dimensionality that is equal to the number of spectral bands. The range of spectral angle is (0 to 1). Unlike the distance metrics it is possible to have a zero spectral angle even when the two vectors are not identical. Therefore this approach is not sensitive to albedo effect and illumination because the angle between two vectors is not variant according to the length of the vectors. SAM between two spectral signatures with L bands $\mathbf{S}_i = (\mathbf{S}_{il}...\mathbf{S}_{il})^T$ and $\mathbf{S}_j = (\mathbf{S}_{jl}...\mathbf{S}_{jl})^T$ is defined as,

$$\mathbf{SAM}\ (\mathbf{S}_i, \mathbf{S}_j) = \mathrm{Cos}^{-1}(\Theta_{si,sj}) \qquad (1)$$

where,

$$\theta(s_i, s_j) = \left(\frac{\sum_{i,j=1}^{L} s_i s_j}{\left[\sum_{i=1}^{L} s_i^2\right]^{\frac{1}{2}} \left[\sum_{j=1}^{L} s_j^2\right]^{\frac{1}{2}}} \right)$$

This classifier is robust against the solar illumination effect. Therefore it is said that the SAM is a robust classifier as it is effective to overcome to the shading effect to target reflectance characteristics. The major disadvantages created by SAM is that it considers the end members taken to classify an hyperspectral image by creating the genuine spectra of a reference material, where the heterogeneous earth's surface are present in many ways and make of mixed pixels.

B. *Support Vector Machine(SVM)*

This classification considers separate data sets for testing sets and training set. All the training set carries different target value and different characteristics. To predict the target values of the given test data by SVM that is the main goal of SVM. The nonlinear SVM is used to calculate the inner product of the mapped samples which avoids the difficulty in determining the mapping function ϕ and also the cost for calculation of the mapped samples and their inner product. Several typical kernels including linear, polynomial, and radial basis function (RBF). Though SVM is initially developed for two-class problems, It has been extended to deal with multiclass classification based on either combination of decision results from multiple two-class classifications or optimization on multiclass based learning [3]. In the first phase, SVM is used for initial training and classification. In the second phase, these results have been used for probability based modeling. The open-source library livSVM is used for initial training and classification of the aforementioned four datasets, and both the linear and the Gaussian radial basis (RBF) kernels are tested. For each group of datasets, all the data are normalized to [−1, 1] before SVM is applied. Through 10-fold cross validation, the best group of parameters, including the cost and the gamma value, are optimally determined. Eventually, the optimal parameters are used for classification. Without misclassification, a linear hyper-plane is unable to separate classes that are common cases. For this special case, a nonlinear separating hyper-plane can able to separate those classes. By using a nonlinear transformation function, data are mapped into a high dimension space. Data are expanded in high dimension space so that a linear separating hyper-plane can be developed [14]. This is the idea of a principle *Cover's theorem* on the *separability of patterns*. In the figure below that describes one that when two classes are in an input space, linear separating hyper-plane cannot separate two classes. A linear separating hyper-plane can be acquired wherein data has to be mapped in the high dimensional feature space [14].

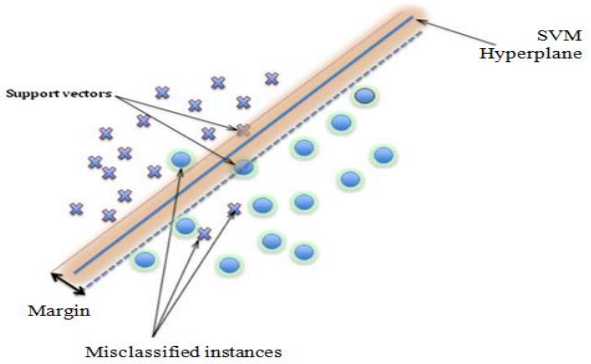


Figure 1: Linear support vector machine example.

Now Θ is a nonlinear transformation function that maps all data into a high dimensional feature space. Let there is a function *T*, that is said to be a kernel function, so that:

$$T(\mathbf{X}_i, \mathbf{X}_k) \equiv \Theta(\mathbf{X}_i) \cdot \Theta(\mathbf{X}_k) \qquad (2)$$

The kernel function Θ is replaced by the explicit form of Θ and dot product of the transformed vectors. Again, the less computationally intensive is provided by the use of the kernel function [15-17]. *Mercer's theorem* is used by a special case of dot product and the formation of the kernel function.

III. METHODOLOGY

In this paper, the research includes with three stages: data preprocessing, using standard feature selection method (statistical information measure feature selection), using Spectral Angle Mapper and SVM RBF kernel for classification and comparison of the results of those two methods.

A. *Data Pre-Processing*

We have worked with NASA Airborne Visible Infrared Spectrometer (AVIRIS) image. We have manipulated the image using Multispec tools application and 8 classes (Soy, Woods, Corn-nottill, Hay, Soy-till, Corn, Wheat, Corn-min) were taken in the image area among of all 16 classes are available in the dataset. We have taken five features more specifically five channels and studied on the classification accuracy results given by SAM and SVM based on them. That's why it is necessary to extract the best possible features by using standard feature selection method.

B. *Feature Extration*

During the selection and manipulation of some particular bands and features of our specified data set, our study has been involved with the feature extraction from the

specialized NASA Airborne Visible Infrared Spectrometer (AVIRIS) image data. For this we have used the mutual information based feature selection approach [16]. This approach is selected in this case for the feature selection because it selects the features that have enhanced discrimination power. The classification accuracy of MIFS features is also higher than the classification in the original space [16]. That is why we have used this method for the extraction of feature from our AVIRIS data set.

C. *Data Operation in Spectral Angle Mapper*

The SAM classifier is one of most emerging classification technique cause of calculating the similarity of spectra by creating the robustness against the influence of shading to underscore the target reflectance characteristics. For hyperspectral image classification, the Multispec tools are used.

D. *Data Operation in Support Vector Machine (RBF kernel)*

On the statistical information based feature selection of hyperspectral remote sensing image, the Kernel Support Vector Machine (SVM) is suitable for classification [14]. Nowadays, Kernel support vector machine is very emerging technique for hyperspectral image classification [17-19]. SVM is implemented to classify the multiclass problem by trickily creating KSVM directly, but due to the number of the classes essential for being correctly classified increases, parameters number for being estimated increases. KSVMs finally affect the classification performance in terms of accuracy. We use livSVM in matlab tools. We have implemented this operation by three steps; firstly, we have performed the input data scaling in the dimension [0 1], secondly we have performed the cross validation check to find out the best value of c(cost estimation factor value) and the gamma to carry out the testing of our referred AVIRIS data set. The values of c and gamma will be used to train our selected data values in accordance with the band value selected by feature selection. Finally we will calculate the training accuracy with best c and gamma value and get the best training accuracy for the specified values referred to as our extracted band values.

IV. Experimental Analysis

We have used NASA Airborne Visible Infrared Spectrometer (AVIRIS) hyperspectral data for our experimental analysis. Spectral Angle Mapper and Support Vector Machine (RBF Kernel) was used for classification. The performance of the proposed approach was evaluated using one real hypersepectral image data sets. Image consisted of 220 channels and was collected by an AVIRIS sensor over the agriculture area of Indian Pine in the north part of Indiana Sixteen classes is defined in the ground truth [23-26]. In this experiment, the numbers of eight classes are used. They are Soy, Woods, Corn-nottill, Soy-till, Hay, Corn, Wheat and Corn-min.

A. *Accuracy Assement*

For assessing the classification accuracy, the confusion matrix was used by using the ground truth image. In the experiment, the percentage of correctly classified pixels refers to the overall accuracy.

B. *Dataset Description*

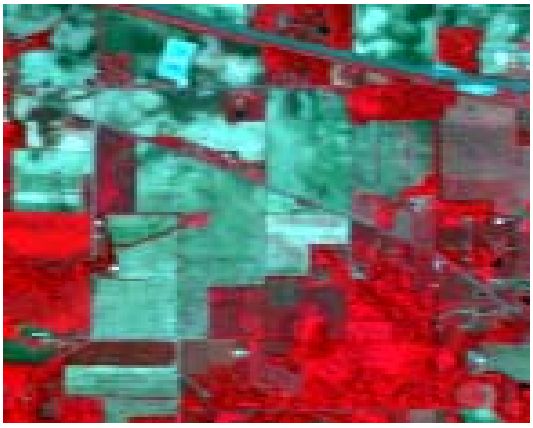
The performance of the proposed approach was evaluated using one real hypersepectral image data sets. Image consisted of 220 channels and was collected by an AVIRIS sensor over the agriculture area of Indian Pine in the north part of Indiana Sixteen classes is defined in the ground truth [23-26]. In this experiment, the numbers of eight classes are used. They are Soy, Woods, Corn-nottill, Soy-till, Hay, Corn, Wheat and Corn-min.

Table 1: Classes used in the experiments for NASATA AVIRIS dataset

AVIRIS Dataset			
Class Name	**Class No**	**Train**	**Test**
Soy	1	96	168
Woods	2	170	130
Corn-nottill	3	96	126
Hay	4	120	128
Soy-till	5	110	105
Corn	6	56	30
Wheat	7	63	84
Corn-min	8	90	80
Total	8	801	851



(a)



(b)

Figure 2: (a) Ground Truth Image (b) Original AVIRIS 220 band hyperspectral Image

C. *Experimental Result and performance analysis*

Spectral Angle Mapper classifier and KSVM is applied on AVIRIS Indian pine dataset ('92AV3C.lan) after performing extracting feature by using Standard feature extraction algorithm (Statistical Information based Feature selection-SIBFS) is used for best feature. The selected image bands are 193, 65, 31, 170, 2. In the experiments, 8 classes are used Soy, Woods, Corn-nottill, Soy-till, Hay, Corn, Wheat and Corn-min for the evaluation of the suitability of the feature selection. Table 2 shows the overall accuracy of the SIBFS selected features when using Kernel SVM. Error matrics are calculated for both training and testing performance shown Table 3 and Table 4 respectively when using Kernel SVM classifier. Table 5 shows the overall accuracy of the SIBFS selected features when using SAM. Error matrics are calculated for both training and testing performance shown Table 6 and Table 7 respectively when using SAM classifier. Finally overall performance of accuracy KSVM vs. SAM based on SIBFS are shown in Table 8 and graphically shown in Figure 3.

Table 2: Shows KSVM accuracy on SIBFS feature band

No. of features	KSVM Accuracy (%)
1(Using images band: 193)	60.549
2(Using images bands:193,65)	79.400
3(Using images band:193,65,31)	98.127
4(Using images band:193,65,31,170)	99.990
5(Using images band : 193,65,31,170,2)	99.997

Table 3: Error matrix of training performance of KSVM on SIBFS

Class No	Pixel No	Number of Samples in Class							
		1	2	3	4	5	6	7	8
1	96	96	0	0	0	0	0	0	0
2	170	0	170	0	0	0	0	0	0
3	96	0	0	96	0	0	0	0	0
4	120	0	0	0	120				
5	110	0	0	0	0	110	0	0	0
6	56	0	0	0	0	0	56	0	0
7	63	0	0	0	0	0	0	63	0
8	90	0	0	0	0	0	0	0	90
Total	801	96	170	96	120	110	56	63	90

Overall training performance (801/801) = 100%

Table 4: Error matrix of testing performance of KSVM on SBFS

Class No	Pixel No	Number of Samples in Class							
		1	1	3	4	5	6	7	8
1	**168**	168	0	0	0	0	0	0	0
2	**130**	0	130	0	0	0	0	0	0
3	**126**	0	0	126	0	0	0	0	0
4	**128**	0	0	0	128	0	0	0	0
5	**105**	0	0	0	0	105	0	0	0
6	**30**	0	0	0	0	0	30	0	0
7	**84**	0	2	0	0	0	0	82	0
8	**80**	0		0	0	0	0	0	80
Total	**851**	168	132	126	128	105	30	82	80

Overall testing performance (849/851) = 99.997%

Table 5: Shows SAM accuracy on SIBFS feature band

No. of features	Testing Accuracy (%)
1(Using images band: 193)	19.7
2(Using images bands:193,65)	47.1
3(Using images band:193,65,31)	62.0
4(Using images band:193,65,31,170)	65.6
5(Using images band : 193,65,31,170,2)	54.6

Table 6: Error matrix of training performance of SAM on SIBFS

Class No	Pixel No	Number of Samples in Class							
		1	2	3	4	5	6	7	8
1	**96**	53	0	10	0	25	2	6	0
2	**170**	0	118	0	0	0	0	0	52
3	**96**	16	0	42	0	17	21	0	0
4	**120**	0	0	0	120	0	0	0	0
5	**110**	42	0	14	0	33	19	2	0
6	**56**	0	0	1	0	3	45	7	0
7	**63**	18	0	0	0	1	13	31	0
8	**90**	0	47	0	0	0	0	0	43
Total	**801**	129	170	165	120	79	100	46	95

Overall training performance (496/801) = 61.9%

Table 7: Error matrix of testing performance of SAM on SIBFS

Class No	Pixel No	Number of Samples in Class							
		1	2	3	4	5	6	7	8
1	**168**	86	0	16	0	53	2	11	0
2	**130**	0	49	0	0	0	0	0	81
3	**126**	16	0	71	0	22	16	0	0
4	**128**	0	0	0	128	0	0	0	0
5	**105**	7	0	12	0	27	50	9	0
6	**30**	0	0	1	1	0	27	2	0
7	**84**	12	0	0	0	1	40	31	0
8	**80**	0	34	0	0	0	0	0	46
Total	**851**	**122**	**83**	**99**	**129**	**103**	**135**	**53**	**127**

Overall testing performance = 54.6%

Table 8: Performance of accuracy KSVM vs. SAM based on SIBFS

No. of features	Accuracy KSVM(%)	Accuracy SAM(%)
1(Using images band: 193)	60.549	19.7
2(Using images bands:193,65)	79.400	47.1
3(Using images band:193,65,31)	98.127	62.0
4(Using images band:193,65,31,170)	99.990	65.6
5(Using images band : 193,65,31,170,2)	99.997	54.6

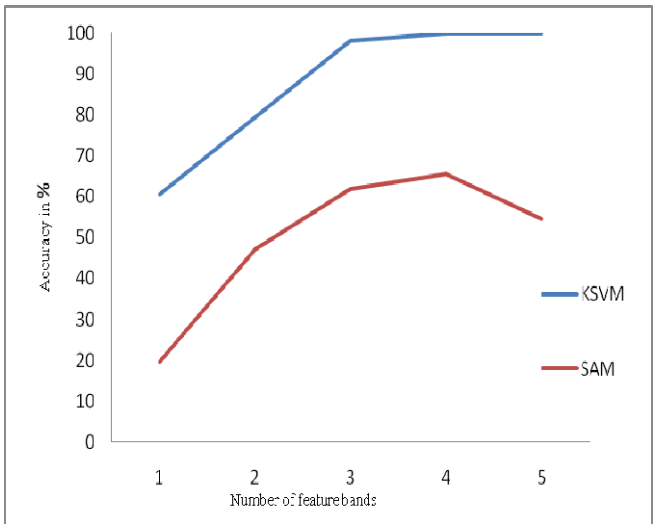


Figure 3: Performance of testing accuracy KSVM vs. SAM based on SIBFS.

V. CONCLUSION

From the performance analysis it is quite clear that KSVM provide better result than SAM. KSVM provide a good classification result due to creating an optimal hyper plane between the target classes and rest of the classes. It has the ability to separate the complex class pairs easily in the high dimensional space. KSVM also gives a good result since some classes are complex in structure and were not linearly separable. The main disadvantages of SAM classifier is that it chose the end members to classify an image by symbolizing the pure spectra of a reference material, in where the earth's surface is heterogeneous and made of mixed pixels. Hence the strength and weakness of the KSVM and SAM are analyzed.

REFERENCES

[1] A. Singh, "Review Article Digital change detection techniques using remotely-sensed data", International Journal of Remote Sensing, vol. 10, no. 6, pp. 989-1003, 1989

[2] King R B, "Land cover mapping principle: a return to interpretation fundamentals", Int. J. Remote Sens. 23,2-3546 in 2002

[3] K.R. Manjula, J. Singaraju J. and A.K. Varma, "Data preprocessing in multi-temporal remote sensing data for deforestation analysis", Global Journal of Computer Science and Technology Software & Data Engineering, vol. 13, no. 6, pp. 1-8, 2013.

[4] N.IS. Bahari, A. Ahmad and B.M. Aboobaider, "Application of support vector machine for classification of multispectral data", 7th IGRSM. International Remote Sensing & GIS Conference and Exhibition. IOP Conf. Series: Earth and Environmental Science 20, 2014.

[5] A. Ahmad and S. Quegan, "The effects of haze on the spectral and statistical properties of land cover classification", Applied Mathematical Sciences, vol. 8, no. 180, pp. 9001-9013, 2014.

[6] A. Ahmad and S. Quegan, "Haze modelling and simulation in remote sensing satellite data", Applied Mathematical Sciences, vol. 8, no. 159,pp. 7909-7921, 2014.

[7] Girouard G, Bannari A, Harti A E and Desrochers A,"Hyperspectral image classification", Istanbul Turkey XXth ISPRS 599 – 604 in 2004

[8] Asmala Ahmad and Shaun Quegan, "Analysis of Maximum Likelihood Classification", Applied Mathematical Sciences, Vol. 6, 2012, no. 129, 6425 – 6436.

[9] Boser, B. E., Guyon, I. M., & Vapnik, V. N.,"*A training algorithm for optimal margin classifiers",* In the Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory, 27-29 July 1992, Pittsburgh, PA, USA.

[10] Vapnik, V. N. (1998),"*Statistical learning theory*", New York: Wiley in 1992.

[11] Hughes, G. F., "On the mean accuracy of statistical pattern recognition." *IEEE Trans. Inform. Theory, 14*(55-63) in 1968.

[12] Kruse F A, Lefkoff A B, Boardman J W, Heidebrecht K B, Shapiro A T, Barloon P J and Goetz A F H ,"*Remote Sens". Environ.* 44 145-163in 1993.

[13] Crosta A P, Sabine C and Taranik J V, "*Remote Sens", Environ.* 65 309-319,1998.

[14] Schwarz J and Staenz K 2001 "Remote sensing *Can. ", J. Remote Sens.* 27 216-224, 2001.

[15] Hunter E L and Power C H 2002 *Int. J. Remote Sens.* 23 2989-3008.

[16] Dr Md Ali Hossain, Mark Pickering,Xiuping Jia, "Improved feature selection based on a mutual information measure for hyperspectral image classification", International Geoscience and Remote Sensing Symposium(IGARSS) in 2011.

[17] Landgrebe D A," Signal Theory Methods in Multispectral Remote Sensing Hoboken" , NJ: John Wiley & Sons in 2003.

[18] Huang C, Davis L S and Townshend J R G 2002 Int. J. Remote Sens.23 725-749

[19] Kluwer Cortes C and Vapnik V, "Support-Vector Networks, Machine Learning", Lorenza Saitta Netherlands Academic Publisher, Boston 20 273-297 in 1995.

[20] Burges C J C, "A tutorial on support vector machines for pattern recognition", Data Mining Knowedge Discovery 2 121–167 in 1998.

[21] C. Chang and C.-J. Lin, " LIBSVM: a library for support vector machines", ACM Transactions on Intelligent Systems and Technology, 2(3):27:1 27:27, 2011.

Analysis of Performance-Improvement of Microstrip Antenna at 2.45 GHz Through Inset Feed Method

A. K. M. Baki[1], Md. Nurur Rahman[2], Shawon Kumar Mondal[3]
Dept. of Electrical and Electronic Engineering, Ahsanullah University of Science and Technology (AUST)
141 & 142 Love Road, Tejgaon I/A, Dhaka, Bangladesh
baki.akalam.eee@aust.edu[1], md.n.rahman13@gmail.com[2], aoschu94@gmail.com[3]

***Abstract*—Microstrip patch antenna (MSA) has numerous applications such as aircraft, biomedical engineering, connected vehicles, cellular phones, satellites, smart grid, and spacecraft. Single MSA or MSA Array (MSAA) require low-cost materials with simple and inexpensive fabrication techniques. With the help of modern and highly precise printed-circuit technology it is possible to fabricate inexpensive and robust MSA/MSAA, which are compatible with microwave monolithic integrated circuit (MMIC) designs. Rectangular patch, one of the most popular patches, has different feeding techniques for impedance matching. Inset fed rectangular patch does not require any additional complex microwave circuit. Inset feeding mechanism can easily be optimized with the help of 3D simulation software. Very few investigations are made on inset fed MSAA, though several research papers can be found on inset fed rectangular MSA. In this paper a comparative analysis of edge fed and inset fed MSA/MSAA at 2.45 GHz frequency band is done using FR-4 and air substrates. It was found that the inset fed MSA with air substrate performs better when the directivities, gains, bandwidths and return losses are considered. Though FR-4 substrate is cheap and miniaturization of MSAA is better with FR-4 substrate.**

***Index Terms*—antenna radiation patterns, directivity, gain, microstrip patch, impedance matching.**

I. Introduction

While most of the advances in Microstrip patch antenna (MSA) and MSA array (MSAA) are due to military and space applications, the utility of this technology has been rapidly gaining interest in commercial sectors [1]. On the other hand the 5th generation (5G) wireless systems will depend on a huge number of smaller cells to offload the traffic from the cellular and from other wireless networks [2]. One of the distinctive features of 5G technology is its intrinsic flexibility of frequency spectrum, which will allow it to support different application in an efficient way, either using low-band below 1 GHz, mid-band from 1 GHz to 6 GHz, or high-frequency band above 6 GHz [3].

MSA and MSAA have applications in different technological areas such as aircraft, biomedical engineering, connected vehicles, cellular phones, satellites, smart grid, and spacecraft. A microstrip patch generally consists of a metallic patch, which may take many different forms, on a grounded plane [1]. A dielectric medium (with a particular dielectric constant), called the substrate, is placed between the radiating patch and the grounded plane.

Some of the characteristics of MSA/MSAA are radiation pattern, beam width, directivity, gain, bandwidth, quality factor, efficiency, polarization, and return loss. The specifications for military and space applications typically put emphasis on best performance with almost no constraint on cost. On the other hand commercial applications seek low cost components sacrificing some electrical performances [1]. MSA/MSAA for commercial applications require low-cost materials with simple and inexpensive fabrication techniques. Microstrip patch can be planar or non-planar. With the help of modern printed-circuit technology it is possible to fabricate inexpensive and robust MSA/MSAA, which are compatible with microwave monolithic integrated circuit (MMIC) designs. There exists different shapes of patches for different applications. Improvements in the properties of the dielectric materials and in design techniques have led to enormous popularity of MSA [4]. Among different shapes, the rectangular MSA is the most popular [1]. Fig.1 illustrates the basic structure of a rectangular shape MSA.

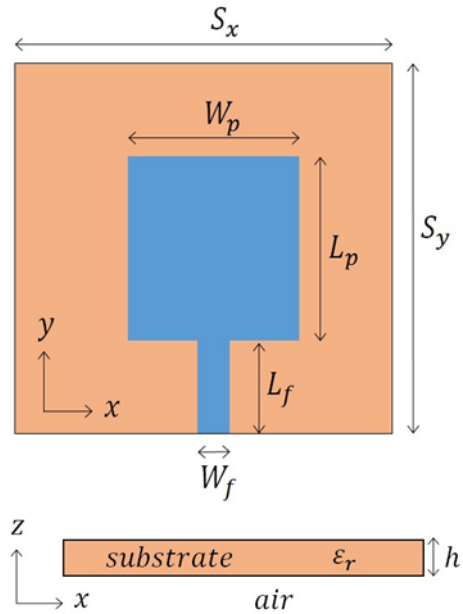


Fig. 1: Rectangular shape microstrip patch antenna.

There are several feeding techniques for microstrip patch antennas such as:

a. Aperture coupled feed
b. Coaxial probe feed
c. Microstrip transmission line feed:
    i. Edge feed-
        1. Center feed
        2. Offset feed
    ii. Inset feed
d. Proximity coupled feed
e. Quarter-wave line feed

Design complexities of MSA/MSAA increase with the miniaturization of the antenna. MSA is a resonant device and suffers from a number of limitations such as low bandwidth, low gain, low power handling capacity and higher return loss. An effort to improve one parameter, such as return loss, results in other complexities. Detailed methods of antenna designs and optimization processes are described in [1], [5]- [8].

In this paper we have investigated the performances of edge fed and inset fed rectangular patches at 2.45 GHz frequency band. We have used the HFSS software as a simulation tool.

The paper is organized as follows. Section II describes the performances of inset fed rectangular patch in terms of its input impedance. Section III makes a comparative analysis of edge fed and inset fed MSA using two different substrates, one is air and the other is FR-4. In section IV an analysis of 4×1 edge fed MSAA and 4×1 inset fed MSAA are made. Section V describes the fabrication of an inset fed rectangular MSA with copper material and air substrate. Finally, conclusion is made in section VI.

II. Analysis of the Input Impedance of Inset Fed Rectangular Patch Antenna

Each feeding method has its own merits and demerits. The feed structure affects the bandwidth [1], directivity, gain, and return loss of antenna. With inset feed patch, a good matching condition can be obtained [1], [6] and a wide impedance bandwidth can also be achieved [6]. Inset feed method does not require any additional matching arrangement. It is achieved by properly controlling the inset position and width. Fig. 2 shows an inset microstrip line fed antenna. Inset feed scheme is an easy feeding scheme since it provides ease of fabrication with simplicity in modeling as well as impedance matching [1].Though improvement procedures of inset feed patch have not been widely investigated in case of array antenna.

Two different feed mechanisms of patch antennas, resonating around 720 MHz, using a probe feed and an inset microstrip-line feed, have been discussed in [6]. The notch cut of the inset fed rectangular patch decreases the effective length of the excited patch surface current, and the effect is larger for a larger notch [6]. An antenna occupying more space in a spherical volume will have a wider bandwidth (BW) [1]. One way of increasing the BW is to increase the height (h) of the substrate. Increase in height also results in more efficient antenna. However higher value of h may result in surface waves that travel within the substrate. These phenomena result in undesired radiations which may cause other problems.

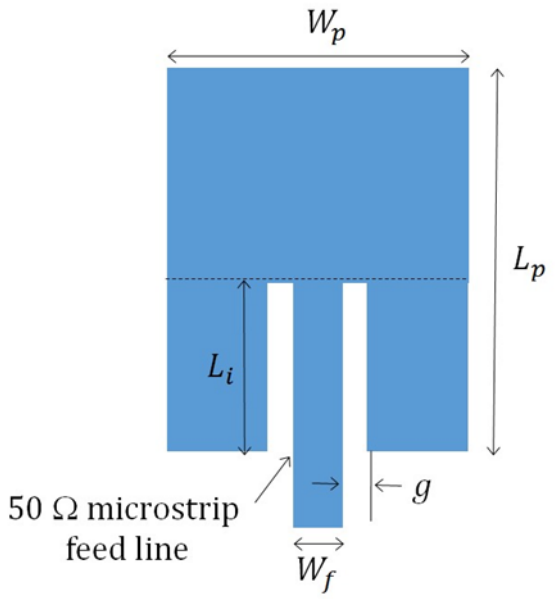


Fig. 2: Inset fed Microstrip Patch Antenna.

By inset feeding of the MSA, the input impedance can be decreased. Let's assume that the patch length is $L_p = 0.61\lambda$ (i.e. $\lambda = 1.64L_p$). The phase change in the length L_i (Fig. 2) is expressed as $\frac{2\pi L_i}{1.64L_p} = 1.22\pi(\frac{L_i}{L_p})$. If sinusoidal current distribution is assumed, the departure from the end by a distance L_i will increase the current by $\cos(1.22\pi\frac{L_i}{L_p})$. At the same time the voltage will also decrease in magnitude by the same amount as that when the current increases [1].

The input impedance ($Z = \frac{V}{I}$) can be expressed as [1]:

$$Z_{in}(L_i) = \cos^2(1.22\pi\frac{L_i}{L_p})Z_{in}(0) \quad (1)$$

Where $Z_{in}(0)$ is the input impedance if the patch is fed at the end.

As an example let $L_i = 0.245L_p$, $\cos(1.22\pi\frac{L_i}{L_p}) = 0.59$ and $\cos^2(1.22\pi\frac{L_i}{L_p}) = 0.35$. Therefore a $0.245L_p$ inset feed would decrease the input impedance by 35%.

III. Performance comparison of edge fed and inset fed rectangular patches using different substrates

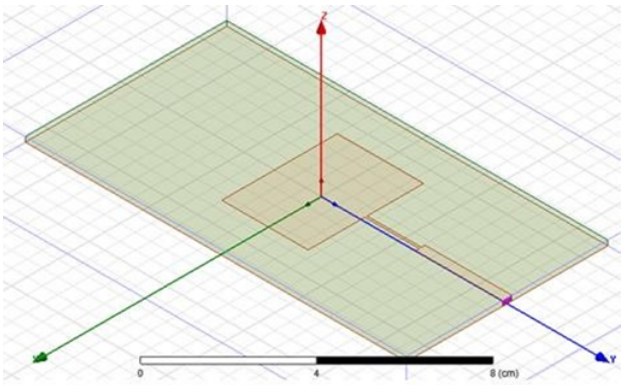
In this section, we have compared the performances of edge fed and inset fed rectangular MSA at the resonant frequency 2.45 GHz ($\lambda = 122.45mm$). Equations for the different dimension of patch antenna can be found in [5], [10]. With inset feed patch, a good matching condition can be obtained and a wide impedance bandwidth can be achieved [6]. We've used HFSS antenna design kit [11] to calculate the dimensions of patch and feed lines. We have considered two different substrates in the simulations, one is FR4 ($\epsilon_r = 4.4$) and the other one is air ($\epsilon_r = 1$). Thickness (h) of the substrate was kept constant at 1.6 mm in every scenario.

Different dimension of edge fed and inset fed patches (with FR4 and air substrates) are summarized in Table I. Use of substrate with lower relative permittivity results in lower impedance, better radiation pattern, increased antenna bandwidth, and higher antenna efficiency [1]. A higher value of permittivity should be used in the locations where no power is to be used radiated (e.g. transmission lines). Higher values of permittivity also result in 'shrinking' of the patch antenna which is good for cell phone use where a very little space is available for antenna placement. Fig.3 shows the snapshots of the patches from HFSS simulations. Fig.4 shows the return losses of the edge fed and inset fed rectangular patches with the FR-4 and air substrates. A summary of the simulation results is included in Table II.

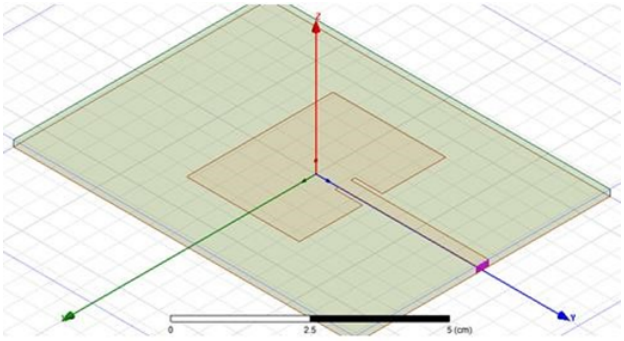
TABLE I: Dimensions (in mm) of edge fed and inset fed patches with FR-4 and air substrates.

Type of substrate ⇒		**FR-4 substrate** ($\epsilon_r = 4.4$)		**Air substrate** ($\epsilon_r = 1$)	
Feed-method ⇒		**Edge feed**	**Inset feed**	**Edge feed**	**Inset feed**
50 Ω feed line	Length	27.96	27.96	51.02	27.96
	Width	3.06	3.06	7.87	3.06
100 Ω feed line	Length	17.6	-	30.61	-
	Width	0.72	-	4.06	-
Patch	Length	27.85	28.7	57.5	58.9
	Width	37.3	37.3	61.2	61.2
Ground/ Substrate	Length	122.96	86.86	227.73	164.97
	Width	65	65	101	101
Inset cut	Length	-	7	-	15.8
	Width	-	1	-	1

It can be observed from Fig.4 and Table II that the inset fed MSA performs better than the edge fed one from all the perspectives of impedance bandwidth, directivity, gain, and return loss. The results with air substrate are also better than the corresponding results with FR-4 substrate. The gain of the inset feed with air substrate was 7.18 dB. In [1] the gain of a rectangular patch with air substrate was estimated within the range of 7∼9 dB. If the effective permittivity increases by a factor of 4, the required length decreases by a factor of 2 [1]. One disadvantage of using the air substrate is the larger antenna size that can be observed from Table I. Generally the higher value of permittivity is used to miniaturize the patch [1]. During the simulations, the length of the inset fed patch had to be increased slightly in order to keep the constant resonant frequency of 2.45 GHz. The aspect ratio (≈ Length/Width) of the patch controls the resonant frequency [1].



(a) Edge fed microstrip patch.



(b) Inset fed microstrip patch.

Fig. 3: Edge fed and inset fed microstrip patches from HFSS simulations (substrate: FR-4).

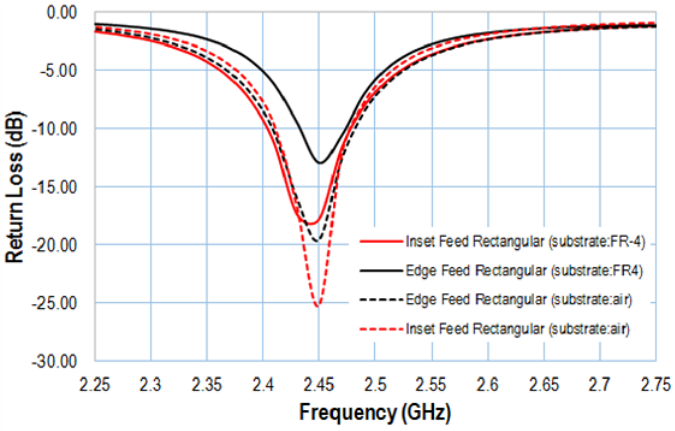


Fig. 4: Return loss characteristics of edge fed and inset fed microstrip patch antennas with FR-4 and air substrates (resonant frequency: 2.45 GHz).

IV. Performance comparison of 4×1 array antenna with edge fed and inset fed rectangular patches

Array antennas are used in order to enhance the gain and reduce the beamwidth. Generally the larger the number of antenna elements used, the higher the gain and the narrower the beamwidth can be achieved. A power efficient MSAA with minimum side lobe level (SLL) and narrower beamwidth will be required in next generation wireless systems. Some existing MSAA applications have drawbacks of larger beamwidth, higher SLL and lower gain. For example, Traffic Collision Avoidance System (TCAS) requires compact MSAA with higher gain, narrower beamwidth, and perfect directional radiation pattern [12]. MSAA should optimize multiple criteria simultaneously such as beam forming gain with very narrow beamwidth, lower side lobe level [13], less interference [14], polarization, power efficiency, and required bandwidth.

We have designed 4×1 MSAA with two different feeding techniques mentioned in the previous sections and with FR-4 substrate . We have used quarter-wave transformer with the feed systems for impedance matching purposes. Fig. 5 shows 4×1 MSAA with inset feed technique. Inter-element spacing of the array was 0.5λ in this case. Simulation results of the two MSAA are summarized in Table II. It can be seen from Table II that the values of all the four parameters (return loss, directivity, gain, and bandwidth) have been improved in case of MSAA in comparison to those of single patch antenna cases. Though the return loss is the lowest in case of inset fed MSAA (Fig. 8), other parameters are comparatively inferior to those of edge fed MSAA. Fig. 6 shows the directivities and 3D radiation patterns of two different MSAA.

V. Development of inset fed rectangular patch with copper material and air substrate

Our research group is in the process of developing and testing inset fed MSA/MSAA which resonate at 2.45 GHz and other ISM bands. Fig.7 shows an inset fed rectangular MSA antenna which was fabricated from copper plate and air was used as substrate between the patch and ground plate. We have chosen the copper material since it is available in the local market. We have selected the air as a substrate in order to check the performance of the inset fed patch with air. The results of the antenna testing will be included in future publications.

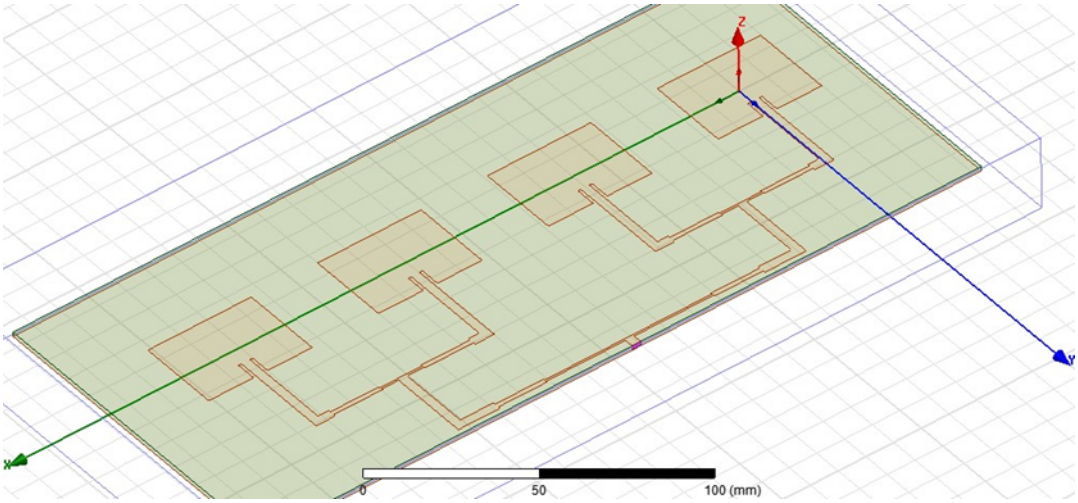
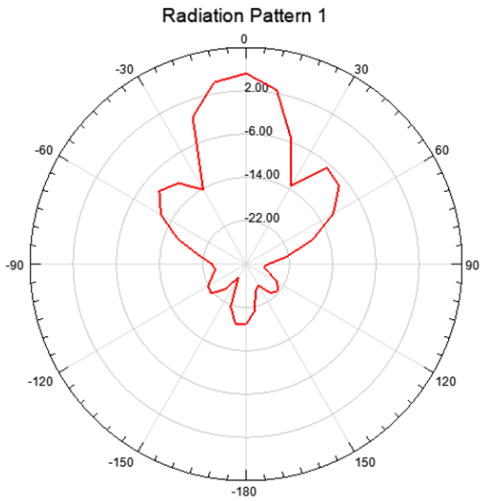
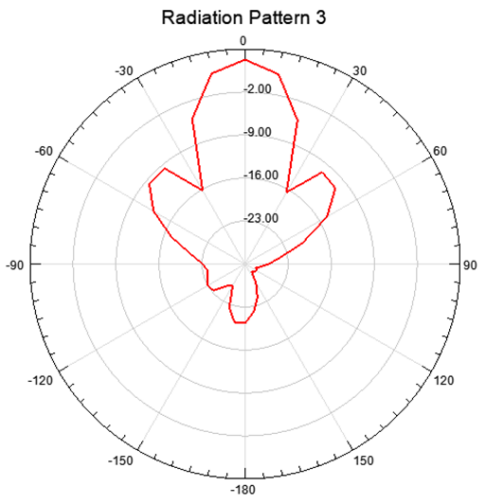


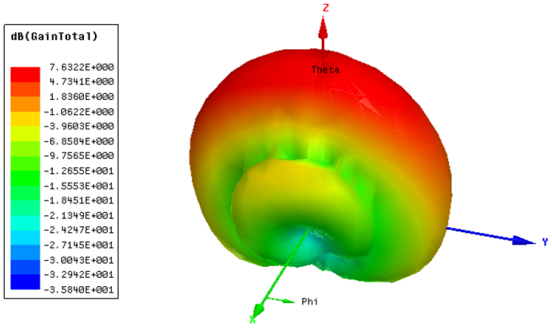
Fig. 5: 4×1 inset fed microstrip patch array from HFSS simulations (substrate: FR-4).



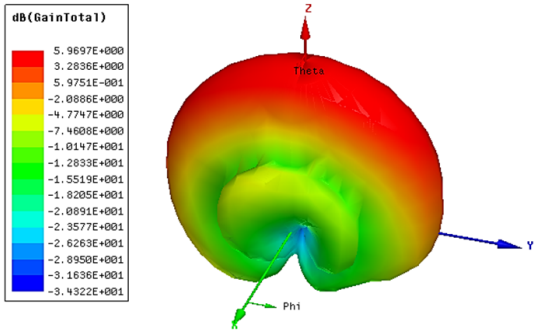
(a) Directivity of edge fed MSAA.



(b) Directivity of inset fed MSAA.

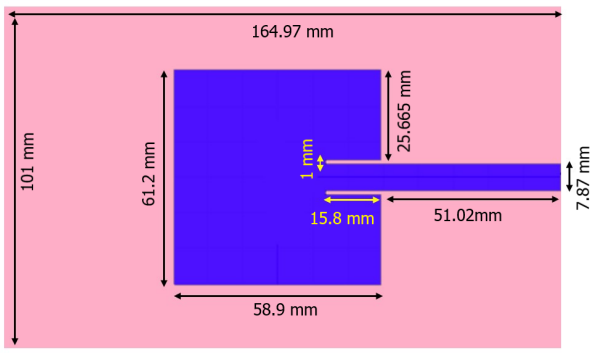


(c) Radiation pattern of edge fed MSAA.

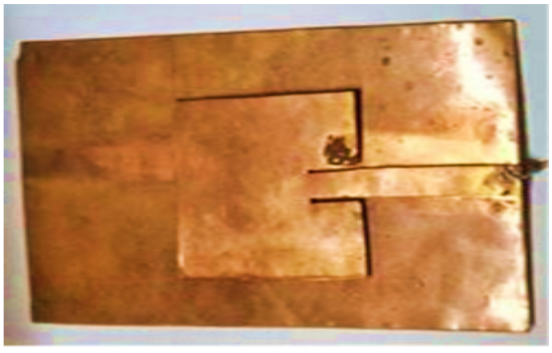


(d) Radiation pattern of inset fed MSAA.

Fig. 6: Directivities and 3D radiation patterns of 4×1 microstrip patch arrays with edge feed and inset feed techniques (substrate:FR-4).



(a) Dimensions of inset fed MSA with air substrate.



(b) Fabricated inset fed MSA with copper material and air substrate.

Fig. 7: Fabricated inset fed microstrip patch with air substrate.

TABLE II: Summary of performance-parameters of edge fed and inset fed rectangular patches at 2.45 GHz.

Antenna type ⇒	**Single MSA**				**4×1 MSAA**	
Substrate type ⇒	**FR-4**		**Air**		**FR-4**	
	Edge feed	**Inset feed**	**Edge feed**	**Inset feed**	**Edge feed**	**Inset feed**
Return Loss (dB)	-12.97	-17.80	-19.60	-25.28	-18.28	-26.00
Directivity (dB)	5.41	5.62	6.6758	7.1913	11.032	9.876
Gain (dB)	2.306	2.5	6.6756	7.1822	7.63	5.97
Bandwidth (MHz)	40	80	70	70	110	90

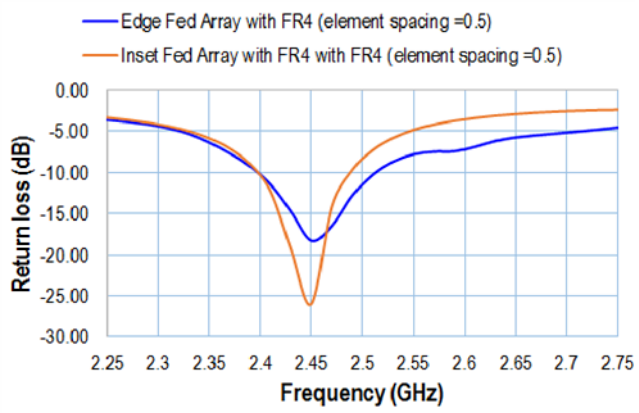


Fig. 8: Return losses of edge fed and inset fed microstrip arrays (4×1) with FR-4 substrate (resonant frequency 2.45 GHz).

VI. Conclusions

In this paper a comparative analysis of different characteristics of edge fed and inset fed rectangular microstrip patches are done at 2.45 GHz. The investigation results, which have been described in the paper, revealed that the better impedance matching can be made through inset feed of rectangular array without any complex modification of microwave circuits. Inset fed rectangular patch antenna performed better than the edge fed patch when impedance bandwidths, directivities, gains and return losses were compared. Though research findings on single patch can be found in different publications, research investigations on edge fed and inset fed rectangular array have not yet been widely published. It has been seen through HFSS simulations that return loss in case of inset fed array, which is -26 dB, is better than that of edge fed array. Return loss in case of edge fed array was -18.28 dB. Better return loss of inset fed array was obtained through optimization of the dimensions of inset fed patch. The investigation results reported in this paper would be helpful in implementation of inset fed rectangular microstrip antennas in future wireless applications.

References

[1] John D Kraus, Ronald J Marhefka, Ahmad S Khan, 'Antenna and Wave Propagation', Tata McGraw Hill Education Pvt. Ltd., 2010.

[2] Vitaly Petrov et al., "Interference and SINR in Millimeter Wave and Terahertz Communication Systems With Blocking and Directional Antennas", IEEE Trans. Wireless Commun. vol. 16, pp. 1791-1808, March 2017.

[3] Antonio Morgado et al. 'A survey of 5G technologies: regulatory, standardization and industrial perspectives', Elsevier Digital Communications and Networks, issue 4, pp. 87–97, 2018.

[4] Simon R Saunders, 'Antennas and Propagation for Wireless Communication Systems', John Wiley and Sons Ltd., 2007.

[5] Constantine A. Balanis, 'Antenna Theory: Analysis and Design', 3rd Edition, John Wiley, 2005.

[6] Kin Lu Wong, Compact and Broadband Microstrip Antennas, John Wiley & Sons Inc., 2002.

[7] K. F. Lee, Principles of Antenna Theory, John Wiley & Sons Ltd, 1984.

[8] W. L. Stutzman and G. A. Thiele, Antenna theory and design, John Wiley & Sons Ltd, 1981.

[9] G. S. Raghuvanshi, 'Microwave Engineering', CENGAGE Learning, 2012.

[10] Sulistyaningsih, Prasetyo Putranto, Mashury W, "Microstrip Patch Array Antenna with Inset Fed and Perturbation for a 3 GHz S-Band Coastal Radar", International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET), October 2017.

[11] HFSS Antenna Design Kit, 'https://www.ansys.com/campaigns/ansys-hfss-for-antenna- design'.

[12] Debji De, Prasanna Kumar Sahu, 'A novel Approach towards the Designing of an Antenna for Aircraft Collision Avoidance System', Int. J.Electron Commun. (AEÜ) , issue 71, pp. 53-71, 2017.

[13] Wei Hong et al., "Multibeam Antenna Technologies for 5G Wireless Communications," IEEE Trans. Antennas Propagat., vol. 65, pp. 6231–6249, Dec. 2017.

[14] Vasanthan Raghavan et al., "Millimeter Wave Channel Measurements and Implications for PHY Layer Design," IEEE Trans. Antennas Propagat., vol. 65, pp. 6521–6533, Dec. 2017.

Ultra High Efficient 2x1 Graphene Patch Antenna Arrays for Single and Dual Band Operation

Md. Abdul Kaium Khan*, Md Ibrahim Ullah, Towqir Ahmed Shaem, Rifat Kabir, Mohammad Abdul Alim
Department of Electrical & Electronic Engineering
University of Chittagong, Chittagong-4331, Bangladesh
abdulkaiumkhan7@gmail.com*, iuemon96@gmail.com, towqirahmedshaem@gmail.com, bejoykabir@gmail.com, mohammadabdulalim@cu.ac.bd

Abstract— **Graphene has opened the floodgate for a plethora of new applications in the field of communication due to its outstanding electrical and optical properties. In this paper, three types of Graphene antenna arrays, i.e. rectangular, circular and triangular are designed to operate around 3THz on a silicon (ϵ_r = 11.9) substrate. A single band to dual band operation is achieved for all three types of Graphene antenna arrays at THz frequency range by changing the height of the silicon substrate, which has not been reported before. In comparison to rectangular and circular antenna arrays, triangular antenna array exhibits better performance in terms of S_{11}, VSWR, BW and Radiation Efficiency with an S_{11} of -39.44dB, VSWR of 1.024, BW of 103.4GHz and a staggering efficiency of 99.11%, which is the highest Radiation Efficiency, reported so far for Graphene antenna array application.**

Keywords—graphene; array; antenna

I. INTRODUCTION

Over the last decades, wired communication mediums are reaching their limits and according to Edholm's law, the wireless data rates have doubled in every 18 months [1]. Future generation communication systems will need to have data rates as high as 100 Gbit s^{-1} or more. With the increasing demand of high-speed data transmission, exploitation of the new higher band of frequencies is becoming a must. For this purpose, Terahertz communication has been considered a decisive wireless technology. Its spectral band lies in between the microwave and mid infrared with a span of frequencies between 0.1-10 Terahertz [2].

The main constraint of Terahertz communication is its high path loss due to atmospheric molecular absorption, which restricts its distance of communication. That is why Terahertz communication is applicable to LAN and PAN networks. In addition to this, implementation of Terahertz transceiver and efficient antennas remains challenging. GaN based HEMTs have already shown f_c >450GHz, f_{max} >600GHz [3] and InP based HEMTs have shown f_c >600GHz, f_{max} >1.2THz [4]. This transistor technology is already used to build MMIC amplifiers, which are reported to exhibit data rate as high as 25 Gbit s^{-1} [5]. The implementation of sources, detectors and modulators operating at Terahertz band will enable Terahertz communication feasible within few years.

Typical metallic antennas of micrometer sizes tend to resonate at optical range of frequencies of several hundreds of Terahertz [6]. These antennas suffer from a significant amount of shortcomings such as signal and channel attenuation due to their low carrier mobility at such high resonant frequencies.

Graphene based antennas can be a solution for this purpose. Graphene is an astonishing material with stunning electrical and optical properties. It is a 2D version of Graphite with carbon atoms arranged in a honeycomb lattice of single layer [7]. Most importantly, it can resonate at Terahertz frequency band [8]. Graphene possesses extremely high carrier mobility in the range of 8000-10000 $cm^2 V^{-1} s^{-1}$, which may reach as much as 200,000 $cm^2 V^{-1} s^{-1}$ at room temperature, and a Young's modulus of 1.5 TPa, which makes graphene one of the strongest materials with immense carrier velocity [2]. This astonishingly high carrier mobility enables Graphene based devices to operate at Terahertz band. The amazing thing about Graphene is that, it is tunable by changing its chemical potential and relaxation time and it possesses better radiation efficiency than typical metallic antennas [9]. It also supports SPP wave propagation, which makes it a precious material in the research field of communication. In the terahertz band, plasmonic nature of the graphene causes extreme solidity, where metals behave as non-plasmonic conductors. Graphene makes the creation of dynamically tunable devices possible [10].

In recent times, there have been some works on Graphene based rectangular, triangular and circular patch antennas at Terahertz range [11-13]. Works are also done on Graphene antenna arrays. In [14], multiple resonance is achieved by changing the number of Graphene layers for UT shaped Graphene antenna array. In [15], rectangular patch Graphene antenna array for 5G application is investigated. Previously, Graphene based dual band antennas are realized by creating slots and stubs near the radiating edges and using backing cavity [16].

In this paper, we are presenting Graphene based triangular, rectangular and circular 2x1 patch antenna arrays with Graphene as a radiating patch on a silicon substrate (ϵ_r = 11.9) operating around 3THz. We achieved a single band to dual band conversion for changing the substrate height of each type of 2x1 patch antenna arrays at Terahertz range and compared the results. Achieving dual band operation by changing the substrate height of different types of Graphene antenna arrays at Terahertz band haven't been investigated before and it provides an easier way of achieving dual band operation at micrometer size than creating slots and this signifies the novelty of this work. Previously, the highest radiation efficiency reported for Graphene antenna array application was 93.14% for rectangular Graphene patch on a Kapton substrate [15]. We have achieved higher radiation efficiency of 99.11% for triangular 2x1 patch antenna array for silicon (ϵ_r = 11.9) substrate which is also a novelty of this work.

II. GRAPHENE MODELLING

We used Computer Simulation Software (CST) Microwave Studio for our 2x1 Graphene antenna array design. CST provides Graphene material for optical applications. In [9], Llatser et al. showed that the resonance of Graphene depends heavily on its chemical potential (μ_c) and relaxation time (τ). The conductivity of Graphene is defined as,

$$\sigma(\omega) = \frac{2e^2 K_B T i}{\pi\hbar^2\left(\omega + i\tau^{-1}\right)} \ln\left[2\cosh\left[\frac{\mu_C}{2K_B T}\right]\right] \quad (1)$$

Here, the electron charge is **e**, the temperature is **T**, the Boltzmann constant is $\mathbf{K_B}$, the chemical potential is $\boldsymbol{\mu_c}$ and the relaxation time is $\boldsymbol{\tau}$. From Equation (1), we can see that chemical potential has a direct relation to the conductivity of Graphene hence a change of chemical potential will change the conductivity of Graphene. The increase of chemical potential increases the resonant frequency and increase in relaxation time produces strong resonant characteristics. Chemical doping is used for changing chemical potential of Graphene. Our Graphene parameters are: Temperature (T) = 300K, Chemical Potential (μ_c) = 0eV, Relaxation Time (τ) = 2ps and Graphene sheet thickness = 0.1um.

III. RECTANGULAR PATCH ARRAY DESIGN

Our rectangular antenna is designed to resonate around 3THz. On the basis of cavity model, the resonant frequency of a rectangular antenna of length L is given by [17] as,

$$f_{10} = \frac{c}{2L\sqrt{\varepsilon_r}} \quad (2)$$

Where c is the light speed and ϵ_r is substrate's dielectric constant. The width W of the antenna has a minor contribution to the resonant frequency. The geometry of the antenna array is given in Fig. 1 and the dimensions are given in Table I.

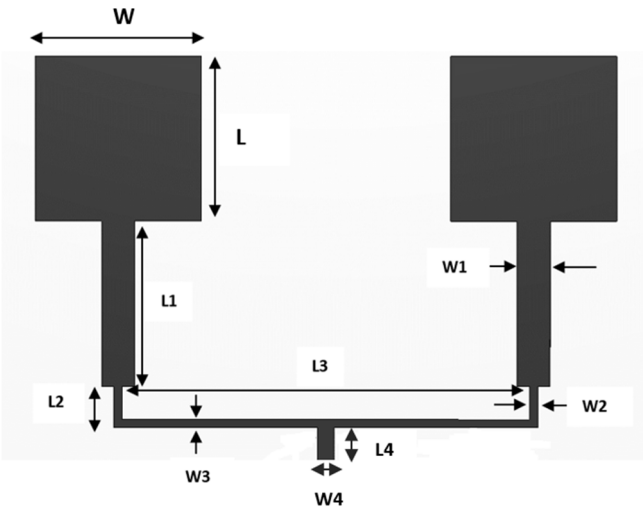


Figure 1. Geometry of 2x1 rectangular Graphene antenna array designed in CST

TABLE I. DIMENSIONS OF PATCH ANTENNA ARRAYS

Dimensions	Rectangular (μm)	Circular (μm)	Triangular (μm)
Width*Length(W*L) / Radius(a) / Side Length(a)	10*10	5	10
Patch height	0.1	0.1	0.1
W1	2	2	2
L1	10	10	10
W2	0.5	0.4	0.5
L2	2.05	2	2.5
W3	0.5	0.5	0.5
L3	25.5	25.5	25.5
W4	1	1	1
L4	2	2	2

IV. CIRCULAR PATCH ARRAY DESIGN

Our circular antenna is designed to resonate around 3THz. The resonant frequency calculation is given by [18] as,

$$f_r = \frac{1}{2\pi\sqrt{\mu\varepsilon}}\left(\frac{X'_{mn}}{a}\right) \quad (3)$$

Where ***a*** is the radius of the patch. For the lowest order mode, Bessel function zeros X'_{mn} becomes 1.8414 and equation (3) becomes,

$$f_r = \frac{1.8411c}{2\pi a\sqrt{\varepsilon_r}} \quad (4)$$

The geometry of the antenna array is given in Fig. 2 and the dimensions are given in Table I.

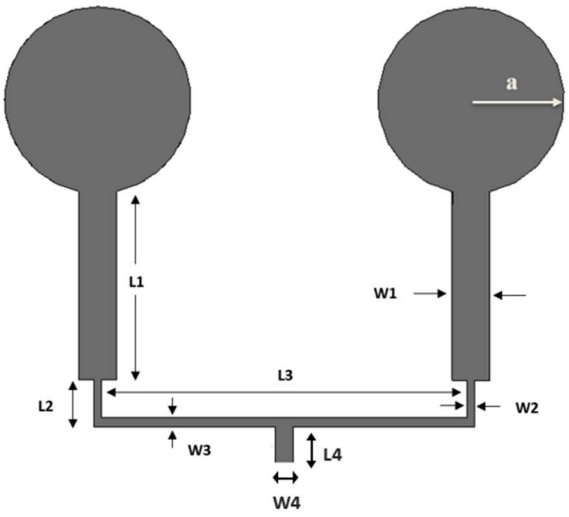


Figure 2. Geometry of 2x1 circular Graphene antenna array designed in CST

V. TRIANGULAR PATCH ARRAY DESIGN

Our triangular antenna is designed to resonate around 3THz. The geometry of the antenna array is given in Fig. 3 and the dimensions are given in Table I.

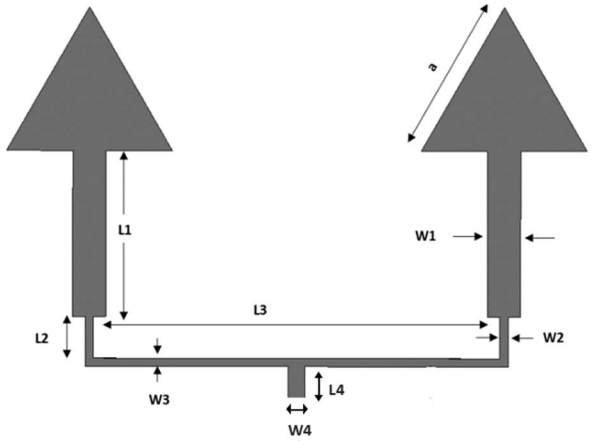


Figure 3. Geometry of 2x1 triangular Graphene antenna array designed in CST

The resonant frequency calculation for an equilateral antenna of the lowest mode is given by [19] as,

$$f_{10} = \frac{2c}{3a_e\sqrt{\varepsilon_r}} \qquad (5)$$

Where a_e is the effective side length of the antenna and is given as,

$$a_e = a\left[1+2.199\left(\frac{h}{a}\right)-12.853\left(\frac{1}{\sqrt{\varepsilon_r}}\right)\left(\frac{h}{a}\right)+16.434\left(\frac{1}{\varepsilon_r}\right)\left(\frac{h}{a}\right)+6.182\left(\frac{h}{a}\right)^2-9.802\left(\frac{1}{\sqrt{\varepsilon_r}}\right)\left(\frac{h}{a}\right)^2\right] \qquad (6)$$

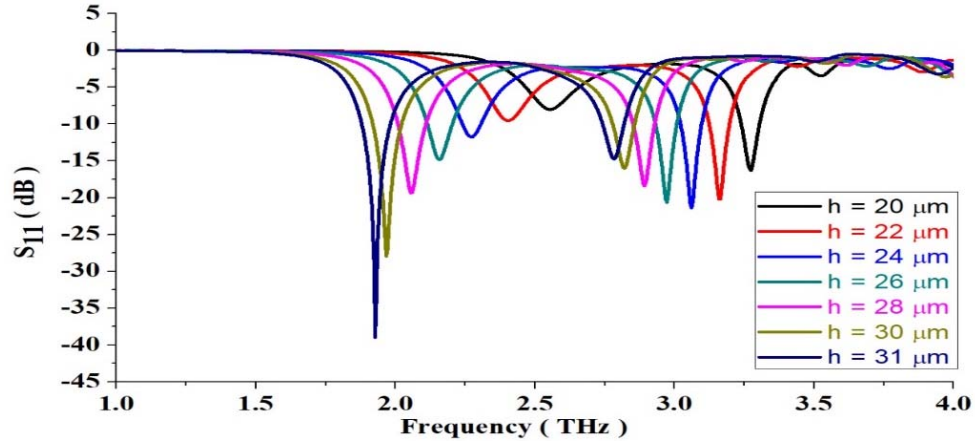
VI. SIMULATION & RESULTS

The complete simulation process is carried out by CST microwave studio, which supports EM field solver. We used microstrip line feeding to excite the antenna arrays with discrete port as exciter and we used quarter wave (λ/4) transformers for matching the input impedance of 50Ω to the patch antenna's load impedance. We simulated each of the antenna arrays and analyzed different antenna parameters, e.g. Return Loss S_{11}, Voltage Standing Wave Ratio (VSWR), Bandwidth, Radiation efficiency, Directivity and Gain for different height of the Silicon substrate.

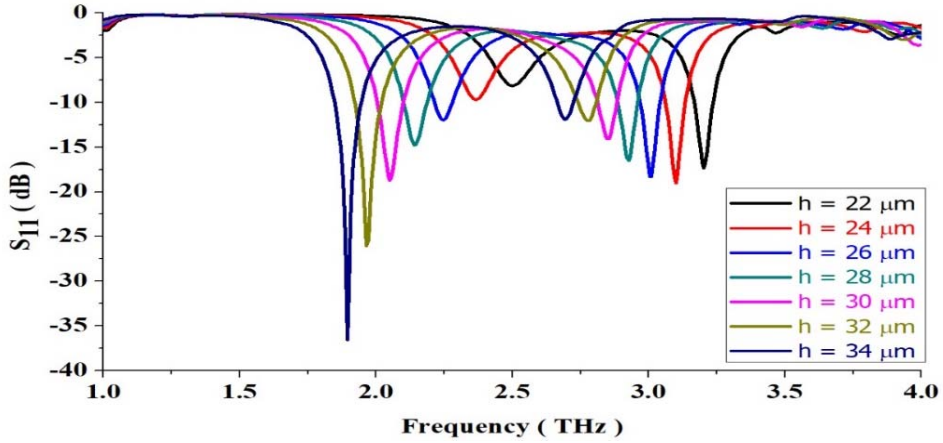
For antennas, the Return Loss S_{11} should be less than -10dB to be considered as workable frequency band. The VSWR should be close to 1 for perfect impedance matching. The Radiation Efficiency should be as high as possible (less negative in dB), where 0dB means perfect antenna with 100% efficiency. The higher the values of Directivity, BW and Gain the better the antenna performance. Table II shows the antenna performance for each antenna array with different values of the substrate height.

TABLE II. PERFORMANCE OF GRAPHENE ANTENNA ARRAYS

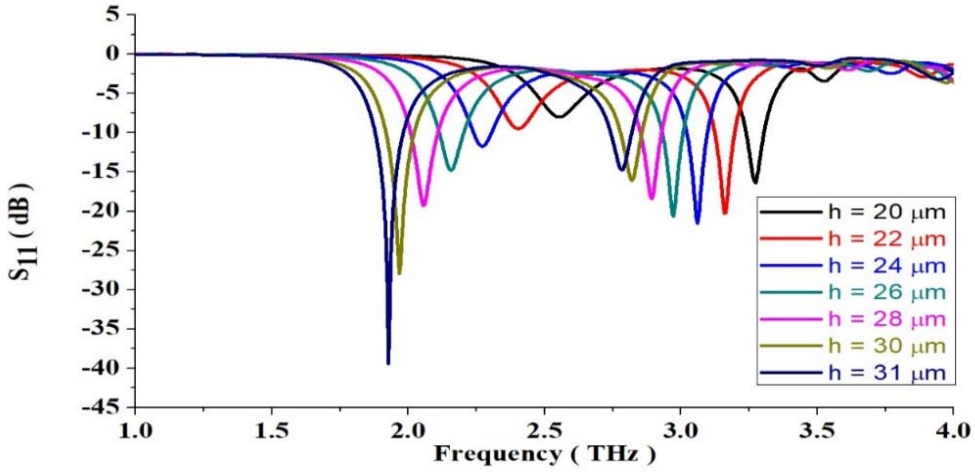
Array Type	Substrate Height-h(μm)	Resonant Frequency-f_c (THz)		Return Loss-S_{11}(dB)		VSWR		Bandwidth (GHz)		Radiation Efficiency(dB)		Directivity(dB)		Gain(dB)	
		Band 1	Band 2	Band 1	Band 2	Band 1	Band 2	Band 1	Band 2	Band 1	Band 2	Band 1	Band 2	Band 1	Band 2
Rectangular	20	3.28	-	-16.24	-	1.3626	-	65.4	-	-0.05617	-	5.392	-	5.335	-
	22	3.168	-	-20.19	-	1.2168	-	72.2	-	-0.0614	-	5.448	-	5.386	-
	24	2.272	3.06	-11.75	-21.41	1.696	1.185	66.25	75.78	-0.038	-0.0836	2.651	5.307	2.613	5.223
	26	2.154	2.968	-14.82	-20.60	1.442	1.205	89.4	76.3	-0.0427	-0.0457	2.599	5.136	2.556	5.090
	28	2.067	2.91	-19.33	-18.33	1.241	1.274	99.82	77.49	-0.0445	-0.0459	2.573	5.090	2.529	5.044
	30	1.97	2.827	-27.97	-16.03	1.0831	1.374	103.4	78.38	-0.0663	-0.0526	2.601	5.151	2.535	5.099
	31	1.924	2.77	-38.98	-14.74	1.022	1.448	102.8	76.84	-0.0574	-0.05601	2.598	5.328	2.540	5.182
Circular	22	3.202	-	-17.34	-	1.3143	-	70.1	-	-0.3633	-	5.628	-	5.265	-
	24	3.1	-	-18.99	-	1.2532	-	72.69	-	-0.3485	-	5.548	-	5.199	-
	26	2.248	3.004	-11.97	-18.28	1.6746	1.276	64.71	73.4	-0.4182	-0.3434	2.854	5.389	2.436	5.045
	28	2.140	2.926	-14.76	-16.51	1.4474	1.351	87.17	73.2	-0.465	-0.3766	2.850	5.451	2.385	4.999
	30	2.05	2.85	-18.74	-14.06	1.2636	1.494	95.95	69.80	-0.4852	-0.4272	2.815	5.457	2.330	5.029
	32	1.966	2.776	-26.03	-12.06	1.1051	1.665	97.82	67.30	-0.4875	-0.5487	2.867	5.616	2.380	5.067
	34	1.894	2.692	-36.57	-11.88	1.0301	1.684	98.48	59.68	-0.5330	-0.4467	2.916	5.515	2.383	5.069
Triangular	20	3.272	-	-16.34	-	1.3595	-	65.67	-	-0.0545	-	5.389	-	5.335	-
	22	3.16	-	-20.26	-	1.2148	-	72.81	-	-0.0596	-	5.448	-	5.388	-
	24	2.27	3.04	-11.72	-21.51	1.699	1.183	65.88	75.83	-0.0369	-0.0832	2.655	5.308	2.618	5.225
	26	2.156	2.972	-14.80	-20.65	1.444	1.204	89.22	76.31	-0.0427	-0.0445	2.599	5.134	2.556	5.090
	28	2.056	2.892	-19.27	-18.42	1.243	1.272	100.2	77.3	-0.0404	-0.0388	2.594	5.100	2.553	5.062
	30	1.968	2.82	-27.91	-16.07	1.083	1.37	103.4	78.2	-0.0649	-0.0521	2.6	5.150	2.535	5.098
	31	1.928	2.784	-39.44	-14.76	1.021	1.44	102.5	77.08	-0.0557	-0.0554	2.596	5.236	2.540	5.181



(a)



(b)



(c)

Figure 4. S_{11} of (a) Rectangular (b) Circular (c) Triangular 2x1 Graphene antenna array

From table II and Fig. 4 we can see that for each type of antenna array, as we increase the substrate height, second band arises from a single band operation. As the substrate height increases, it creates enough space for the fringing field to excite another higher order mode to be propagated. The fringing-filed effect of antenna is related to patch dimension and substrate height [20]. The new order mode resonates alike the first mode but at a separate frequency because of distinct phase constants. Thus, a single to dual band conversion is achieved for all three-antenna arrays.

The highest single band gain is 5.388dB for triangular antenna array operating at 3.16THz with substrate height h=22 µm. The highest dual band gain is exhibited by rectangular array with h=24µm. The second band operates at 3.06THz with a Gain of 5.223dB. Fig. 5 shows the 3D & 2D radiation pattern for rectangular array with h=24µm. The best S_{11} and VSWR exhibited by triangular array with h=31µm. The first band operates at 1.928THz with a very good S_{11} of -39.44dB and VSWR of 1.021. The fringing field increases the BW of the antenna and we can see the proof from the results in table II. The highest single band BW is 72.81GHz for triangular array with h=22µm. For dual band operation, the highest BW is 103.4GHz for triangular array with h=30µm operating at 1.968THz. The highest Radiation efficiency is exhibited by triangular array with h=28µm. The second band operates at 2.892THz with an astonishing Radiation efficiency of -0.0388dB or 99.11% and a Gain of 5.062dB. Fig. 6 shows the 3D & 2D radiation pattern for triangular array with h=28µm. The amazing thing is that the lowest Radiation efficiency for triangular Graphene antenna array is -0.0832dB or 98.1%. The lowest Radiation efficiency for rectangular Graphene antenna array is -0.0836dB or 98.09% and in case of circular array the lowest Radiation efficiency is -0.5487dB or 88.1%. So, we can see that, all the three types of Graphene antenna arrays are highly efficient.

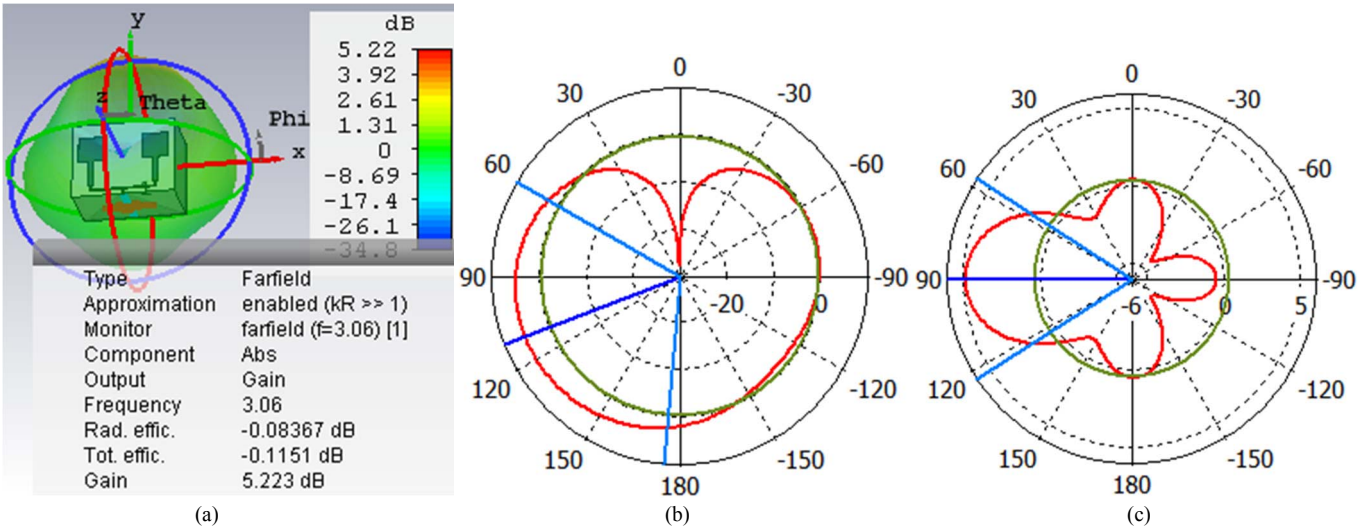


Figure 5. Radiation pattern of 2x1 rectangular array for h=24µm (a) 3D (b) E-plane (Y-Z plane) with Φ=90° (c) H- plane (X-Y plane) with Θ=90°

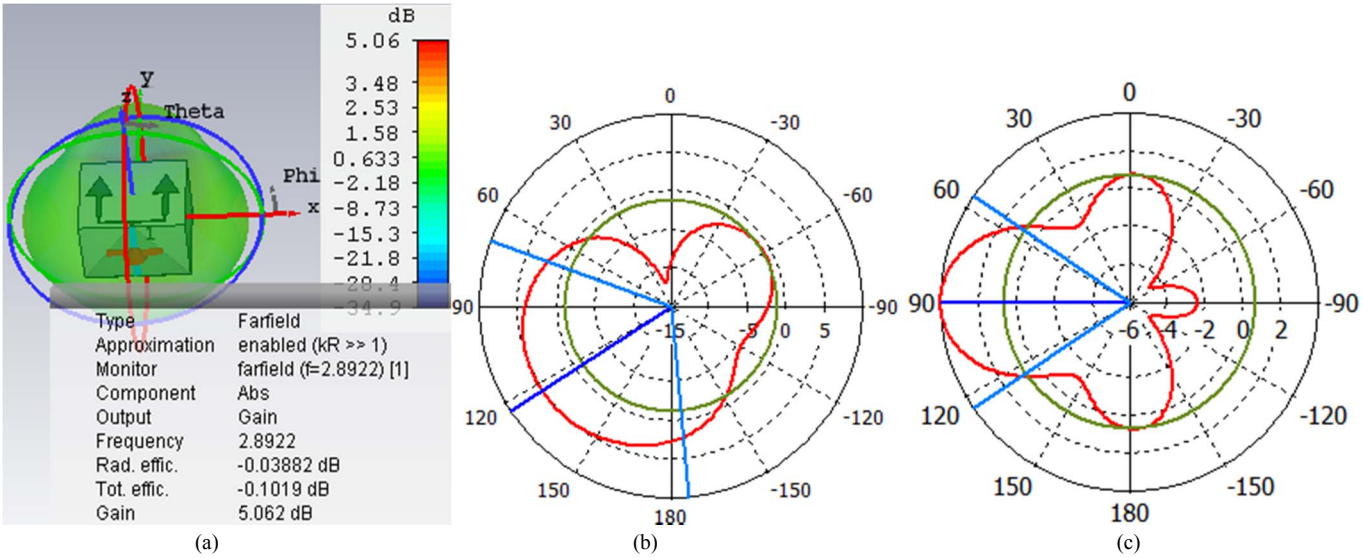
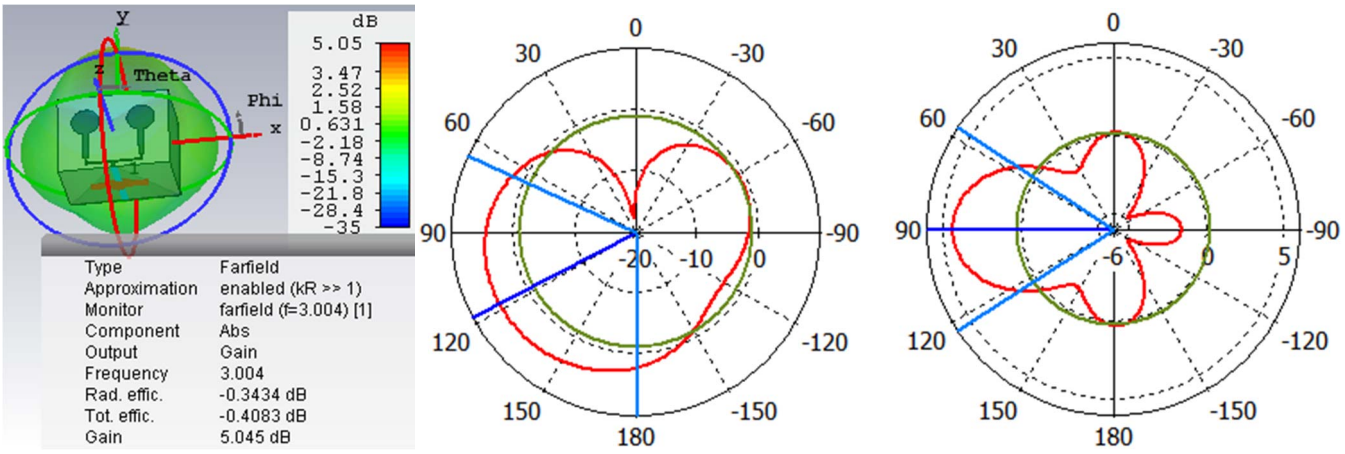


Figure 6. Radiation pattern of 2x1 triangular array for h=28µm (a) 3D (b) E-plane (Y-Z plane) with Φ=90° (c) H- plane (X-Y plane) with Θ=90°



(a) (b) (c)

Figure 7. Radiation pattern of 2x1 circular array for h=26µm (a) 3D (b) E-plane (Y-Z plane) with Φ=90° (c) H- plane (X-Y plane) with Θ=90°

VII. CONCLUSION

In this paper, we presented an approach for gaining dual band operation from a single band operation by increasing the height of the silicon substrate for three types of antenna arrays. We have seen from the results that for all the three types of Graphene antenna arrays, i.e. rectangular, circular, triangular a single band to dual band conversion is achieved. In comparison to other two types of antenna arrays, the triangular antenna array shows better performance in case of Radiation efficiency, S_{11}, VSWR and BW. A staggering 99.11% antenna efficiency is achieved for triangular antenna array operating at 2.892THz for substrate height h=28µm. Circular antenna array exhibits Radiation efficiency greater than 88.1% for different substrate heights and rectangular antenna array exhibits Radiation efficiency greater than 98.09% for different substrate heights. All the three antenna arrays are highly efficient.

REFERENCES

[1] S. Cherry, Edholm's law of bandwidth, IEEE Spectr. 41 (7) (2004) 58–60.

[2] I. F. Akyildiz, J. M. Jornet, C. Han, "Terahertz band: Next frontier for wireles communications," Physical Communication 12 (2014) 16–32.

[3] K. Shinohara, D. Regan, Y. Tang, A. Corrion, D. Brown, J. Wong, J. Robinson, H. Fung, A. Schmitz, T. Oh, S. Kim, P. Chen, R. Nagele, A. Margomenos, M. Micovic, Scaling of gan HEMTs and schottky diodes for submillimeter-wave mmic applications, IEEE Trans. Electron. Devices 60 (10) (2013) 2982–2996

[4] V. Radisic, K. Leong, X. Mei, S. Sarkozy, W. Yoshida, W. Deal, Power amplification at 0.65THz using inp HEMTs, IEEE Trans. Microw. Theory Tech. 60 (3) (2012) 724–729.

[5] I. Kallfass, J. Antes, D. Lopez-Diaz, S. Wagner, A. Tessmann, A. Leuther, Broadband active integrated circuits for terahertz communication, in: Proc. of 18th European Wireless Conference European Wireless, EW, 2012, pp. 1–5.

[6] S. Abadal, I. Llatser, A. Mestres, H. Lee, E. Alarc´on, A. Cabellosaparicio, "Time-Domain Analysis of Graphene-based Miniaturized Antennas for Ultra-short-range Impulse RadioCommunications," IEEE Transactions on Communications 63 (4) (2015)1470–82.

[7] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, A. K. Geim, "The electronic properties of graphene," Reviews of Modern Physics 81 (1) (2009) 109–162.

[8] J. M. Jornet and I. F. Akyildiz, ''Graphene-based nano-antennas for electromagnetic nanocommunications in the terahertz band,'' in *Proceedings of the EuCAP '10*, 2010, pp. 1-5.

[9] I. Llatser, C. Kremers, D. Chigrin, J. M. Jornet, M. C. Lemme, A. Cabellos-Aparicio, and E. Alarc´on, ''Radiation Characteristics of Tunable Graphennas in the Terahertz Band,'' *Radioengineering Journal*, 2012, vol. 21, no. 4.

[10] Sergi Abadal, Seyed Ehsan Hosseininejad, Albert Cabellos-Aparicio, Eduard Alarcón, "Graphene-Based terahertz antennas for area-constrained applications", 40th International Conference on Telecommunications and Signal Processing (TSP) – 2017, DOI: 10.1109/TSP.2017.8076102

[11] R. Bala, A. Marwaha, "Investigation of Graphene-based miniaturized terahertz antenna for novel substrate materials", Eng. Sci. Technol. Int. J., Mar. 2016, vol. 19, no. 1, pp. 531-537.

[12] Rajni Bala, A. Marwaha, "Development of Computational Model for Tunable Characteristics of Graphene-based Triangular Patch Antenna in Terahertz Regime" Springer on Journal of Computational Electronics, ISSN- 1569-8025, indexed by SCI , Thomson Reuter journal list, Impact Factor 1.520, DOI 10.1007/s10825-015-0761-6, Print ISSN 1569-8025, Online ISSN 1572-8137, pp 1.6, online 2015

[13] Sakibul Azam, Md. Abdul Kaium Khan, Towqir Ahmed Shaem, Abdullah Zowad Khan, "Graphene-based circular patch terahertz antenna using novel substrate materials", Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT), DOI: 10.1109/ICIEV.2017.8338605

[14] Ekşioğlu, Y., Cetin, A.E. & Durmaz, H. Plasmonics (2018) 13: 1081. https://doi.org/10.1007/s11468-017-0607-0.

[15] Sa'don, S.N.H., Kamarudin, M.R., Ahmad, F. et al. Appl. Phys. A (2017) 123: 118. https://doi.org/10.1007/s00339-016-0749-5

[16] Yanfei Dong, et al., Dual-band reconfigurable terahertz patch antenna withgraphene-stack-based backing cavity, IEEE Antennas Wirel. Propag. Lett. 155(2016).

[17] Garg, R., Bhartia, P., Bahl, I. J., & Ittipiboon, P. (2001). Microstrip Antenna Design Handbook. Artech House, Boston, London.

[18] C. Balanis, *Antenna Theory*, Wiley Interscience, Hoboken, NJ 2005.

[19] Chen, W., K. F. Lee and J. S. Dahele, "Theoretical and Experimental Studies of the Resonant Frequencies of the Equilateral Triangular Microstrip Antenna, "IEEE Trans. On Antennas and Propagation, Vol. AP-40, 1992, pp. 1253-1256.

[20] K.P. Ray Girish Kumar, Broadband Microstrip Antennas, Artech House Inc, 2003 (Chapter 2).

Performance Analysis of Tx-Rx Isolated Distributed Antenna System Implementing In-band Full-duplex for Up-link Communication to Mitigate Self-interference in 5G

Ruhin Chowdhury[*,1], A. K. M. Sharoar Jahan Choyon[*,2]
[*]*Dept. of EECE, Military Institute of Science & Technology (MIST), Dhaka, Bangladesh*
[1]ruhio138@gmail.com , [2]choyonsharoar@gmail.com

***Abstract*—Full-duplex (FD) is significantly better than half duplex (HD) as it transmits and receives data signal simultaneously at the same frequency channel. But one of the major challenges in this technique is self-interference (SI). To our best knowledge, for the first time, we introduce a system model for 5th generation up-link communication exploiting distributed antenna system (DAS) to reduce SI to an extent and thus making the overall network performance better. This proposed DAS is modified version of the conventional one. It uses separate transmitter and receiver whereas traditional DAS uses transceiver which is highly exposed to SI. After comprehensive simulation of the proposed system, its performance is evaluated and compared with the existing DAS technology under different parameters and scenarios. The SINR performance of proposed system model appears to be better than the traditional DAS.**

***Index Terms*—Full-duplex; Distributed Antenna System; Self-Interference; Signal to Interference plus Noise Ratio; 5G.**

I. Introduction

Rapid growth of consumer demand for worldwide wireless communication has motivated the development of fifth generation mobile communication system. For meeting the demand of increasing users, in-band FD communication has been proposed as a new scheme in 5G which doubles the capacity compared to HD [1-2]. The main aspect of FD is the simultaneous utilization of transmission and reception at the same frequency by a single user increasing the transmission rate which will enable mobile operators to maximize the use of the limited frequency bandwidth they are given [3]. Although using FD, user capacity will be raised, a major drawback of this technology, SI degrades the overall network performance rendering this technology impractical. To suppress this additional deleterious signal, numerous methods have been proposed for SI cancellation [4-6].

Since wireless networks use transceiver for communication, the closely-spaced transmitter and receiver antennas makes the system suffer high magnitude of SI. This paper proposes a new model of DAS implementing FD by isolating the transmitter and receiver and therefore, reducing the effect of SI in the main signal. Analysis is presented and the proposed system model is compared with the traditional DAS using FD and path loss parameters standardized by 3rd Generation Partnership Project (3GPP) demonstrating the effect of SI on both models [7].

Traditional DAS has been designed to increase network capacity solving coverage problem by re-using existing infrastructure [8]. Although, DAS have been implemented originally to cover dead spots in indoor wireless communication systems, this system can also be utilized to expand system capacity with reduced transmit power providing a better coverage in cellular communication. Because of analytical simplicity, DAS is used for uplink performance [9-10]. It can be recognized through investigation that our proposed model produces better performance.

This paper is organized as follows. Section II illustrates an elaborate description of our proposed system model. Section III elucidates the path loss model as well as the systemic evaluation of SINR calculation for the proposed model. Section IV provides the performance analysis and compares it with traditional DAS. Finally, Section V concludes the paper by summing the analysis.

II. System Model

Fig.1 illustrates the proposed in-band FD system model for up-link communication using DAS. The proposed system model exploits a macro cell in a homogeneous network consisting of a receiver approximately at the center and multiple transmitters distributed within the receivers perimeter. This model presents separate transmitter and receiver where every transmitter and receiver works individually not as a transceiver since SI increases as the distance between the transmitter and receiver decreases. Furthermore, different cell uses different frequency band which are orthogonal to each other resulting in zero inter-cell interference. In this model a single user is placed randomly in the cell at a time considering no user to user interference. Here a User Terminal (UT) connects to the nearest transmitter which transmits signal and the receiver which receives signal, both working individually. Since same frequency is used for transmission and reception simultaneously, the receiver of DAS will receive the received signal of UT as well as a portion of the transmitted signal from the transmitter as SI. For a traditional DAS depicted in Fig. 2,

in case of implementing FD, interference degrades the SINR output resulting in a far worse signal strength and crosstalk as all the nodes work as transceivers. During a phone call, UT sends and receives signal from the same node resulting in major SI.

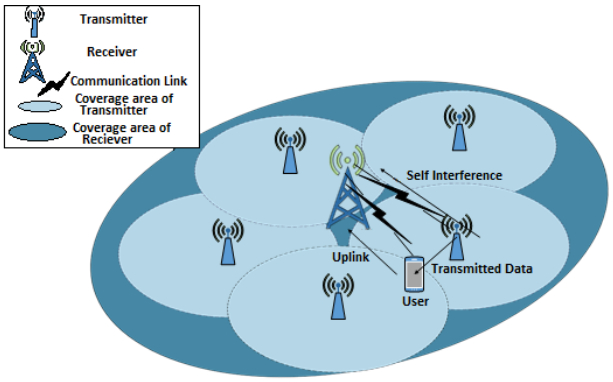


Fig. 1. Proposed model for full-duplex up-link communication exploiting DAS

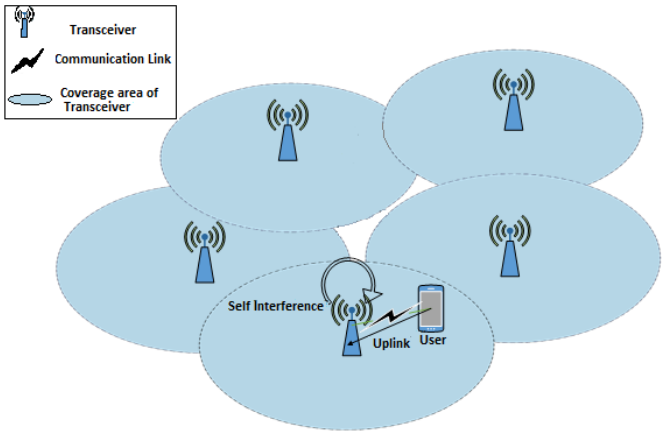


Fig. 2. Full-duplex up-link communication exploiting Traditional DAS

III. Theoretical Analysis

To meet the demand of the rapid growth of consumers 3GPP has introduced a new path loss model for 5G. This model supports higher frequencies (above 6 GHz up to 100 GHz). This model is also applicable for Line of Sight (LOS) and Non Line of Sight (NLOS) scenarios having log normal shadow fading with some basic criteria like distance between Base Station (BS) and UT, UT height, BS height, etc. [7]. In our proposed method, we have taken this path loss model into consideration and evaluated the received power by the Receiver from UT and the SI power from transmitter. Considering the peak to peak distance between Receiver/Transmitter and UT is three dimensional distance (d_{3D}), the linear distance between Receiver/Transmitter and UT is two dimensional distance (d_{2D}), the following equations for outdoor Users are as follows: [7]

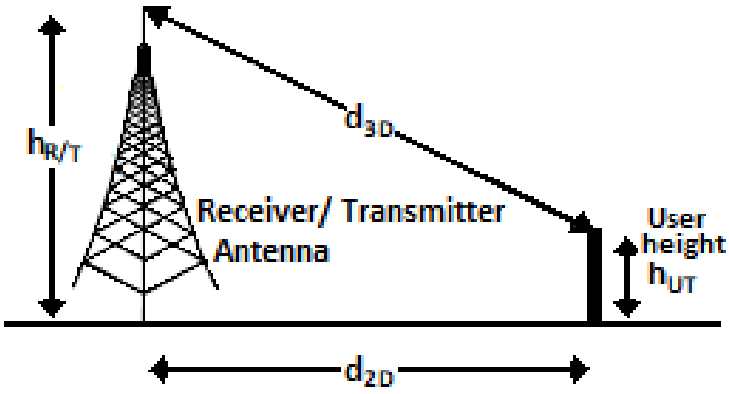


Fig. 3. d_{2D} and d_{3D} for outdoor users.

From Fig. 3,

$$d_{3D} = \sqrt{(d_{2D})^2 + (h_{R/T} - h_{UT})^2} \tag{1}$$

here,
$h_{R/T}$ = Receiver / Transmitter height
h_{UT} = UT height

When the receiver is connected to UT we take receiver height (h_R) into consideration and transmitter height,(h_T) when UT is connected to transmitter for this path-loss model. The path-loss model equations for LOS and NLOS for Urban Macro (UMa) and Rural Macro (RMa) outdoor along with all the parameters are described below:

A. Path-Loss Equation for UMa:

1) For LOS:

$$\text{PL}_{UMa-LOS} = PL_1 \quad \text{for}\,(10m \le d_{2D} \le d'_{BP}) \tag{2}$$

$$\text{PL}_{UMa-LOS} = PL_2 \quad \text{for}\,(d'_{BP} \le d_{2D} \le 5Km) \tag{3}$$

where,

$$PL_1 = 32.4 + 20\log_{10}(d_{3D}) + 20\log_{10}(fc) \tag{4}$$

$$PL_2 = 32.4 + 40\log_{10}(d_{3D}) + 20\log_{10}(fc) - 10\log_{10}((d'_{BP})^2 + (h_{R/T} - h_{UT})^2) \tag{5}$$

$$\text{and} \quad d'_{BP} = 4h'_{R/T}h'_{UT}fc/c \tag{6}$$

Here,

Center frequency is fc ;
The effective antenna heights at the Receiver/Transmitter (in meters), $h'_{R/T} = h_{R/T} - h_E$;
The effective antenna heights at the UT (in meters), $h'_{UT} = h_{UT} - h_E$;
Breaking point distance is d'_{BP} ;
The variance, $\sigma_{SF} = 4$ for shadow fading.

2) For NLOS: Keeping the parameters unchanged, the path-loss for UMa-NLOS scenarios, $PL_{UMa-NLOS}$ is given by:

$$PL_{UMa-NLOS} = \max(PL_{UMa-LOS}, PL'_{UMa-NLOS}) \quad for\ (10m \leq d_{2D} \leq 5Km) \tag{7}$$

where,

$$PL'_{UMa-NLOS} = 13.54 + 39.08\log_{10}(d_{3D}) + 20\log_{10}(fc) - 0.6(h_{UT} - 1.5) \tag{8}$$

For this case, the variance σ_{SF}= 6 for shadow fading.

B. Path-Loss Equation for RMa:

1) For LOS:

$$\text{PL}_{RMa-LOS} = PL_1 \quad \text{for}\,(10m \leq d_{2D} \leq d_{BP}) \tag{9}$$

$$\text{PL}_{RMa-LOS} = PL_2 \quad \text{for}\,(d_{BP} \leq d_{2D} \leq 5Km) \tag{10}$$

where,

$$PL_1 = 20\log_{10}(4\pi d_{3D} fc/3)) + 0.002\log_{10}(h)d_{3D} + \min(0.03h^{1.72}, 10)\log_{10}(d_{3D}) - \min(0.044h^{1.72}, 14.77) \tag{11}$$

here, variance σ_{SF}= 4 for shadow fading.

$$PL_2 = PL_1(d_{BP}) + 40\log_{10}(d_{3D}/d_{BP}) \tag{12}$$

$$\text{and} \quad d_{BP} = 2\pi h_{R/T} h_{UT} fc/c \tag{13}$$

here, d_{BP} is Breaking point distance and variance σ_{SF} becomes 6 for shadow fading.

2) For NLOS: The parameters and ranges are similar to LOS. The path loss equation for NLOS-RMa scenarios, $PL_{RMa-NLOS}$ is:

$$PL_{RMa-NLOS} = \max(PL_{RMa-LOS}, PL'_{RMa-NLOS}) \quad for\ (10m \leq d_{2D} \leq 5Km) \tag{14}$$

where,

$$PL'_{RMa-NLOS} = 161.04 - 7.1\log_{10}(W) + 7.5\log_{10}(h) - (24.37 - 3.7(h/h_{R/T})^2)\log_{10}(h_{R/T}) + (43.42 - 3.1\log_{10}(h_{R/T}))(\log_{10}(d_{3D}) - 3) - (3.2(\log_{10}(11.75h_{UT}))^2 - 4.97) + 20\log_{10}(fc) \tag{15}$$

here, variance σ_{SF} becomes 8 for shadow fading.

C. Received Power and SINR Calculation:

The following equation is used for calculating both the received power and SI power,

$$P_r = P_T - P_L - X \tag{16}$$

where,

$$P_r = Received\ Signal\ Power,$$
$$P_T = Transmitted\ Power$$
$$P_L = Path\ Loss,$$
$$X = Shadow\ Fading$$

Calculation of received power from the UT for up-link ,

$$P_r = P_T(UT) - P_L(UT) - X \tag{17}$$

Calculation of received SI power from the transmitter,

$$P_i = P_T(T) - P_L(T) - X \tag{18}$$

The SINR calculation for our proposed model is as follows:

$$SINR = \frac{P_r}{P_i + P_n} \tag{19}$$

where, P_n is the thermal noise power for one resource block in LTE.
In case of traditional DAS system, the value of SI power of transceiver is same as the value of transmitted power from the transmitter, $P_T(T)$ as there is no presence of path loss or shadow fading.

IV. Results & Discussions

Using the path loss model equations for LOS and NLOS for UMa and RMa outdoor along with all other parameters standardized by 3GPP provided in Section III, SINR performance is evaluated for our proposed Tx-Rx isolated DAS and a comparison with traditional DAS is demonstrated to mitigate the effects of SI. Parameters used for this analytical simulation with their values are displayed in Table I and II for UMa and RMa respectively.

The following Fig. 4 and Fig. 5 delineate the comparison of SINR performance between our proposed and traditional DAS for different frequencies. The graphs are shown for UMa LOS and NLOS scenarios respectively. From these plots it is clear that although SINR performance enhances with the decrease in frequency for both models, the output of traditional DAS is

TABLE I
SIMULATION PARAMETERS FOR UMa SCENARIOS [7]

Parameters	Value
Antenna height of Transmitter (h_T)	20m
Antenna height of Receiver (h_R)	25m
Antenna height of UT (h_{UT})	2m [1.5m $\leq h_{UT} \leq 22.5m$]
Receiver range	4Km
Effective environment height (h_E)	1m
Effective Antenna height of Transmitter (h'_T)	19m
Effective Antenna height of Receiver (h'_R)	24m
Effective Antenna height of UT (h'_{UT})	1m
Light velocity (c)	$3 \times 10^8 ms^{-1}$
Transmitted power of Transmitter ($P_T(T)$)	43 dBm
Transmitted power of UT ($P_T(UT)$)	23 dBm
Thermal noise power (P_n)	-121.45 dBm

TABLE II
SIMULATION PARAMETERS FOR RMa SCENARIOS [7]

Parameters	Value
Antenna height of Transmitter (h_T)	20m [10m $\leq h_{DAS} \leq 150m$]
Antenna height of Receiver (h_R)	35m [10m $\leq h_R \leq 150m$]
Antenna height of UT (h_{UT})	1.5m [1m $\leq h_{UT} \leq 10m$]
Receiver range	4Km
Average building height (h)	15m [5m $\leq h \leq 50m$]
Average street width (W)	20m [5m $\leq W \leq 50m$]
Light velocity (c)	$3 \times 10^8 ms^{-1}$
Transmitted power of Transmitter ($P_T(T)$)	43 dBm
Transmitted power of UT ($P_T(UT)$)	23 dBm
Thermal noise power (P_n)	-121.45 dBm

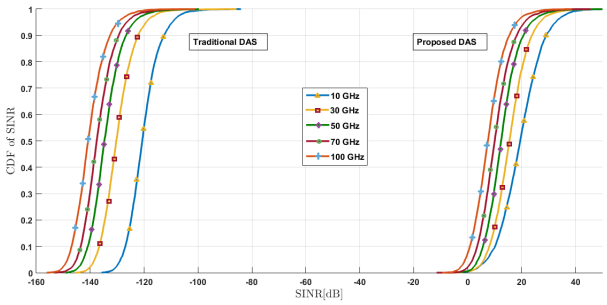


Fig. 4. Comparison between proposed and traditional DAS varying frequency for Urban Macro (LOS).

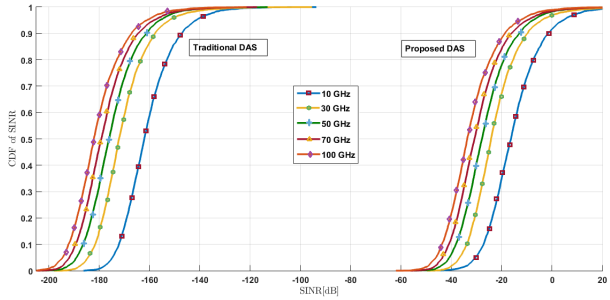


Fig. 5. Comparison of SINR performance between proposed and traditional DAS varying frequency for Urban Macro (NLOS).

always far worse due to the receiver antenna of the transceiver being highly exposed to the unwanted transmitted signal of its transmitter antenna which is SI. On the other hand, the proposed model gives a better output as the transmitter and receiver are separated and SI is reduced by creating distance between them.

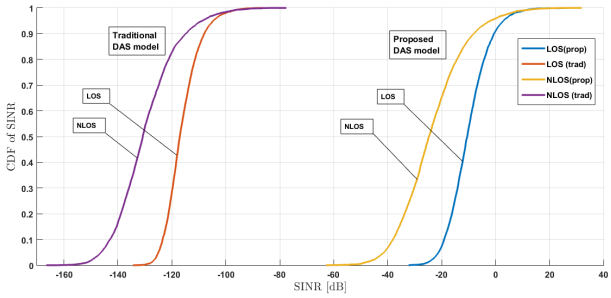


Fig. 6. Comparison of SINR performance between proposed and traditional DAS for Rural Macro (LOS and NLOS) keeping the center frequency fixed at 7 GHz.

Fig. 6 compares the SINR output between our proposed and traditional DAS for Rural Macro LOS and NLOS scenarios keeping the center frequency fixed at 7 GHz. According to the graph the traditional DAS shows high SI and worse performance while the proposed model gives a better output.

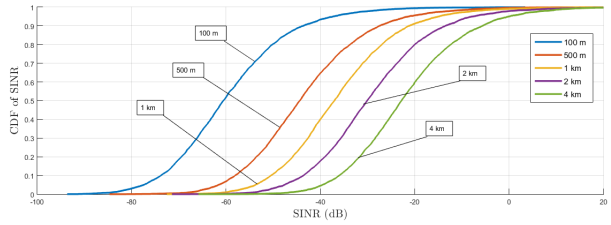


Fig. 7. SINR performance of proposed DAS for RMa-NLOS varying transmitter distance at 7GHz.

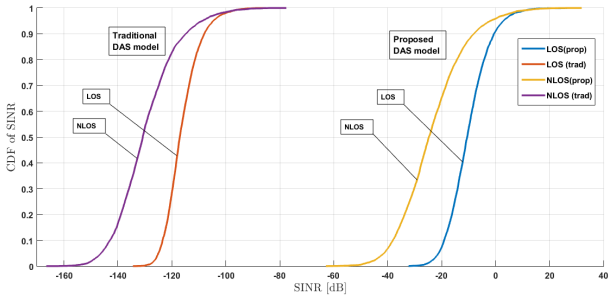


Fig. 8. SINR performance of proposed DAS for UMa-LOS varying transmitter distance at 50GHz.

SINR performance of proposed DAS for RMa-NLOS and UMa-LOS varying transmitter distance from receiver keeping the frequency fixed at 7 GHz and 50 GHz respectively are explained by using Fig. 7 and Fig. 8 respectively. As the distance increases between the transmitter and the receiver the

SI decreases resulting in a better performance. In the figures, the performance becomes saturated as the distance increases. It is also observed that SI is lowest and the SINR performance is quite better when the distance between transmitter and receiver is greatest (4 km).Varying distance for traditional DAS has no effect on SINR performance as every node works as transceiver.

V. Conclusion

The proposed system model presented in this paper comprises of two technologies: FD communication and DAS. Simulations are evaluated for both Urban and Rural outdoor case scenarios with various parameters to enumerate their impacts on the proposed and traditional network. After analysis, it is found that, for traditional DAS (where the nodes perform as transceiver), the SINR was worse due to a huge amount of SI. But for the proposed DAS (with the same parameters), the SINR was much better in comparison. The performance of SINR for 5th generation communication network is better when a lower frequency is used. Furthermore, SI degrades the overall system performance when transmitter is closest to the receiver. Gradual increase in distance reduces SI and enhances the output of the system. This model can be further used for indoor scenarios and down-link channel.

References

[1] M. G. Sarret, M. Fleischer, G. Berardinelli, N. H. Mahmood, P. Mogensen, H. Heinz, On the Potential of Full Duplex Performance in 5G Ultra-Dense Small Cell Networks, 24th European Signal Processing Conference, Sept. 2016

[2] N. H. Mahmood, M. G. Sarret, G. Berardinelli, P. Mogensen, Full Duplex Communications in 5G Small Cells. 13th IEEE International Wireless Communications and Mobile Computing Conference Valencia, Spain, June 2017

[3] B. Debaillie, B. van Liempd, B. Hershberg, J. Craninckx, K. Rikkinen, D.J. van den Broek, E.A.M. Klumperink, B. Nauta , In-Band Full-Duplex Transceiver Technology for 5G Mobile Networks, 41st European Solid-State Circuits Conference, Sept. 2015

[4] C. Mosquera, Wideband Self-Interference Cancellation for Better Spectrum Use. World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, July 2016.

[5] S. Hong, J. Brand, J. Choi, M. Jain, J. Mehlman, Applications of Self-Interference Cancellation in 5G and Beyond. pp-114, IEEE Communications Magazine, Feb 2014.

[6] Z. Zhang, X. Chai, A. V. Vasilakos, L. Hanzo, Full Duplex Techniques for 5G Networks: Self-Interference Cancellation, Protocol Design, and Relay Selection, IEEE, May 2014.

[7] 3GPP, Study on channel model for frequency spectrum above 6 GHz,TR 38.900 V14.3.1 (2017-07).

[8] S.K. Ramachandran, Establishing a Regulatory Framework for Distributed Antenna Systems, Telecommunications Policy Research Conference, July 2008

[9] A.A.M. Saleh, A. Rustako, R. Roman, Distributed Antennas for Indoor Radio Communications. IEEE Transactions on Communications, Vol-35, Issue-12, December 1987.

[10] W. Choi, J. G. Andrews, Downlink Performance and Capacity of Distributed Antenna Systems in a Multicell Environment. IEEE transactions on wireless communications, vol-6, issue-1, Jan 2007.

Performance Analysis of a CPDM-QPSK Direct Detection Optical Transmission System under the effects of Cross-Polarization

A. K. M. Sharoar Jahan Choyon[*,1], S M Raiyan Chowdhury[**], Ruhin Chowdhury[*]
[*]*Dept. of EECE, Military Institute of Science & Technology (MIST), Dhaka, Bangladesh*
[**]*Dept. of EEE, Bangladesh University of Engineering & Technology (BUET), Dhaka, Bangladesh*
[1]choyonsharoar@gmail.com

***Abstract*—Analysis is developed for a Circular Polarization Division Multiplexing Quadrature Phase Shift Keying (CPDM-QPSK) with direct detection optical transmission system considering the impact of the crosstalk(Xtalk) owing to cross-polarization(XPol) induced crosstalk by arbitrary misalignment of the SOPs. A thorough analysis is demonstrated to acquire the output expression of the Signal to Cross-talk plus Noise Ratio (SCNR) for a direct detection CPDM-QPSK receiver. Considering the Maxwellian distribution of the random misalignment angles, the outcome of Bit Error Rate (BER) performance is presented ultimately. Moreover, cross-polarization affects the BER performances and thus the system suffers a significant deterioration from the expected results. It can also be observed that, the system suffers power penalty of 0.18, 0.92 and 4.44 dB for misalignment angle of 10, 11 and 12 degrees respectively at BER of 10^{-9}.**

***Index Terms*—cross-polarization; crosstalk; circular polarization division multiplexing; direct detection; quadrature phase shift keying.**

I. Introduction

The transmission capacity demand which is ever-growing in optical transmission systems has made it necessary to increase the spectral efficiency by implementing different techniques of transmission. Circular Polarization Division Multiplexing (CPDM), which involves Right Circular Polarization (RCP) and Left Circular Polarization (LCP) is an efficient technique for making the spectral efficiency of fiber-optic communication systems quadrupled. Every state of polarization Right Circular Polarization and Left Circular Polarization of two independently modulated signals at the same wavelength are transmitted in this technique [1]. By the use of coherent receivers or direct detection receivers, these orthogonally polarized signals are demultiplexed at the receiver. Since compared to the direct detection, the coherent detection is more expensive, a number of efforts have been made in recent times for the development of the direct detection technique for the Polarization Division Multiplexing (PDM) signals [2]. A circularly Polarized (CP) beam, when transmitted through a PBS, produces a set of LP beams as vertical and horizontal polarizations meaning that it is more efficient than PDM for the improvement of the spectral efficiency of the fiber optic system and hence makes the information carrying capacity of the fiber double than that of PDM [3-4].

In this paper, we present a proposal to incorporate CPDM into the QPSK fiber optic communication (FOC) system with a view to quadrupling the spectral efficiency and accordingly the bandwidth of the system without altering the spectral width of the transmitter. This paper analyses the performance of a prevailing CPDM-QPSK system with direct detection affected by XPol induced crosstalk. At the output of direct detection receiver, SCNR and BER have been evaluated analytically. Considering different system parameters, a numerical estimation is carried out for the decline of the BER performance by cross-polarization. Similar estimation is carried out for power penalties suffered by cross-polarization.

This is the description of the organization of the paper. The detailed system model for CPDM-QPSK direct detection is mentioned in Section II. Theoretical analysis of the system with configuration is done in Section III by analytical evaluation of the BER and SCNR. Section IV depicts the results found from the simulation. Ultimately, Section V concludes the paper by precising the analysis.

II. System Model

The block diagram of a CPDM-QPSK transmitter is depicted in Fig. 1 where, two orthogonal and circularly polarized input lasers electric fields are separated by the Circularly Polarization Beam Splitter (CPBS) into two components, RCP and LCP. Consequently, these two components individually pass through the Polarization Beam Splitter (PBS) modules and are divided into two linearly polarized beams separately. Using the Mach Zender Modulator linearly polarized signals are QPSK Modulated. Then the RCP and LCP signals are combined using a Circularly Polarization Beam Combiner (CPBC).

The receiver shown in Fig. 2 follows direct detection method for demodulating the transmitted signal. The electric fields of signals received from transmitter section are divided into two components (RCP and LCP) by the help of a CPBS and are further split into two components by two PBSs. Each output of PBS is fed into QPSK demodulators. Hence, the original electrical signal is recovered from its transmitted data. In CPDM-QPSK we get 4 pairs of orthogonal components which eventually quadruples the transmission rate without altering the spectral width of the transmitter.

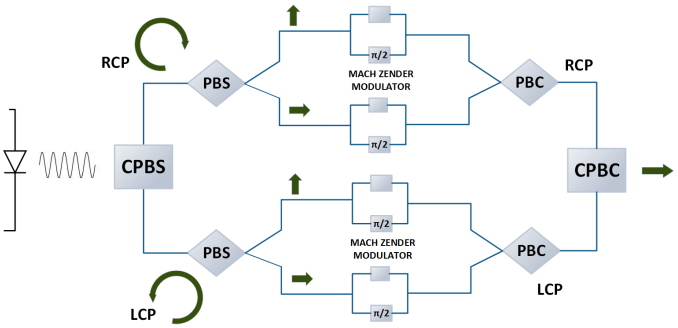


Fig. 1. CPDM-QPSK transmitter section

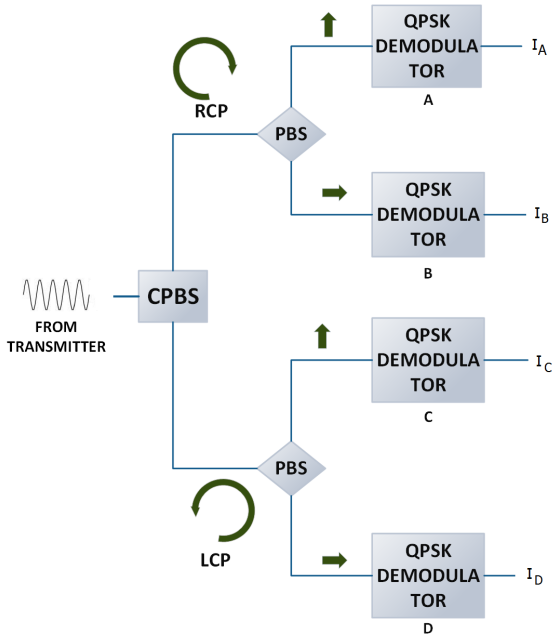


Fig. 2. CPDM-QPSK Direct Detection receiver section

III. Theoretical Analysis

The uneven rotations of the sub-channel SOPs cause misalignment of the eigenmode of PBS of the receiver which deteriorates the signals by inducing crosstalk and fading between the sub-channels.Taking $|e'_x\rangle$ and $|e'_y\rangle$ are the eigenmodes of the PBS, the sub-channel electric field E_x will be the component of the main field along e'_x. Thus [4],

$$|e'_x\rangle \langle e'_x|E|e'_x|E\rangle = \varepsilon_F E_x |e'_x\rangle + \varepsilon_X E_y |e'_x\rangle \tag{1}$$

where,

$$\varepsilon_F = \langle e_x{}'|e_x\rangle = \text{complex co-efficient of fading}$$
$$\varepsilon_X = \langle e_x{}'|e_y\rangle = \text{complex co-efficient of crosstalk}$$

The magnitude of ε_F and ε_X can be determined as [4],

$$|\varepsilon_F| = \cos(\theta/2) \text{ and } |\varepsilon_X| = \sin(\theta/2) \tag{2}$$

here, the random misalignment angle, θ satisfies the Maxwell-Boltzman pdf with mean θ_m. Thus, the pdf of θ can be written as [5],

$$p(\theta) = \frac{32}{\pi^2 {\theta_m}^3}\theta^2 \exp\left\{\left(\frac{-4\theta^2}{\pi {\theta_m}^2}\right)\right\} \tag{3}$$

The output signal at the PBS [1],

$$\begin{aligned} E_{RCP,x} &= \text{Re}[\varepsilon_F E_s e^{j\phi} + j\varepsilon_X E_s e^{j\phi}] \\ &= \varepsilon_F E_s \cos(\omega t - kz) - \varepsilon_X E_s \sin(\omega t - kz) \end{aligned} \tag{4}$$

$$\begin{aligned} E_{RCP,y} &= \text{Re}[\varepsilon_F E_s e^{j(\phi+\frac{\pi}{2})} + j\varepsilon_X E_s e^{j(\phi+\frac{\pi}{2})}] \\ &= -\varepsilon_F E_s \sin(\omega t - kz) - \varepsilon_X E_s \cos(\omega t - kz) \end{aligned} \tag{5}$$

where, $\phi = (\omega t - kz)$.
Likewise,

$$E_{LCP,x} = \varepsilon_F E_s \cos(\omega t - kz) - \varepsilon_X E_s \sin(\omega t - kz) \tag{6}$$

$$E_{LCP,y} = \varepsilon_F E_s \sin(\omega t - kz) + \varepsilon_X E_s \cos(\omega t - kz) \tag{7}$$

The PBC input is as follows,

$$\begin{aligned} E_{RCP,1} &= -(\varepsilon_X E_s \cos\phi + \varepsilon_F E_s \sin\phi) \\ &\quad + (\varepsilon_X E_s \sin\phi - \varepsilon_F E_s \cos\phi) \end{aligned} \tag{8}$$

$$\begin{aligned} E_{RCP,2} &= -(\varepsilon_X E_s \cos\phi + \varepsilon_F E_s \sin\phi) \\ &\quad - (\varepsilon_X E_s \sin\phi - \varepsilon_F E_s \cos\phi) \end{aligned} \tag{9}$$

The PBC output is,

$$\begin{aligned} E_{RCP} &= a_{even}(\varepsilon_X E_s \cos\phi + \varepsilon_F E_s \sin\phi) \\ &\quad + a_{odd}(\varepsilon_X E_s \sin\phi - \varepsilon_F E_s \cos\phi) \end{aligned} \tag{10}$$

Similarly,

$$\begin{aligned} E_{LCP} &= b_{even}(\varepsilon_X E_s \cos\phi + \varepsilon_F E_s \sin\phi) \\ &\quad + b_{odd}(\varepsilon_X E_s \sin\phi - \varepsilon_F E_s \cos\phi) \end{aligned} \tag{11}$$

For direct detection, the received signal from the transmitter,

$$r_A = s\exp(-\alpha d) \tag{12}$$

where, $s = E_{RCP} + E_{LCP}$
α = attenuation constant, d = distance from the transmitter to receiver.

The output photo-current from the QPSK demodulator,

$$\begin{aligned} I_A &= R_d|r_A|^2 \\ &= R_d|a_{odd}(\varepsilon_X E_s \sin\phi - \varepsilon_F E_s \cos\phi)|^2 \\ &= R_d(\varepsilon_X^2 P_s + \varepsilon_F^2 P_s) \\ &= I_{sig,A} + I_{xtalk,A} \end{aligned} \tag{13}$$

where, the responsivity of photo-diode is,

$$R_d = \frac{e\eta}{h\nu}$$

where,

η = quantum efficiency, ν = operating frequency

e = electronic charge, h = Plank's constant

$$P_s = \frac{E_s^2}{2},\ P_{LO} = \frac{E_{LO}^2}{2}$$

where, the singal and the crosstalk currents are,

$$I_{sig,A} = R_d \varepsilon_F^2 P_s$$
$$I_{xtalk,A} = R_d \varepsilon_X^2 P_s$$

So, combining XPol-induced crosstalk for a chosen BER and mean misalignment angle θ_m, the conditional $SCNR(\theta)$ can be calculated by,

$$SCNR(\theta) = \frac{|I_{sig,A}|^2}{|I_{xtalk,A}|^2 + \sigma_{noise}^2} \tag{14}$$

where, the total noise power,

$$\sigma^2{}_{noise} = \sigma^2{}_{thermal} + \sigma^2{}_{shot}$$

where , $\sigma^2{}_{thermal} = 4kTB/R_L$ and $\sigma^2{}_{shot} = \sqrt{2}eBI_{sig,A}$

Using the following formulas we can calculate both the $BER(\theta)$ and the average BER,

$$BER(\theta) = \frac{1}{2}\text{erfc}(\sqrt{\frac{SCNR(\theta)}{2}}) \tag{15}$$

The average BER is,

$$\text{BER} = \int_0^{\pi/2} BER(\theta) p(\theta) d\theta \tag{16}$$

IV. Results & Discussions

Subsequent to the analytical obtainment given in previous section, Section IV elucidates the simulation results of SCNR(θ), BER(θ) and power penalty influenced by cross-polarization induced crosstalk. Table I displays the parameters with appropriate values for this simulation.

TABLE I
Simulation Parameters and their Standard Values

Parameters	Value
Charge of electron , e	1.6021 × 10^{-19} Coul
Responsivity , R_d	0.85
Temperature , T	300K
Boltzmann Constant , k	1.38065 × 10^{-23} J/K
Load Resistance , R_L	25 Ω

Fig. 3 delineates the plots of SCNR (with and without the XPol-induced crosstalk) vs misalignment angle for different values of optical received signal power of -20 dBm, -15 dBm and -10 dBm. Due to XPol-induced crosstalk there is a significant drop in the performance of SCNR as the SCNR decreases with the rise of misalignment angles which is clear from the equation (14). When there is no crosstalk, SCNR is higher and it rises with the rise in optical received signal power, also supported by the equation (14). The maximum SCNR is found to be 33.2 dB when there is no cross-polarization at a received signal power of -10 dBm.

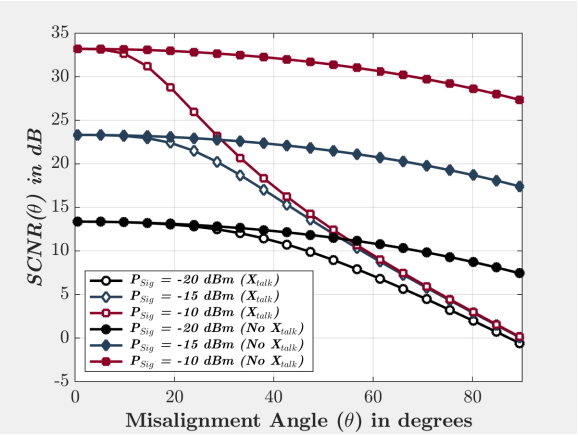


Fig. 3. SCNR (with and without crosstalk) vs misalignment angle for different received signal power.

The increment of BER with the rise of mean misalignment angles is shown in Fig. 4. When the received signal power is higher it causes lower BER and at a fixed received signal power BER rises with rise in crosstalk that is misalignment angles. It is evident in the graph that at a given signal power of -6 dBm , the BER rises from 10^{-20} to almost 10^{-1} for incrementing mean misalignment angles from 9 to 45 degrees. So, it is evident that increasing misalignment angle also gradually increases the BER for an adjusted signal power since misalignment angle increments Xtalk power. The degrading performance of BER is caused by the emersion in crosstalk power which satisfies Fig. 4 as well as the equations (14) and (15).

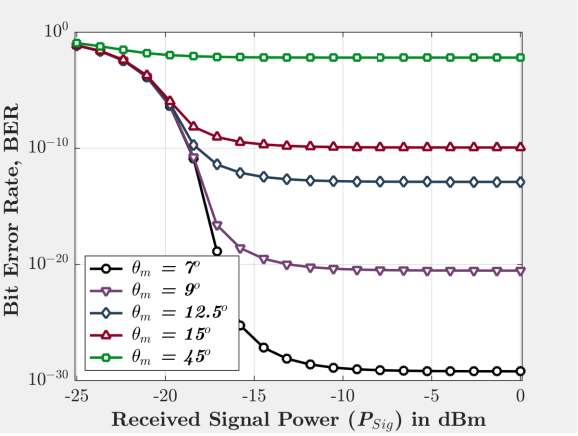


Fig. 4. BER vs Received signal power for different misalignment angles.

Fig. 5 signifies the loss of SCNR as a function of misalignment angle for multiple values of signal power. We see that the losses of SCNR is extra at greater values of received signal power and higher values of misalignment angles. The

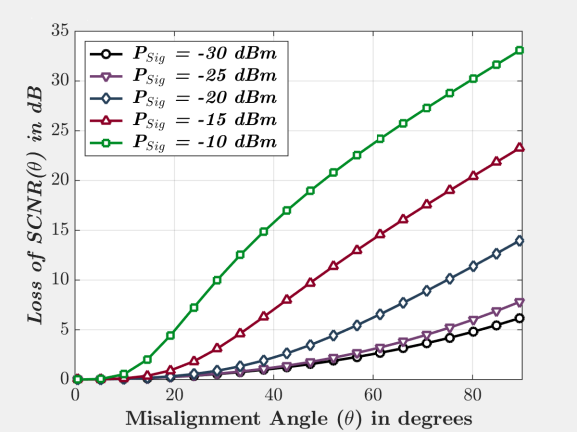


Fig. 5. Loss of SCNR vs misalignment angle for multiple values of signal powers.

loss of SCNR is 8.66 dB from 9.99 dB to 1.33 dB when the received signal power is lessened from -10 dBm to -30 dBm at a misalignment angle of 30 degree.

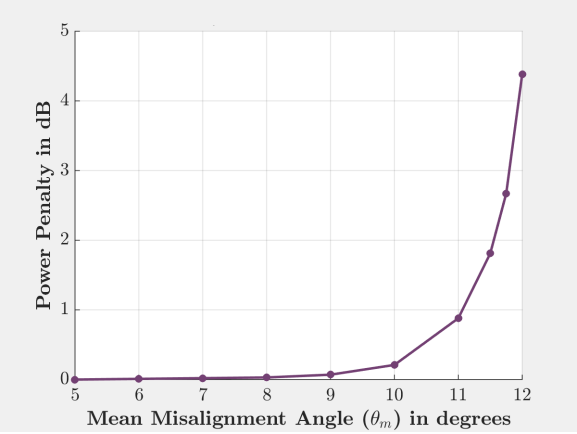


Fig. 6. Power penalty vs mean misalignment angle at BER 10^{-9}.

Fig. 6 represents the power penalty vs mean misalignment angles at a fixed BER of 10^{-9}. This plot elucidates the result of XPol-induced crosstalk which suffers power penalty. According to the estimation, the system undergoes power penalty of 0.18, 0.92 and 4.44 dB for the mean misalignment angles of 10, 11 and 12 degrees respectively at a BER of 10^{-9}. The given analysis makes it explicit that, as the misalignment angle augments, the BER performance declines for the interference of crosstalk into the main signal.

V. Conclusion

This paper introduces CPDM-QPSK for direct detection FOC system intending to quadruple the spectral efficiency. According to the analysis of SCNR, BER and power penalty, it is observed that the interference of crosstalk into main signal weakens the system resulting in the decline of BER and SCNR performance. Therefore, the system suffers power penalty for different misalignment angles. This eventually lessens the receiver sensitivity to acquire a BER of 10^{-9}. However, using this method the spectral bandwidth can be improved to a great extant. The current approach stands on the theoretical and mathematical analysis of circular polarization in light which can provide compelling advantages in the future of faster FOC system.

References

[1] A.K.M. Sharoar, S.M. Raiyan, S.P. Majumder, "Performance of a CPDM-QPSK Coherent Homodyne Optical Transmission System due to Cross Polarization Effects", International Conference on Computer, Communication, Chemical, Material and Electronic Engineering, Feb. 2018.

[2] D. Che, A. Li, X. Chen, Q. Hu, Y. Wang, and W. Shieh, Stokes vector direct detection for short-reach optical communication, Opt. Lett., vol. 39, no. 11, pp. 31103113, Jun. 2014.

[3] Krummrich PM, Schmidt ED, Weiershausen W, Mattheus A. Field trial results on statistics of fast polarization changes in long haul WDM transmission systems. In OFC/NFOEC 2005; OThT6.

[4] Winter M, Bunge CA, Setti D, Petermann K. A statistical treatment of cross-polarization modulation in DWDM systems. J Lightwave Technol 2009;27:373951

[5] Masayuki M. Polarization-mode dispersion and its mitigation. J Natl Inst Inf Commun Technol 2006;53:4652.

Design and Simulation of an UWB Vivaldi Antenna for Cancer Detection and Treatment

Antora Hossain
Department of Electronics & Telecommunication Engineering,
Rajshahi University of Engineering & Technology,
Rajshahi-6204, Bangladesh
antorahossain94@yahoo.com

Md. Kamal Hosain
Department of Electronics & Telecommunication Engineering,
Rajshahi University of Engineering & Technology,
Rajshahi-6204, Bangladesh
kamaleteruet@gmail.com

***Abstract*—An UWB Vivaldi antenna is proposed for cancer detection and treatment. The antenna operates at the resonant frequency of 4.76 GHz within the bandwidth of 4.63 GHz (3.13 GHz-7.76 GHz) in free space. In addition, the antenna possesses a high radiation efficiency of 81.34% and a gain of 5.81 dB in free space condition. The overall dimension of the proposed antenna is 110.46 x 96.77 x 1.67 mm^3. The antenna is simulated with 1.6 mm thick FR-4 dielectric substrate since it is cost-effective and easily available. Furthermore, a six-layer biological tissue model comprising of skin, fat, outer cortical bone, cancellous bone, inner cortical bone, and muscle with a tumor of 10 mm radius is modeled to assess antenna performance for cancer detection and treatment. In simulation, a specific gap of 2.5 mm is maintained between the proposed antenna and the proposed phantom model in order to avoid skin burn and other side-effects. Cancerous cells in tumor are detected and killed because of higher SAR value of the cancerous tissue than the normal healthy tissue.**

Keywords—UWB Vivaldi antenna, phantom model, SAR, cancer detection, cancer treatment.

I. INTRODUCTION

The underlying concept of cancer detection and treatment is to diagnose and cure cancerous cells ensuring minimum cumulative side-effects of the nearby healthy tissues. The malignant tumor is responsible for generating cancer which is the second leading cause of death [1]. Antenna can be used as the basic element of cancer detection and treatment. For the detection of cancer, microwave imaging is one of the most prominent methods where the tumor shows electrical properties that are substantially different from those of healthy tissues. One leading feature of microwave imaging antenna is that the antenna must have ultra-wideband (UWB) characteristics to enhance the resolution of the image [2]. Cancer treatment can also be performed by applying high doses of antenna radiation beam on the detected tumor region. This process is called radiotherapy. It is executed very carefully so that the adjacent healthy tissues do not get damaged along with the cancerous tissues. In recent years, many antennas including Vivaldi antenna, Dipole antenna, Monopole antenna, Bow-tie antenna, Pyramidal horn antenna, Microstrip patch antenna, Log periodic antenna, and Spiral antenna have been designed for microwave imaging and radiotherapy applications [3].

The Vivaldi antenna is one of the most popular directive antennas which is widely used for commercial UWB applications like microwave imaging and radiotherapy. This antenna provides significant advantages such as high efficiency, high gain, wide bandwidth, simple geometry and so on. In a recent paper, Alzabidi *et al.* [2] proposed an optimum UWB Vivaldi antenna for tumor detection. The antenna had an antipodal configuration and it achieved better performance for an overall dimension of 329.25 x 153 x 1.6 mm^3. The bandwidth of the antenna was nearly 0.8 GHz after simulating with phantom model. In another paper, Alsulaiman *et al.* [4] proposed an ultra-wideband balanced antipodal Vivaldi antenna for hyperthermia treatment. The overall volume of the antenna was 82 x 92 x 1.6 mm^3. After simulating this antenna with phantom model, a bandwidth of 3.7 GHz was obtained. In this paper, the aim is to design an UWB Vivaldi antenna assuring improved performance (e.g., bandwidth) than the existing ones for cancer detection and treatment.

In cancer detection and treatment, biocompatibility analysis for human safety is vital. One of the biocompatibility analysis factors is specific absorption rate (SAR). The term SAR refers to a parameter used to measure the power absorbed per mass of body tissue [5]. This calculation is performed as a post-processing step after simulation. According to IEEE C95.3 method, a cube is grown around the specific point until it contains the required mass of body tissue corresponding to a reference power. Tissue with malignant tumor shows higher value of SAR than that of normal healthy tissue [2]. This variation of SAR gives an indication of the presence of tumor and an indication of absorption of power in tumor. The value of SAR can be calculated by using a biological tissue model around the antenna.

The organization of this paper is as follows: Section II presents the theory of Vivaldi antenna, the design equations and configuration of the proposed Vivaldi antenna. Section III illustrates the antenna performance in free space and with the proposed six-layer biological tissue model. Ultimately, discussion and conclusion are provided in Section IV and Section V respectively.

II. VIVALDI ANTENNA DESIGN

A. Theory of Vivaldi Antenna

The basic Vivaldi antenna has an exponentially tapered slot profile. Primarily, the antenna consists of a dielectric substrate and a slot-line for feeding purpose. The basic configuration of this antenna is shown in Fig. 1. The antenna is also known as travelling wave antenna which can be categorized into propagating section and radiating section [6]. The slot width indicates the separation between the conductors. In the propagating section, the slot width is less than one-half of free space wavelength ($\lambda_0/2$) while the opposite situation occurs in the radiating section. Therefore, the bond between the conductors and the curved portion of antenna is stronger in the propagating section than that in the radiating section. The Vivaldi antenna is most fruitful

with slot-line feeding because of design flexibility. In this scheme, a transition is designed to couple signals to the slot-line from the transmitter or receiver. The transition causes low loss over a wide range of frequency so that it does not limit the operating bandwidth.

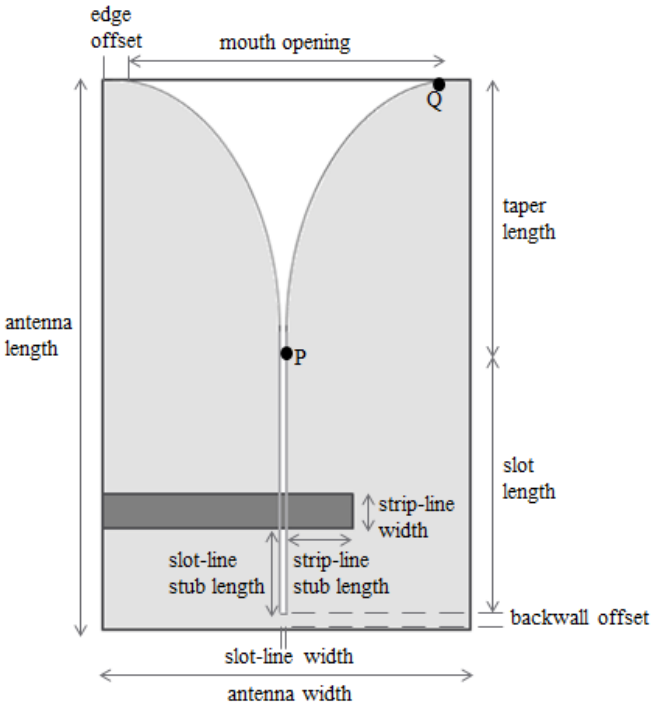


Fig. 1. Basic configuration of Vivaldi antenna.

(i) Substrate width (*subw*): In order to accomplish the desired radiation performance, the substrate width (*subw*) should be greater than one-half wavelength at the lowest frequency of operation (λ_{min}) [6].

$$subw > \frac{\lambda_{min}}{2} \tag{1}$$

(ii) Taper length (*TL*): In this design, the taper length (*TL*) is equal to the substrate width (*subw*) as given below:

$$TL = subw \tag{2}$$

(iii) Taper rate (*r*): The taper rate (*r*) can be defined by an exponential equation as [7],

$$Y = C_1 e^{-rx} + C_2 \tag{3}$$

where, $C_1 = \frac{y_2 - y_1}{e^{rx_2} - e^{rx_1}}$ and $C_2 = \frac{e^{rx_2} y_1 - e^{rx_1} y_2}{e^{rx_2} - e^{rx_1}}$

(iv) Strip-line width (*MW*): Depending on the dimensions of the transmission line, the characteristic impedance Z_0 can be calculated as [7]:

$$Z_0 = \frac{60}{\varepsilon_r} ln\left(\frac{8h}{MW} + \frac{MW}{4h}\right) \tag{4}$$

where, *MW* = strip-line width and *h* = substrate thickness

(v) Slot-line width (*2S*): It indicates the separation between the slot-line conductors before the tapered section [6]. This parameter should be determined carefully in order to get the desired return loss response.

(vi) Slot-line stub length (*QWS*): The length of slot-line stub is determined by this parameter [6]. The stub length should be open circuited by means of a short circuited quarter wavelength stub in order to improve field coupling.

(vii) Strip-line stub length (*QWM*): The strip-line prolongs a finite length and ends to slot-line transition [6].

(viii) Slot length (*SL*): In this design, the slot length (*SL*) is calculated by summing stripline width (*MW*) and slot-line stub length (*QWS*) as follows:

$$SL = MW + QWS \tag{5}$$

(ix) Microstrip length (*ML*): In this design, the microstrip length (*ML*) is obtained by summing quarter wave microstrip (*QWM*) i.e., strip-line stub length, half of slot width (*s*) and half of substrate width (*subw*) as follows:

$$ML = QWM + s + subw/2 \tag{6}$$

(x) Backwall offset (*ext*): The backwall offset (*ext*) refers to the extra metallization that is implanted at the origin of the slot-line [6]. This parameter should be optimized in order to prevent any deterioration in return loss response.

(xi) Antenna length (*L*): Antenna length (*L*) is the summation of taper length (*TL*), strip-line width (*MW*), strip-line stub length (*QWS*) and backwall offset (*ext*). Antenna length should be greater than a free space wavelength at the lowest frequency of operation (λ_{min}) to ensure an improved gain and beamwidth performance [6]. Mathematically, it is expressed as follows:

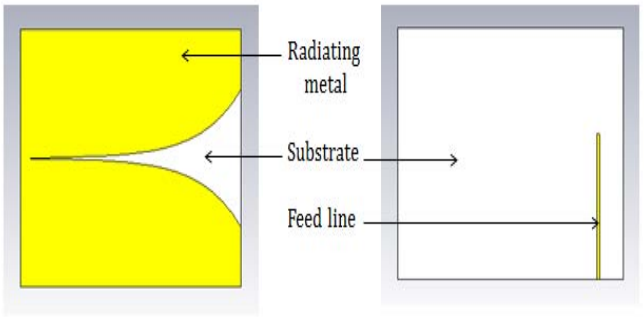
$$L = TL + MW + QWS + ext \tag{7}$$

(xii) Antenna width (*W*): In this design, antenna width (*W*) is equal to the substrate width (*subw*) as given below:

$$W = subw \tag{8}$$

B. Antenna Configuration

The configuration of the proposed Vivaldi antenna is shown in Fig. 2. The design parameters of this antenna are listed in Table I which have been calculated from the design equations given in Sub-section A. In this antenna design, the annealed copper with a thickness of 0.035 mm is used as the metal or radiating element and the FR-4 lossy dielectric material with a thickness of 1.6 mm is used as the substrate. The microstrip to slot-line feed is used as the feeding technique to obtain the resonance at the selected frequency of 4.76 GHz. The total volume of the antenna is 110.46 x 96.77 x 1.67 mm^3. The top metallic part provides the current flow path. The feedline is located at 0.035 mm distance from the substrate at the bottom part. The iterative simulation method is used to match the antenna impedance with the feeder impedance of 50 Ω.



(a) (b)

Fig. 2. Configuration of the proposed Vivaldi antenna (a) Top view (b) Bottom view.

TABLE I. DESIGN PARAMETERS OF THE PROPOSED VIVALDI ANTENNA

Design parameter	Symbol	Value (mm)
Substrate thickness	*h*	1.6
Substrate width	*subw*	96.77
Taper length	*TL*	96.77
Taper rate	*r*	0.05
Stripline width	*MW*	1.3068
Half slot-line width	*S*	0.20565
Slot-line stub length	*QWS*	7.38
Stripline stub length	*QWM*	7.38
Slot length	*SL*	8.6868
Microstrip length	*ML*	55.97
Backwall offset	*ext*	5
Antenna length	*L*	110.46
Antenna width	*W*	96.77
Metal thickness	*t*	0.035
Antenna thickness	*T*	1.67

III. ANTENNA PERFORMANCE

A. In Free Space

The design and simulation of the proposed antenna is done in the commercial software package of CST Studio Suite 2015. After simulating the antenna in free space, the following results are obtained:

Fig. 3 shows the S_{11} parameter of the proposed Vivaldi antenna in free space where the antenna resonates at 4.76 GHz with a return loss of -22.019946 dB. The simulated bandwidth is 4.63 GHz (3.13 GHz-7.76 GHz) below -10 dB return loss which meets the need of ultra-wideband. Such a bandwidth is achieved by optimizing the design parameters of the antenna which is particular in the minimization of taper rate (*r*). The radiation efficiency of the proposed Vivaldi antenna at 4.76 GHz in free space is -0.8969583 dB, i.e., 81.34 % because of a good impedance matching.

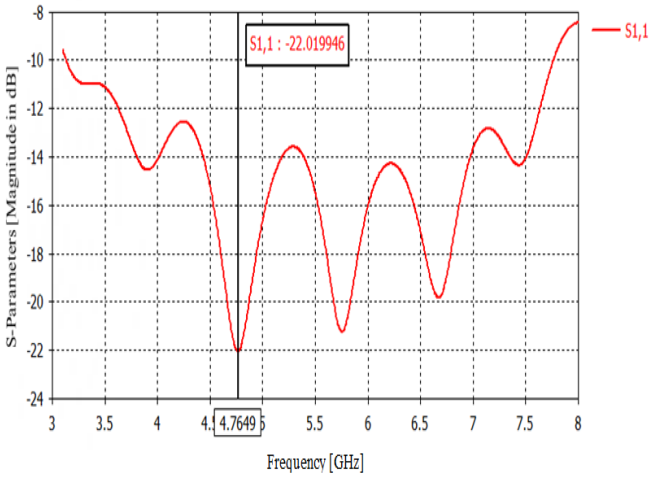
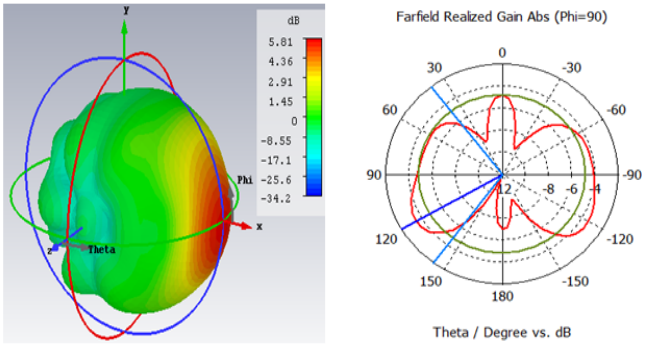


Fig. 3. Simulated return loss response of the proposed Vivaldi antenna in free space.

Fig. 4 (a) shows the 3-D gain pattern and Fig. 4 (b) presents the 2-D gain pattern (recorded at $\theta = 0^0$, $\varphi = 90^0$) at 4.76 GHz of the proposed Vivaldi antenna in free space. From Fig. 4 (a), it is obvious that the maximum gain is 5.81 dB and the maximum radiation is along the central axis of the tapered slot. From Fig. 4 (b), it is apparent that the main lobe magnitude is -3.47 dB along the direction of -122^0. Besides, the 3-dB beamwidth is 106.9^0 and the side lobe level has a small value of -1.3 dB. These characteristics reveal that the antenna exhibits end-fire radiation pattern as required for cancer detection and treatment.

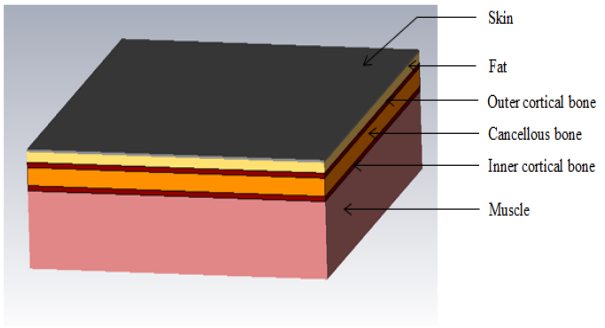


(a) (b)

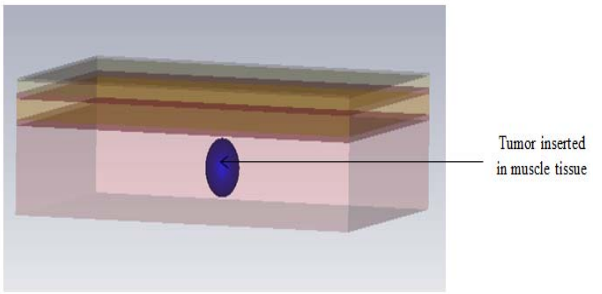
Fig. 4. Simulated far-field gain pattern of the proposed Vivaldi antenna in free space (a) 3-D view (b) 2-D view.

B. Proposed Biological Tissue Model

A six-layer biological tissue model with rectangular shape consisting of skin, fat, outer cortical bone, cancellous bone, inner cortical bone and muscle is modeled as shown in Fig. 5 (a). A tumor with spherical shape is inserted in the muscle tissue layer as shown in Fig. 5 (b). The model is designed using the dielectric properties of tissue as indicated in Table II at the resonant frequency of 4.76 GHz. The thicknesses of the tissue layers and the radius of the tumor are relevant to human body. In order to assess the performance of the antenna with the six-layer biological tissue model, the antenna is pulled on the phantom maintaining a distance of 2.5 mm from skin [8] to avoid skin burn.



(a)



(b)

Fig. 5. (a) Six-layer tissue model (b) Six-layer tissue model with tumor.

TABLE II. DIMENSIONS AND DIELECTRIC PROPERTIES OF SIX-LAYER BIOLOGICAL TISSUE MODEL AT 4.76 GHZ [9]-[10]

Layer name	Volume (mm^3)	Dielectric properties			
		Relative permitivity	Loss tangent	Conductivity (S/m)	Mass density (kg/m^3)
Skin	200x200x1	35.969	0.3021	2.8774	1100
Fat	200x200x4	5.0518	0.17006	0.2275	910
Outer cortical bone	200x200x2	10.154	0.33665	0.9052	1900
Cancellous bone	200x200x7	16.257	0.39757	1.7115	570
Inner cortical bone	200x200x2	10.154	0.33665	0.9052	1900
Muscle	200x200x30	49.852	0.28673	3.7851	1041
Tumor (10 mm radius)	4188.79	50.821	0.26665	4	1041

C. *Antenna Performance with Proposed Biological Tissue Model*

The antenna is simulated with biological tissue model. The proposed Vivaldi antenna resonates at 6 GHz with phantom and tumor which is different from the free space response. Hence antenna size is modified to resonate the antenna at 4.76 GHz like in free space. For this purpose, the dimension of the antenna is merely changed, i.e., the substrate width (*subw*) is increased from 96.77 mm to 99.5 mm and the taper rate (*r*) is reduced from 0.05 to 0.04. Other parameters are kept unchanged. Then the simulated resonant frequency is obtained at 4.76 GHz with a return loss of -19.176292 dB as shown in Fig. 6. Here the bandwidth of 4 GHz (3.24 GHz-7.24 GHz) is found below -10 dB return loss which is slightly lower than 4.63 GHz as obtained in free space. The other performance parameters described below are also simulated with this changed dimension of the antenna.

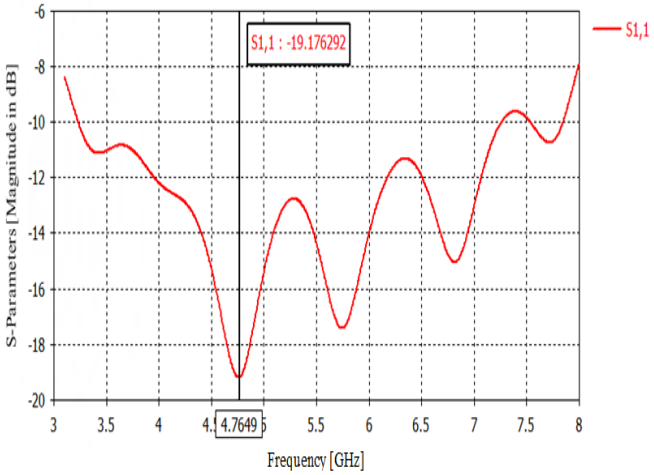
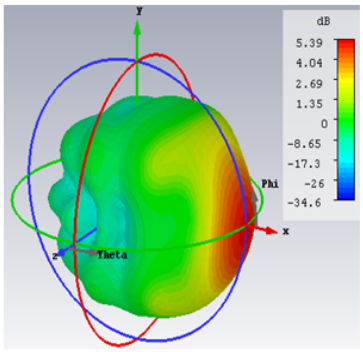
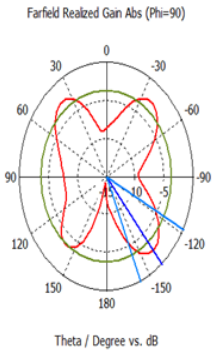


Fig. 6. Simulated return loss response of the proposed Vivaldi antenna with phantom and tumor.

The radiation efficiency of the proposed Vivaldi antenna with phantom and tumor at 4.76 GHz is -1.8206393 dB, i.e., 65.76%. Fig. 7 (a) shows the 3-D gain pattern and Fig. 7 (b) shows the 2-D gain pattern (recorded at $\theta = 0^0$, $\varphi = 90^0$) at 4.76 GHz of the proposed Vivaldi antenna with phantom and tumor. From Fig. 7 (a), it is obvious that the maximum gain is 5.39 dB. It is evident from Fig. 7 (b) that the main lobe magnitude is -2.43 dB along the direction of -140^0. Besides, the 3-dB beamwidth is 39.2^0 and the side lobe level is -1.2 dB.



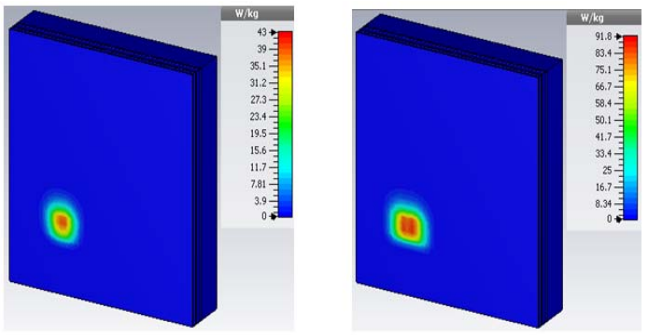
(a)



(b)

Fig. 7. Simulated far-field gain pattern of the proposed Vivaldi antenna with phantom and tumor (a) 3-D view (b) 2-D view.

The SAR distribution of the proposed Vivaldi antenna (i) with phantom (ii) with phantom and tumor is simulated. Fig. 8 (a) shows 10 g average SAR determined at 4.76 GHz and at operating power of 0.5 W. From Fig. 8 (a), it is obvious that the maximum SAR is 43 W/kg which is less than 91.8 W/kg obtained in Fig. 8 (b). This dissimilation is an indication of the presence of tumor and absorption of power in it. The identified tumor is gradually killed as a consequence of absorbing the massive radiation beam of antenna. In this procedure, the detection and treatment of cancer are performed in a comprehensive manner.



(a) (b)

Fig. 8. Simulated SAR distribution of the proposed Vivaldi antenna (a) with phantom (b) with phantom and tumor.

IV. DISCUSSION

The bandwidth (i.e., 4 GHz) of the proposed Vivaldi antenna with biological tissue model is much higher than the bandwidth (i.e., ≈ 0.8 GHz) of that proposed by Alzabidi *et al.* [2] and the bandwidth (i.e., 3.7 GHz) of that proposed by Alsulaiman *et al.* [4]. In addition, the proposed Vivaldi antenna has reduced physical dimension (i.e., 110.46 x 96.77 x 1.67 mm^3) compared to the physical dimension (i.e., 329.25 x 153 x 1.6 mm^3) of that proposed by Alzabidi *et al.* [2].

V. CONCLUSION

In this paper, an UWB Vivaldi antenna is proposed for cancer detection and treatment. Moreover, a six-layer biological tissue model consisting of skin, fat, outer cortical bone, cancellous bone, inner cortical bone, and muscle with a tumor of 10 mm radius is modeled. From the simulation results, it is obvious that the proposed structure has better performance than the existing ones particularly in terms of bandwidth and volume. Also it achieves a high gain and high radiation efficiency. Cancer detection and treatment are assured with the value of SAR. The cancerous tissue exhibits higher SAR value than normal tissue. It can be concluded that the proposed UWB Vivaldi antenna can enhance the resolution of microwave image for cancer detection. This antenna can be used for cancer treatment as well by means of emitting high radiation beam on cancerous cells in tumor. Thus it can provide a bright future in the field of cancer detection and treatment.

REFERENCES

[1] Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S and Mathers C, "Cancer Incidence and Mortality Worldwide," *IARC Cancer Base,* vol. 1, no. 11, 2012.

[2] M. A. Alzabidi, M. A. Aldhaeebi and I. Elshafiey, "Optimization of UWB Vivaldi Antenna for Tumor Detection," *Journal on Artificial Intelligence,Modelling and Simulation (AIMS),* pp. 71-75, 2013.

[3] Nouralhuda A. Hassan, Moustafa Mohamed Ahmed and Mazher B. Tayel, "Basic Evaluation of Antennas Used in Microwave Imaging for Breast Cancer Detection," *Third International Conference on Computer Science and Engineering,* pp. 55-63, August 2016.

[4] Khaled A. Alsulaiman, Mohammed Al alShaykh, Sulaiman Alsulaiman, and Ibrahim Elshafiey, "Design of Ultra Wideband Balanced Antipodal Vivaldi Antenna for Hyperthermia Treatment," *PIERS Proceedings,* pp. 12-15, August 2013.

[5] He Yu, Guo-hui Yang, Qun Wu and Ming-hui Su, "Design and Optimization of UWB Vivaldi Antenna for Brain Tumor Detection," *IEEE Transactions on Antennas and Propagation,* 2016.

[6] Yakup Erdogan, "Parametric Study and Design of Vivaldi Antennas and Arrays," *METU,* pp. 1-149, March 2009.

[7] Rajveer Dhawan and Gurkirandeep Kaur, "Vivaldi Antenna Simulation on Defining Parameters, Parametric Study and Results," *I J C T A,* vol. 9, no. 11, pp. 5129-5138, January 2016.

[8] M.J. Ammann, S. Curto, X.L. Bao and P. McEvoy, "Antenna Design Considerations for High Specific Absorption Rate in Local Hyperthermia Treatment," *IEEE Xplore Conference,* pp. 1-4, July 2008.

[9] C.Gabriel and S.Gabriel, "Compilation of the dielectric properties of body tissues at RF and microwave frequencies," *Occupational and environmental health directorate,* June 1996.

[10] Mohamed M. Elsewe, "Evaluation of EM absorption loss over breast mass for breast cancer diagnosis," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* August,2011.

Hybrid-core Microstructure Fiber for Chemical Identification in Terahertz Regime

Md. Ahasan Habib
Dept. of Electrical and Electronic Engineering
Rajshahi University of Engineering and Technology, Rajshahi-6204
Email: habib.eee.116.ah@gmail.com

Md. Selim Reza
Dept. of Electrical and Electronic Engineering
Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj-8100
Email: selimreza.bsmrstu.eee@gmail.com

***Abstract*—In this paper, a new kind of hybrid-core porous hexagonal microstructure fiber is presented for different chemical identification by using terahertz signal. In order to analyze the guiding characteristics, finite element method (FEM) based Comsol *V*5.3a software is utilized. By filling the core air holes with aqueous analytes (Benzene and Ethanol) and tuning various designing parameters, maximum relative sensitivity of 79.2% and 78.9% is achieved at 1.5 THz for benzene and ethanol respectively. Along with the high sensitivity the proposed fiber offers very low confinement loss of 7.4×10^{-11} cm^{-1} and low bending loss of 4.9×10^{-36} cm^{-1} at optimal condition for Benzene. This proposed fiber is very easy to realize and can be easily fabricated by using the recent popular fabrication technique. It seems that, the excellent characteristics will made the proposed fiber a good candidate in different chemical identification applications.**

Keywords—relative sensitivity, porous core, terahertz, bending loss, confinement loss.

I. INTRODUCTION

Very recently a particular narrow frequency band of electromagnetic spectrum have become a hot topic to the researchers. This frequency band is known as terahertz radiation band and the frequency range of the promising band is 0.1 to 10 THz [1]. The main reason behind the popularity of this narrow frequency band is that this terahertz signal can be effectively used in various applications including communication, security, imaging, spectroscopy, cancer cell detection [2-7] etc. In recent years, terahertz signal is effectively used in different chemical identification such as benzene, ethanol, formalin etc. [8-10]. In case of photonic crystal fiber (PCF), air holes are arranged in different geometric orientation to maintain dielectric condition. According to light guiding mechanism the PCF can be classified into two different category and they are photonic bandgap (PBG) and indexed guiding PCF. When the effective refractive index of the core section is higher than the cladding then the light will be forwarded by using total internal reflection mechanism. If the cladding's refractive index is lower than that of core, then the guiding mechanism will be PBG and this kind of PCF is known as porous core photonic crystal fiber (PC-PCF).

Highly sensitive chemical sensor is essential to identify toxic and flammable chemicals in different industrial applications [11]. That's why the optoelectronics researchers have proposed different types of chemical sensors for different chemical detection. In 2016, Asaduzzaman *et al*. proposed a hybrid core PCF where elliptical shaped air holes are introduced in the core region to create asymmetry in the core [12]. This proposed fiber exhibited low birefringence of 0.0015, relative sensitivity of 49.71% for ethanol, low confinement loss of 2.75×10^{-10} dB/m at optimum geometric condition. A hexagonal lattice chemical sensor with six elliptical air holes in core was proposed by Arif et al. in 2017, which offered maximum sensitivity of 53.35%, low birefringence of 0.0028, low confinement loss of 1.6×10^{-10} dB/m simultaneously at 1.33 μm operating wavelength. After that, Arif *et al*. proposed another PCF based chemical sensor which offered better sensitivity than the previous one [13]. Their proposed hexagonal shaped sensor offered high relative sensitivity of 59% and low confinement loss in the order of 10^{-11} dB/m for optimum designing parameters. Paul et al. proposed a folded shaped silica based porous core photonic crystal fiber [14], which offered very high relative sensitivity of 69% and low confinement loss of 5×10^{-11} dB/cm for benzene at 1.33 μm operating wavelength.

In this paper, we propose a hexagonal type hybrid core microstructure fiber for benzene and ethanol detection by using terahertz signal. The proposed chemical sensor offers very high sensitivity of 78.07%, low confinement loss of 7.4×10^{-11} cm^{-1}, large effective area of 0.94×10^{5} μm^2, very low bending loss of 4.9×10^{-36} cm^{-1} simultaneously at 1.5 THz for optimum geometric condition. The proposed fiber can be fabricated by using Extrusion technique. So, we hope these excellent characteristics will make the chemical sensor a handy one in various sensing applications.

II. GEOMERTY OF PROPOSED SENSOR

The cross-sectional view of the proposed porous-core PCF terahertz sensor with its enlarged version of core and cladding is shown in Fig. 1. Finite element method (FEM) based commercially available software COMSOL Multiphysics v5.3a is used to design the sensor. In cladding region hexagonal type lattice is introduced by using circular shaped air holes whose radius and diameter are denoted by **r** and **d** respectively. To facilitate the fabrication simplicity, sufficient distance between the air holes of the cladding is ensured and the distance between two adjacent air hole in cladding is defined as **p** and the strut thickness of the core is ***a***. The core region consists of a single circular air hole and six triangular air holes. A little edge of the six triangular air holes are detruncated to avoid overlapping. The total diameter of the core is defined as ***D*** which is related to ***r*** i.e. $\boldsymbol{D} = 2.5\boldsymbol{r}$. The chemicals which will be subjected to test will be placed in the core air hole and the light signal will be introduced through the fiber. In the numerical analysis

benzene (n= 1.366) and ethanol (n= 1.354) are used as test analyte and the guiding characteristics are checked. A perfectly matched layer (PML) absorbing boundary which thickness is 10% of the fiber radius is given at the outer part of the cladding because the PML boundary condition itself act as an anti-reflecting layer that absorbs outgoing waves of the PCF. Though there are several number of polymer materials that can be used as a bulk material for designing a sensor, for this proposed sensor Zeonex is chosen as the bulk material. This is because of it's number of useful optical properties such as constant index of refraction of 1.53 in terahertz range, low absorption loss of 0.2 cm^{-1}, high glass transition temperature, negligible water absorption, high biocompatibility and excellent chemical resistance even at elevated temperatures [17].

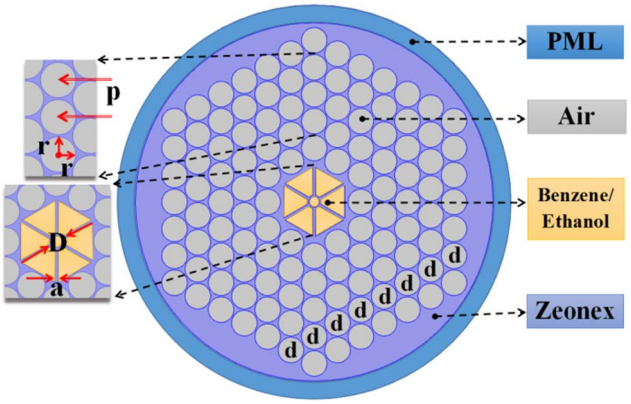
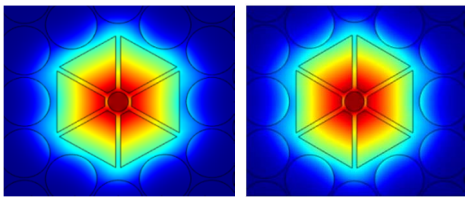


Fig. 1. Cross-sectional view of the proposed hybrid-core microstructure fiber.

III. Simulation Results and Discussion

To analyze the guiding properties full vectorial finite element method with Perfectly Matched Layer (PML) boundary condition is adopted with a FEM based software Comsol *V5.3a*. The mode field distribution of the proposed chemical sensor for benzene and ethanol at 1.5 THz is reported in Fig. 2 which indicates that the light signal is tightly confined through the central portion of the core. The following figure also ensures better interaction between the light signal and the test analyte which is essential for chemical detection.



(a) (b)

Fig. 2. Power flow distribution of the proposed hybrid-core fiber at 1.5 THz when core diameter is 170 µm for (a) benzene and (b) ethanol.

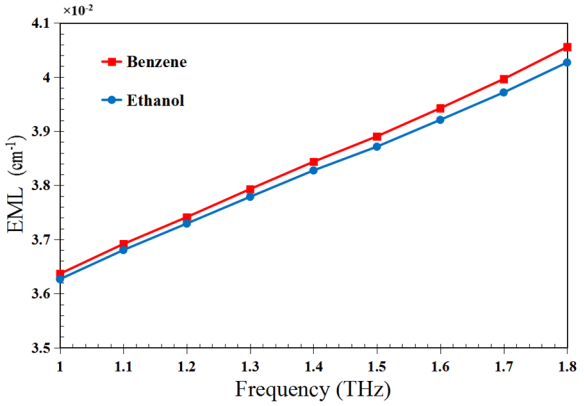


Fig. 3. Effective material loss of the proposed hybrid-core microstructure fiber as a function of frequency for D= 170 µm.

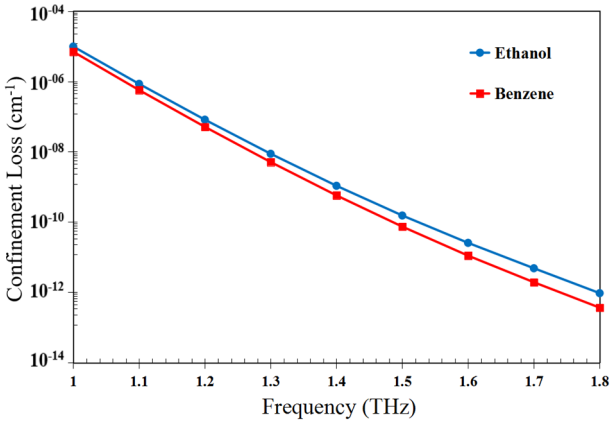


Fig. 4. Confinement loss of the proposed sensor for different operating frequencies at optimum core diameter.

At first, two types of losses will be investigated for the proposed chemical sensor and they are the effective material loss (EML) and the confinement loss. The term EML indicates the amount of light energy captured by the background material of the proposed PC-PCF. This loss can be evaluated by using the following mathematical expression [1],

$$\alpha_{eff} = \frac{\left(\varepsilon_0 / \mu_0\right)^{1/2} \int_{A_{mat}} n\,\alpha_{mat} |\mathrm{E}|^2 dA}{2\int_{All} S_z dA} \quad (1)$$

where ε_0 is for the permittivity of free space, μ_0 for the permeability of vacuum, n is the effective refractive index of the background material, E is the modal electric field, α_{mat} is the bulk material absorption loss and S_z is the z-component of the Poynting vector. The EML of the proposed chemical sensor for different operating frequency is reported in Fig. 3 for 170 µm core diameter. The reason behind the selection of particular core diameter will be reported in the further discussions. The following figure informs that the effective material loss rises with the increase of operating frequency. The physics behind that is the higher frequency electromagnetic signal always wants to propagate through the higher indexed zone. As the refractive index of Zeonex is higher than the analytes so the light of higher frequencies tries to travel through solid material than aqueous analytes. The EML of that proposed fiber are 0.0389 cm^{-1} 0.0387 cm^{-1} at 1.5 THz for benzene and ethanol respectively.

During the propagation of light signal through the core, a fraction of total light travels through the cladding air holes near the core and the air holes consume some of the light

energy. This type of loss is called confinement loss and this can be calculated by using the following expression [12],

$$\alpha_{CL} = 8.686 \times \frac{2\pi f}{c} \mathrm{Im}(n_{eff}) \quad (2)$$

where, c is the speed of the light in free space, f is for the frequency of the light and Im(n_{eff}) is the imaginary part of the refractive index of the guided mode. The confinement loss characteristics of the proposed PC-PCF is reported in Fig. 4 which shows that the variation of loss is inversely related with the frequency variation. This happens because higher frequency light travels through the core and the loss reduces. At D= 170 µm core diameter the confinement loss is 7.4×10^{-11} cm^{-1} which is comparable with [12-15].

Fig. 5 represents the EML and confinement loss of the proposed chemical sensor for different core diameter at 1.5 THz to determine the optimum core diameter. From the reported figure we can see that the EML and confinement loss intersects at 170 µm core diameter. That's why we choose 170 µm as optimum core diameter for the proposed chemical sensor. At optimum condition the numerical values of geometric parameters are **a**= 6 µm, **D**= 170 µm, **r**= 68 µm.

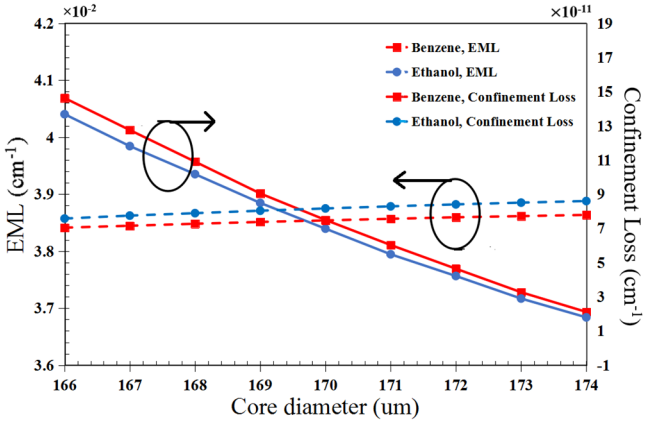


Fig. 5. EML and Confinement loss of the proposed sensor as a function of core diameter at 1.5 THz.

Now, the relative sensitivity of the proposed PC-PCF will be analyzed which is the most important characteristics of a sensor. The term relative sensitivity indicates the interaction between the electromagnetic signal and the analytes. The relative sensitivity of a sensor can be calculated by using the following equation [14],

$$r = \frac{n_r}{n_{eff}} \times P\,\% \quad (3)$$

where r is the relative sensitivity, n_r is the refractive index of the chemical to be sensed, n_{eff} is the effective refractive index of the guided mode and P is the power fraction of the guided mode. The power fraction indicates the amount of total light signal propagates through the core region. The power fraction can be calculated from the following equation [13],

$$P = \frac{\int_{sample} R_e(E_x H_y - E_y H_x)dxdy}{\int_{total} R_e(E_x H_y - E_y H_x)dxdy} \times 100 \quad (4)$$

where E_x, E_y and H_x, H_y are the transverse electric field and magnetic field of the guided light respectively. The integration of the numerator and denominator in Eqn. 4 are done for the liquid to be sensed in the core region and total area of the sensor respectively. The relative sensitivity of the hybrid core PCF as a function of frequency for benzene and ethanol are reported in Fig. 6 and Fig. 7 respectively for different geometric conditions. The relative sensitivity of the proposed sensor for both benzene and ethanol at optimum and optimum ± 2% condition is reported in the following figures. For optimum ± 2% condition all the geometric parameters (except PML thickness) are varied at the same percentage from optimum condition. From the following figure we can see that the relative sensitivity increases with the increase of operating frequency from 1 to 1.5 THz and after that the sensitivity decreases with the increase of frequency. So that 1.5 THz is considered as the optimum operating frequency for the proposed fiber. Both figure shows that the relative sensitivity is higher for optimum + 2% condition but the increased dimension of all air holes will create fabrication more complicated. The following graphical representations show that the relative sensitivity of the proposed sensor are 79.2% and 78.9% at 1.5 THz for benzene and ethanol which is better than the previous reported works in [12-15].

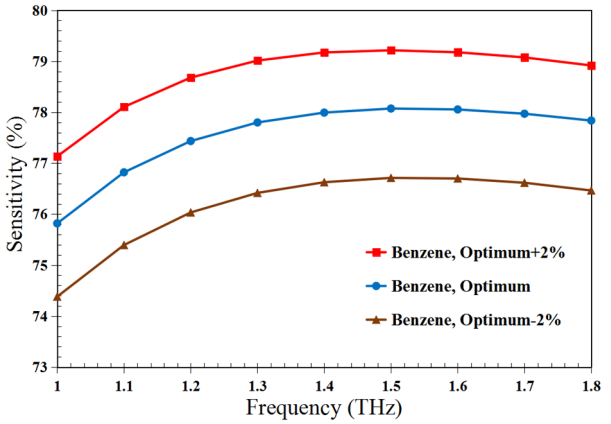


Fig. 6. Relative sensitivity of the proposed sensor at different operating frequencies for benzene at different geometric parameter.

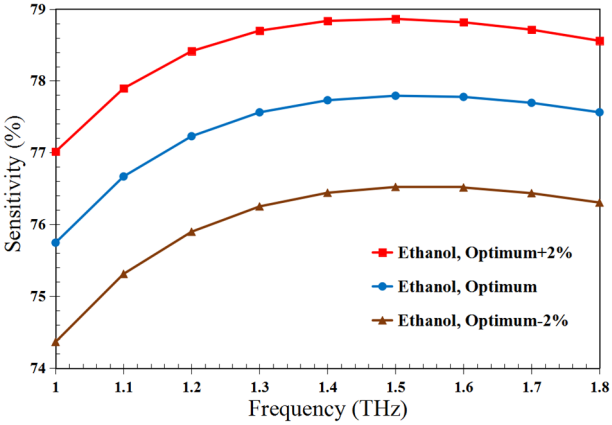


Fig. 7. Relative sensitivity of the proposed sensor at different operating frequencies for ethanol at different geometric parameter.

Now, another kind of loss of the proposed chemical sensor will be discussed which is important for the practical implementation of the sensor. During some practical applications of the fiber, it is required to bent and at the bended region a fraction of light come out from the central region. This type of loss is called bending loss and this can be calculated by using the following equation [5],

$$\alpha_{BL} = \frac{1}{8}\sqrt{\frac{2\pi}{3}}\frac{1}{A_{eff}}\frac{1}{\beta}F\left[\frac{2}{3}R\frac{\left(\beta^2-\beta_{cl}^2\right)^{3/2}}{\beta^2}\right] \quad (5)$$

where R is the bending radius, $F(x) = x^{-1/2}e^{-x}$,the propagation constant β and β_{cl} are defined as $\beta = 2\pi n_{co}/\lambda$ and $\beta_{cl} = 2\pi n_{cl}/\lambda$ respectively and A_{eff} is the effective area of guided mode light. The effective area can be can be evaluated by using the following equation [5],

$$A_{eff} = \frac{\left[\int I(r)rdr\right]^2}{\left[\int I^2(r)dr\right]^2} \quad (6)$$

where, $I(r) = |E_t|^2$ is the electric field of the fundamental guided mode. The effective area of the proposed chemical sensor at optimal designing parameter is reported in Fig. 8. The following figure reported that the effective area decreases with the increase of operating frequency because the light shows better confinement at higher frequencies and the effective area decreases. Fig. 9 shows the bending loss of the proposed sensor for R=1cm and R=2 cm at different operating frequencies for both benzene and ethanol. From the previously reported figure we can see that the bending loss of the proposed sensor is 4.9×10^{-36} cm^{-1} and 3×10^{-37} cm^{-1} for benzene and ethanol at optimum designing condition when the operating frequency is 1.5 THz.

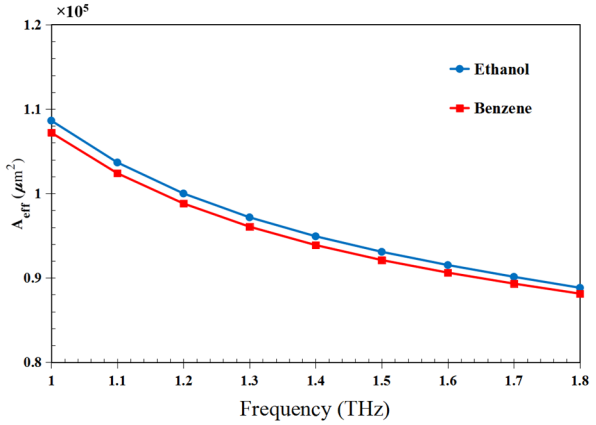


Fig. 8. Effective area of the proposed PC-PCF for different operating frequency at **D**=170 µm.

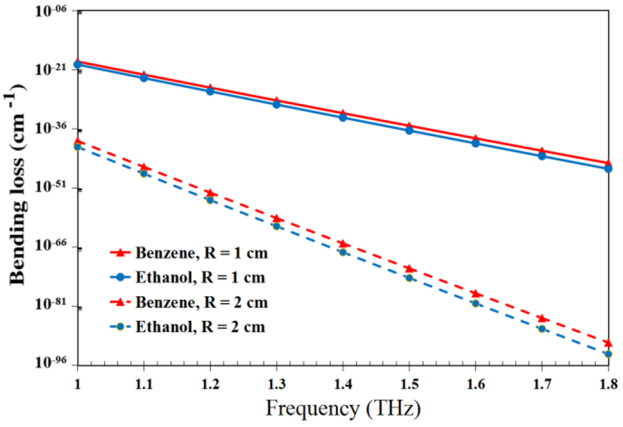


Fig. 9. Bending loss of the proposed chemical sensor as a function of frequency for D = 170 µm

When digital light signal propagates through the optical waveguide, due to the material of the waveguide and the orientation of the air holes the pulse width increases and it is called dispersion. As the effective refractive index of Zeonex is almost constant so only waveguide dispersion of the proposed sensor will be investigated. The dispersion characteristics of any microstructure fiber can be easily calculated by using the following expression [5],

$$\beta_2 = \frac{1}{c}\left(2\frac{dn}{d\omega}+\omega\frac{d^2n}{d\omega^2}\right) \quad (7)$$

where $\omega=2\pi f$ and c is the velocity of light in free space and f is the frequency of the operating light signal. The dispersion characteristic of the proposed microstructure fiber is shown in fig. 10. The graphical representation confirms that the dispersion of the proposed fiber is 0.56± 0.12 ps/THz/cm for benzene and ethanol from 1.4 to 1.8 THz.

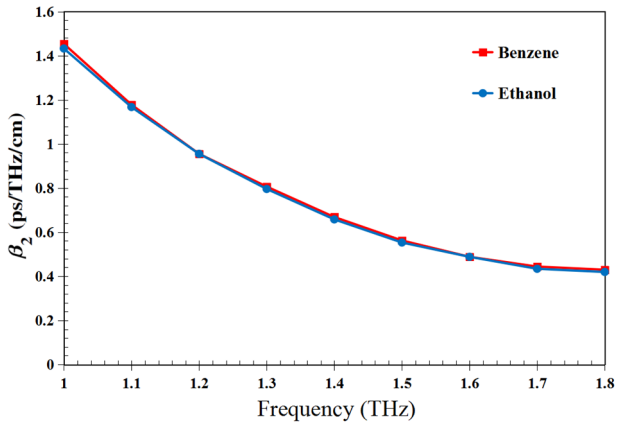


Fig. 10. Dispersion of the proposed chemical sensor for different analytes as a function of frequency at **D**= 170 µm.

At last, the possible fabrication technique of the proposed chemical sensor design. In the present age, various fabrication techniques (drilling and stacking, Extrusion, fiber drawing etc.) have invented and by using these methods various fiber can be fabricated easily and effectively. For the proposed microstructure fiber we propose Extrusion technique [16, 17] to fabricate the fiber because by using this method any kind of structure can be fabricate easily. Finally, a comparison of our proposed fiber with other remarkable chemical sensor is reported in Table. 1 and from the following table we can see that our proposed fiber is better than the reported sensor in the literature.

TABLE 1. COMPARISON OF OUR PROPOSED FIBER WITH SOME OTHER REMARKABLE DESIGNS

Ref.	Year	Operating Range	Relative Sensitivity [%]	Confinement loss [dB/m]
[12]	2016	1.33 µm	49.7	2.75×10^{-10}
[13]	2017	1.33 µm	53.35	1.6×10^{-10}
[14]	2016	1.33 µm	59.07	10^{-11}
[15]	2017	1.55 µm	65.18	5×10^{-9}
Proposed	---	**1.5 THz**	**79.2**	**3.2×10^{-9}**

IV. CONCLUSION

A new kind of porous core microstructure fiber for chemical identification is proposed and analyzed in this paper using terahertz light signal. The numerical results confirms that the proposed fiber offers very high relative sensitivity of almost 79%, very low confinement loss of 3.2×10^{-9} dB/m, very low bending loss of 4.9×10^{-36} cm^{-1} simultaneously at optimum designing parameter. The proposed fiber is very simple to realize and can be easily fabricated by using the Extrusion technique. We hope, this

sensor will be a handy one in different chemical sensing applications.

REFERENCES

[1] M. A. Habib, M. S. Anower, and M. R. Hasan, "Ultrahigh birefringence and extremely low loss slotted core microstructure fiber in terahertz regime," Current Opt. and Photon., vol. 1, no. 6, pp. 567-572, Dec 2017.

[2] M. Walther, B. M. Fischer, A. Ortner, A. Bitzer, A. Thoman, and H. Helm, "Chemical sensing and imaging with pulsed terahertz radiation,"Anal. Bioanal. Chem., vol. 397, no. 3, pp. 1009–1017, 2010.

[3] Y. Yu *et al*., "Some features of the photonic crystal fiber temperature sensor with liquid ethanol filling,"Opt. Express, vol. 18, no. 15, pp. 15383- 15388, 2010.

[4] H. Y. Fu *et al*., "High pressure sensor based on photonic crystal fiber for downhaul applications," Appl. Opt., vol. 49, no. 14, pp. 2639-2643, 2010.

[5] M. A. Habib, M. S. Anower, and M. R. Hasan, "Highly birefringent and low effective material loss microstructure fiber for THz wave guidance," Opt. Commun., vol. 423, pp. 140-144, 2018.

[6] W. Withayachumnankul*et al*., "T-ray sensing and imaging," Proc. IEEE, vol. 95, no. 8, pp. 1528–1558, Aug. 2007.

[7] C.B. Reid, et al., "Terahertz pulsed imaging of freshly excised human colonic tissues," Phys. Med. Biol., vol. 56, no. 14, pp. 4333–4353, 2011.

[8] T. Kiwa *et al.*, "Chemical sensing plate with a laser-terahertz monitoring system," *Appl. Opt.*, vol. 47, no. 18, pp. 3324–3327, 2008.

[9] H. Ademgil, "Highly sensitive octagonal photonic crystal fiber based sensor," *Optik-Int. J. Light Electron Opt.*, vol. 125, no. 20, pp. 6274–6278, 2014.

[10] H. Ademgil and S. Haxha, "PCF based sensor with high sensitivity, high birefringence and low confinement losses for liquid analyte sensing applications," *Sensors*, vol. 15, no. 12, pp. 31833–31842, 2015.

[11] M. B. Pushkarsky, M. E. Webber, O. Baghdassarian, L. R. Narasimhan, and C. K. N. Patel, "Laser-based photoacoustic ammonia sensors for industrial applications," *Applied Physics B*, vol. 75, no. 2-3, pp. 391-396, 2002.

[12] S. Asaduzzaman, K. Ahmed, T. Bhuiyan, and T. Farah, "Hybrid photonic crystal fiber in chemical sensing," SpringerPlus, vol. 5, no. 748, pp. 1-11, 2016.

[13] M. F. H. Arif, and M. J. H. Biddut, "A new structure of photonic crystal fiber with high sensitivity, high nonlinearity, high birefringence and low confinement loss for liquid analyte sensing applications," Sens. and Biosens. App., vol. 12, pp. 8-14, 2017.

[14] M. F. H. Arif, K. Ahmed, S. Asaduzzaman, and M. A. K. Azad, "Design and optimization of photonic crystal fiber for liquid sensing applications," Photonic Sens., vol. 6, no. 3, pp. 279-288, 2016.

[15] B. K. Paul, K. Ahmed, S. Asaduzzaman, and M. S. Islam, "Folded cladding porous shaped photonic crystal fiber with high sensitivity in optical sensing applications: Design and analysis," Sens. and Biosens. App., vol. 12, pp. 36-42, 2017.

[16] H. Ebendorff-Heidepriem, and T.M. Monro, "Extrusion of complex preforms for microstructured optical fibers," *Opt. Express*, vol. 15, no. 23, pp. 15086–15092, 2007.

[17] M. A. Habib, M. S. Anower, "Design and numerical aanlysis of highly birefringent single mode fiber in THz regime," Opt. Fib. Techn., vol. 47, pp. 197-203, 2019.

Single Mode Flattened Dispersion Circular-Shaped Porous Core Photonic Crystal Fiber for THz Regime

Md. Anwar Sadath
Department of EEE
Bangladesh University of Engineering and Technology
(BUET)
Dhaka-1216, Bangladesh
bidduth08@yahoo.com

Mohammad Saiful Islam
Department of EEE
Bangladesh University of Business and Technology
(BUBT)
Dhaka-1216, Bangladesh
saiful.eee14@gmail.com

Md. Hafijur Rahman
Department of EEE
Bangladesh University of Business and Technology
(BUBT)
Dhaka-1216, Bangladesh
hafijurrahman5331@gmail.com

Md. Tariqul Islam
Department of EEE
Bangladesh University of Business and Technology
(BUBT)
Dhaka-1216, Bangladesh
tariqulshahin33@gmail.com

Sabbir Ahmed Khan
Department of EEE
Bangladesh University of Business and Technology
(BUBT)
Dhaka-1216, Bangladesh
sabbirahmedkhan.bubt@gmail.com

Shahadat Hossain
Department of EEE
Bangladesh University of Business and Technology
(BUBT)
Dhaka-1216, Bangladesh
shpatwary2016@gmail.com

***Abstract*— The proposed fiber for THz propagation demonstrates that it has flattened dispersion with effective material loss of 0.063 cm^{-1} and confinement loss of 2.57×10^{-10} cm^{-1} . The dispersion loss of this proposed fiber is 0.5305 ± 0.05 ps/THz/cm. In this design we have demonstrated that the proposed fiber has low absorption loss due to its background material. In this design we also have discussed about other modal characteristics such as v-parameter, power fraction and also fabrication possibilities. This proposed fiber can be used for efficient and convenient transmission of THz waves.**

***Keywords*—** *PCF, Dispersion, Confinement loss, Effective material loss, THz waveguide*

1. INTRODUCTION

THz radiation, sometimes known as sub-millimeter-wave radiation, formed of electromagnetic waves between the range of 0.1-10 THz band [1]. Since THz waves operates between the infrared and microwave ranges, it carries the characteristics from both of them. Now-a-days, THz PCF is getting so much attention due to their convenient use in multidisciplinary fields like imaging [2], remote sensing [3], biomedical engineering [4]. Due to their low absorption loss and low dispersion, THz fiber is being used in biomedical engineering for different purposes like early diagnosis of skin cancer [4] , diagnosis of breast cancers [5], and diagnosis of the colon tissue cancers [6]. But THz wave communication faces various types of problems because it operates in free space medium. The whole THz system uses three parts (THz source, waveguide and detector) which are already accessible for commercial use.

Strong water vapor absorption and the misalignment issues of THz PCF are the major obstacles in THz communication. Researchers are working very hard to make THz communication very convenient with low loss. Dielectric metal - coated tube, Bragg fibers, Metallic wires were used as waveguides. But most of the fields propagates through outside of the waveguide core because these waveguides uses dielectric subwavelength and metallic wires. Now-a-days various polymers are being used as base material like PMMA, Teflon, and TOPAS for THz propagation due to their low absorption loss. Flattened dispersion is desirable in long distance THz wave communication. TOPAS is being used as base material for THz propagation due to their low absorption loss than other materials. Though a wide number PCF have been proposed in recent years, one of the researchers named Atakaramians et al. worked on dispersion properties of a porous core fiber to have low loss transmission through photonic crystal fiber [7]. Therefore porous core photonic crystal fiber is getting greater attention because of the design structure, the waveguide parameters such as core diameter, air hole radius, pitch size and air filling fraction can be determined. In this proposed fiber, we exposed a new design of PCF with circular shaped core with rectangular air hole inside the core which has flattened dispersion, low confinement loss and low effective material loss.

2. DESIGN STRUCTURE AND RESULTS

The physical structure of the proposed PCF is shown in Fig. 1. The finite element based software COMSOL 5.3a is used to make the design of the proposed waveguide. The proposed PCF is consisted of seven rectangular shape air-holes inside the core to obtain low flattened dispersion and low effective material loss (EML). In our simulation, we have used finer mesh to have the better accuracy for our design. In our proposed design, the optimum core diameter (D_{core}) is 350 μm and operating frequency is 1 THz, because the results are better. In Fig. 1, the cladding region is formed of different rectangular shaped air holes whose lengths are defined as L, L_1, L_2, L_3 respectively. The defined lengths in the proposed fibers are 683 μm, 633 μm, 501 μm and 292 μm respectively. Here d is denoted as the center to center distance of the rectangular air hole inside the core which is around 70 μm. We have used Topas as a background

material which is commercially known as cyclic olefin copolymer (COC) in our proposed PCF. It has refractive index of 1.5258 and it is constant over 0.1-2 THz wave propagation. The best advantages of our proposed PCF is its simplicity and low dispersion and confinement loss.

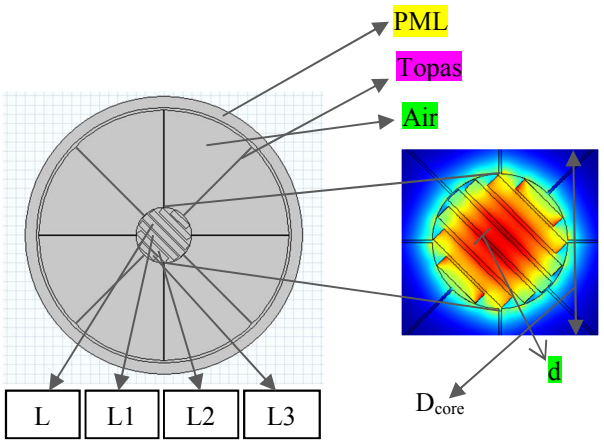


Fig. 1. Cross sectional view of the proposed PCF

THz wave propagation mostly suffers from effective material loss or material absorption loss. EML is the major loss of any proposed fiber that can be calculated by the following expression [8],

$$\alpha_{eff} = \sqrt{\frac{\varepsilon_0}{\mu_0}} \left(\frac{\int_{mat} n_{mat|E^2|\alpha_{mat} dA}}{|\int_{all} S_z dA|} \right) \quad (1)$$

Here, n_{mat}, α_{mat}, ε_0, μ_0 are the refractive index, bulk material absorption loss, permittivity and the permeability of the free space. From our proposed PCF we have found the material loss of 0.063 cm^{-1} at 1 THz frequency that is shown in fig. 2. Fig. 3. Also shows the EML variation at different D_{core} at 1.0 THz frequency. EML increases with the increase in D_{core} because of increase in material in the core region.

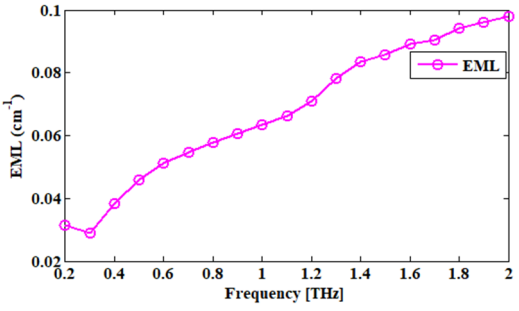


Fig. 2. EML variation at different frequency

Another parameter that is also important for any proposed PCF is confinement loss. It generally depends on the core porosity of the fiber and be calculated by the following expression [9],

$$L_c = 8.686 \left(\frac{2\pi f}{c}\right) Im(n_{eff}) \frac{dB}{m} \quad (2)$$

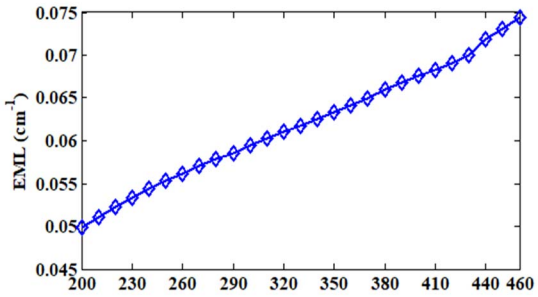


Fig. 3. EML variations at different core diameter

Here, $Im(n_{eff})$ denotes the refractive index and f denotes the operating frequency. From fig. 4. We can see that , the confinement loss is decreasing with the increase of the core diameter. At optimum parameters, we have found that the confinement loss is only 2.57 × 10^{-10} cm^{-1}.

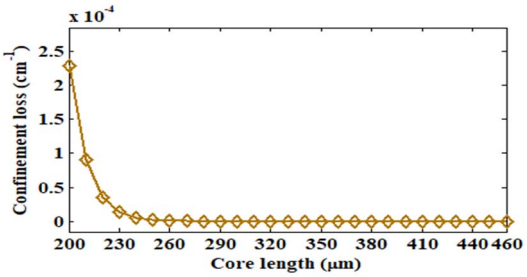


Fig. 4. Confinement loss as a function of frequency at D_{core}=350 μm

Another important parameter for a PCF is power fraction. From the power fraction we can find percent of the light intensity through different region of any fiber. From fig. 5, we can see, the light propagation intensity through different medium at [0.2-1.8] THz frequency range. Power fraction of a photonic crystal fiber can be calculated from the following equation,

$$n' = \frac{\int_x S_z dA}{\int_{all} S_z dA} \quad (3)$$

Where X represents the core, cladding or material region. In our proposed fiber, we have calculated the power fraction (light intensity) through core, cladding and material regions. At optimum operating parameters above 65% light passes through the core air-holes which is much better result compared to recently proposed models.

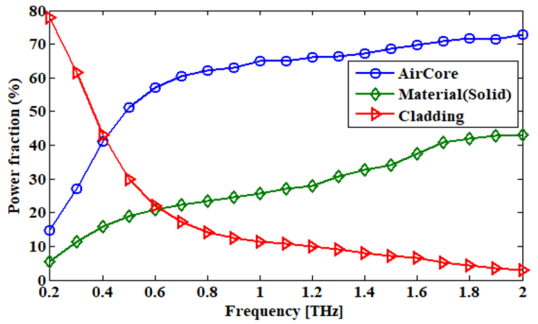


Fig. 5. Power fraction (%) at different frequencies for D_{core}=350 μm

We also have discussed about the waveguide dispersion. Due to consideration of the dispersion, we have given a great attention to the structure of our proposed PCF. The waveguide dispersion of PCF can be obtained by the following equation [9],

$$\beta_2 = \frac{2}{c}\frac{d_{n_{eff}}}{dw} + \frac{w}{c}\frac{d^2 n_{eff}}{dw^2} \qquad (4)$$

From our simulation, we have obtained a nearly zero flat dispersion of 0.5305 ± 0.05 ps/THz/cm in THz frequency range of 1.0 - 1.8 THz which is shown in fig. 6. In our proposed PCF, we have introduced rectangular air shape inside the core to have flat dispersion.

We have designed our fiber for the single mode propagation and that can be calculated by using V-parameter. V-parameter can be expressed by the following expression [10],

$$V = \frac{2\pi r f}{c}\sqrt{n_{co}^2 - n_{cl}^2} \leq 2.405, \qquad (5)$$

Here, f is the operating frequency, c is the velocity of light, n_{co} and n_{cl} indicates the refractive index of the core and cladding. We have assumed n_{cl} = 1 because of suspended core. From fig.7. we can see that single mode always exists up to 1.7 THz.

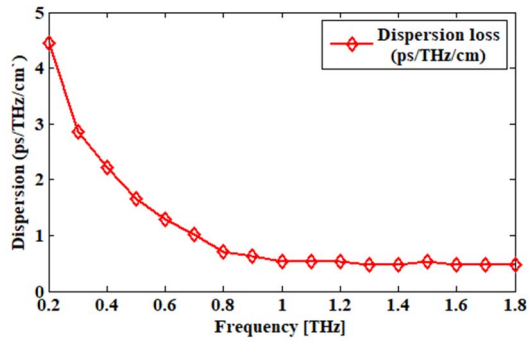


Fig. 6. Waveguide dispersion versus frequency at D_{core} = 350 µm

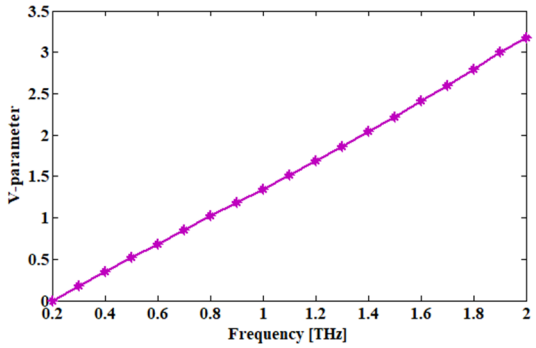


Fig. 7. V-parameter at different frequency for D_{core} = 350 µm

V-parameter as a function of core diameter is also shown in fig. 8. Our proposed PCF acts as a single mode fiber up to 440 µm.

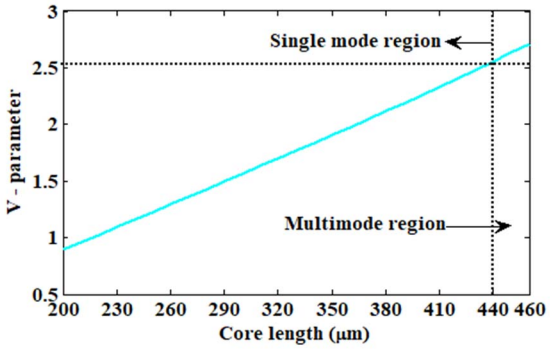


Fig. 8. V parameter of the proposed PCF as a function of core diameter at 1.0 THz frequency

Finally, fabrication is an important issue for every PCF. There are a lot of methods available to fabricate PCFs commercially. Sol-gel, drilling and drawing, stack and draw, extrusion, 3D technology are such methods. As our proposed fiber contains non-circular air-holes, extrusion technique can be used to fabricate our design. In addition, micro-structured polymer optical fiber with asymmetric air holes is also feasible using 3D printing technology

3. CONCLUSION

In summary, a completely new circular shaped air core with rectangular shaped air-hole PCF is proposed in this paper. Since Topas has low bulk material absorption loss and zero material dispersion, it is used as the background material in our design. A thorough numerical investigation has been used which shows an EML of 0.063 cm^{-1} , low confinement loss of 2.57 × 10^{-10} cm^{-1} for 350µm core diameter at 1.0 THz operating frequency. A nearly zero flat dispersion of 0.5305 ± 0.05 ps/THz/cm in frequency range of 1.0 - 1.8 THz is also obtained. The fabrication of this proposed PCF is feasible because of its realistic size and simplicity. Since this proposed fiber has low loss with flattened dispersion, so this PCF can be a potential candidate for applications in long distance THz wave propagation.

REFERENCES

[1] D. Abbott and X.-C. Zhang, "Scanning the issue: T-ray imaging, sensing, and retection," Proceedings of the IEEE 95, 1509–1513 (2007).

[2] K. Fukunaga, N. Sekine, I. Hosako, N. Oda, H. Yoneyama, and T. Sudoh, "Real-time terahertz imaging for art conservation science," J. Eur. Opt. Soc. Rapid Commun. 3, 08027/1–4 (2008)

[3] Liu, J., Dai, J., Chin, S. L. & Zhang, X.-C. Broadband terahertz wave remote sensing using coherent manipulation of fluorescence from asymmetrically ionized gases. *Nature Photon.* **4**, 627–631 (2010).

[4] R. M. Woodward, V. P. Wallace, D. D. Arnone, E. H. Linfield, and M. Pepper, "Terahertz pulsed imaging of skin cancer in the time and frequency domain," J. Biol. Phys. 29(2/3), 257–259 (2003).

[5] P. C. Ashworth, E. Pickwell-MacPherson, E. Provenzano, S. E. Pinder, A. D. Purushotham, M. Pepper, and V. P. Wallace, "Terahertz pulsed spectroscopy of freshly excised human breast cancer," Opt. Express 17(15), 12444–12454 (2009).

[6] C. B. Reid, A. Fitzgerald, G. Reese, R. Goldin, P. Tekkis, P. S. O'Kelly, E. Pickwell-MacPherson, A. P. Gibson, and V. P. Wallace, "Terahertz pulsed imaging of freshly excised human colonic tissues," Phys. Med. Biol. 56(14), 4333–4353 (2011).

[7] S. Atakaramians, S. Afshar, V., B. M. Fischer, D. Abbott, and T. M. Monro, "Porous fibers: a novel approach to low loss THz waveguides," Opt. Express 16, 8845–8854 (2008)

[8] S. Islam, M.R. Islam, M. Faisal, A.S.M.S. Arefin, H. Rahman, J. Sultana, S. Rana, Extremely low-loss, dispersion flattened porous-core photonic crystal fiber for terahertz regime, Opt. Eng. 55 (7) (2016) 076117.

[9] R. Islam et al., "Low-loss rotated porous core hexagonal single-mode fiber in THz regime," Opt. Fiber Technol. 24, 38–43 (2015).

[10] N. Chen, J. Liang, and L. Y. Ren, "High-birefringence, low-loss porous fiber for single-mode terahertz-wave guidance," Appl. Opt. 52, 1641–5302 (2013).

[11] P. I. Antonov and V. N. Kurlov, "New advances and developments in the Stepanov method for the growth of shaped crystals," Crystallography Rep., vol. 47, no. 1, pp. S43–S52, Jan. 2002.

Optimization and energy management of hybrid renewable power generation using HOMER and FLC

Nesar Uddin
Dept. of IET, CUET
Chittagong, Bangladesh.
nesar15cuet@gmail.com

Sayed Abu Nahian
Dept. of IET, CUET
Chittagong, Bangladesh.
nahians@cuet.ac.bd

Md. Saiful Islam
Dept. of ETE, CUET
Chittagong, Bangladesh.
saiful05eee@gmail.com

Jewel Sikder Joy
Dept. of IET, CUET
Chittagong, Bangladesh.
jewelsikderjoyeee@gmail.com

Abstract—**This research paper proposes an energy generation system to utilize renewable energy for off-grid area of Bangladesh which includes hill track areas and Rohinga camps. Most of the hill track areas and Rohinga camps did not access electricity. Again, present system loss of electricity is a great problem in Bangladesh. The main purpose of the paper is to design an optimal hybrid energy management system using HOMER control technique and Fuzzy Logic Controller (FLC). The system consists of PV arrays, Biogas generator, diesel generator, converter and Battery. We have used the HOMER software for effective size calculation of the system components. FLC is used to manage the obtained power from renewable energy sources efficiently. Optimum fuzzy rules are designed to control the shifting of the switches considering the load profiles and availability of renewable sources on a given period of time. For the justification of cost effectiveness and power management system simulation result from our control strategy have been compared with other prevailing systems. In addition, we obtain lower CO_2 emission from the system renewable fraction which indicates the environment friendliness. This simulation result provides a summary to utilize renewable energy with properly on recent energy crises in Bangladesh.**

Keywords—**Hybrid energy generation system; PV array; Biogas; Hybrid optimization of multiple energy resources (HOMER); Fuzzy logic controller (FLC);**

I. INTRODUCTION

The capacity of installed electricity generation in Bangladesh is about 15,755 MW in 2017 which does not exist for total population [1]. About 71.42% electrical energy comes from natural gas followed by furnace oil 15.44%, diesel 2.91%, and imported 5.37% those are limited [1]. Since, energy cannot be created or destroyed, so the source of non-renewable energy will run out very rapidly. It also increases air, water and soil pollution, greenhouse gases with the amount of carbon dioxide emission that attributes to global warming. Again energy consumption is increasing day by day. On the other hand, most of the hill track and remote areas are out of electricity. So, energy generation from renewable sources is the best choice for protecting the environment and fulfilling the total load demand of those areas. It is a solution towards the limited availability of fossil fuel and climate change.

There are many kinds of renewable sources such as solar energy, hydro energy, wind energy, tidal energy, geothermal energy and biomass energy etc. Among them Solar and biogas energy are more suitable for electricity generation in hill track areas, remote areas and Rohinga camps in Bangladesh. Present share of renewable energy in Bangladesh is only 1% which envisions having 10% by 2020 which is shown in Table I [2]. The target of our government about electricity generation from renewable sources for domestic house is 2,470MW within 2021, and 3,864MW within 2041. Again, according to the view of government, the production of biogas from domestic house is 600,000 m^3/day by 2031 which will be increased to 3 million m^3/day by 2041 [2]. In Bangladesh, cattle dung is available from 11.6 million cows, buffaloes, goats and sheep [3]. One kilogram of dung can produce .037m^3 biogas. Therefore, available dung of cattle and buffaloes can produce $2.97x10^8$ m^3 of biogas which is equal to $1.52x10^6$ tons of kerosene or $3.04x10^6$ tons of coal [3]. So, generation of electricity from biogas is an enormous source in Bangladesh. Besides, a substantial amount of biogas can be produced from human and animal excreta, garbage water hyacinth and forest residue. Therefore, the combination of solar energy and biogas energy can play a vital role to fulfill the load demand for those areas. This paper is offered for multi households with 220 kW load demand in a day. The available average solar radiation is 4-7kWh per square meter per day on that area with near 250–300 sunshiny days in a year that can eliminate a precent of the total load demand in this area [4]. Again the fuel of biogas energy is also available in this area which can mitigate a large amount of load demand of required energy.

TABLE.I RENEWABL EENERGY STATUS IN BANGLADESH [2].

Category	Achievement (MW)	Target (MW)
Solar Energy	More than 300	1800
Wind Energy	2	364
Biomass Energy	Less than 1	45
Biogas Energy	Less than 5	9
Hydro Energy	230	237
Other RE	Less than 1	15
Total	539	2,479

To minimize the system loss and increase efficiency, effective energy management system is most important. So, an Artificial Intelligence (AI) as fuzzy logic controller in MATLAB is one of the most widely used applications of fuzzy set theory that is used for management system which must minimize power loss [5]. This management system is one of the more effective systems that optimize energy efficiency. The control strategy based on fuzzy logic theory has been proposed to achieve the optimal result of the battery charging/discharging, PV operating condition, biogas generator and diesel generator off-on performance.

II. METHODLOGY

A. STUDY AREA & RESOURCES ASSESSMENT

In this research, study area has been selected for Bashkhil in Boalkhali Chittagong that is located at 22 degrees 50 minutes North latitude and 92 degrees 22 minutes East

longitude [6]. It is a hill track area in where solar and biogas fuel is available and most of the people are not access of electricity. About 140-150 family can get benefit of electricity from this study by utilizing renewable energy for 220 kWh load demands for a day. Overall research method has been presented in Fig. 1.

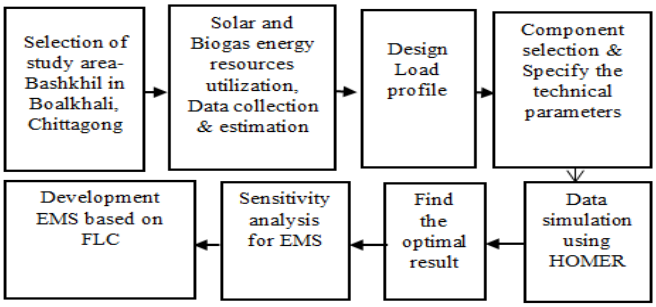


Fig.1 The overall methodological outline of the research.

For this research typical monthly solar radiation data has been collected from NASA by using Homer that is shown in Table II.

TABLE II. AVERAGE SOLAR RADIATION AND BIOMASS FUEL IN PROPOSED AREA [6].

Month	Clearness Index	Daily Solar Radiation (kWh/m²/day)	Daily Biomass (ton/day)
January	0.641	4.523	1.5
February	0.616	5.005	1.2
March	0.605	5.699	1.3
April	0.552	5.781	1.2
May	0.528	5.807	1.2
June	0.390	4.344	1.2
July	0.362	3.987	1.1
August	0.403	4.272	1.3
September	0.417	4.064	1.2
October	0.575	4.874	1.3
November	0.597	4.345	1.3
December	0.649	4.350	1.4
Average	0.641	4.523	1.26

By investigation the study area, total number of the cows is 50. Dung potential for total cows is 500 kg/day. Domestic waste potential is 250 kg/day and forest residue is 400 kg/day. Total compost manufacture is 1500 kg/day. From this manure, the potentiality of biogas is 75 m^3 /day which are shown in Table III. Biogas produces by Photosynthesis process. Photosynthesis can be described by flowing chemical reaction [7]:

$C_6H_{12}O_6 + 6\ O_2 \text{ ---> } 6\ CO_2 + 6\ H_2O + 19\ MJ\ /m^3$ ……(1)

Where, $C_6H_{12}O_6$ represents biomass resource.

TABLE III. BIOGAS RESOURCESS UTILIZATION OF PROPOSED AREA [6]

Specifications	Details
Total Domestic Waste	250kg/day
Total cow	50
Forestry residue	400kg
Dung for one cow	10 kg/day
Overall dung	500 kg/day
Other manure	350 kg/day
Gas profit	.05 m^3/kg
Total amount of biogas	75 m^3
1 m^3	19 MJ
Total electricity produced	1425 MJ = 395 kWh
Efficiency of Generator	35%
Overall electricity produced	138 kWh/day

B. SYSTEM DESCRIPTION

The proposed hybrid system consists of PV arrays, Biogas generator, Diesel generator, converter and batteries which is shown in Fig. 2. According to load demand, PV array and Biogas generator, Diesel generator supply electricity. The extra energy from PV and Biogas generator is stored in the battery until full charge. When generation of electricity from PV and Biogas generator, Diesel generator is unavailable, then battery is used as backup system. The converter is recovered as an inverter and rectifier.

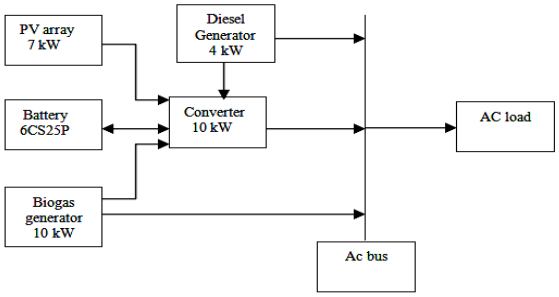


Fig.2 Block diagram of the system.

C. PV array

Photovoltaic (PV) array produce electrical energy through semiconductor materials from light energy based on a phenomenon called photovoltaic effect [8]. Generated power from PV array can be denote by below equation [9]

$$P_{pv}(t) = \eta_{pv} . A_{pv} . S(t) \text{..... ....} (2)$$

Where, η_{PV} – efficiency PV arrays, A_{PV} - total area of PV array in m^2 and S(t) - solar radiation kW/m^2. Average monthly solar radiation from January to December has been presented in Fig. 03.

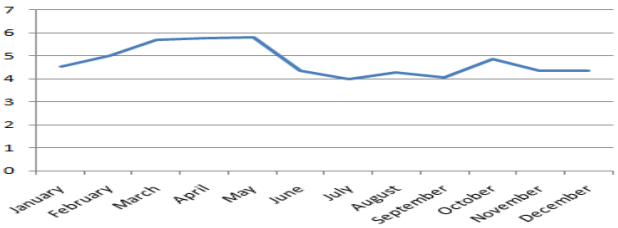


Fig. 3 Average monthly solar radiation.

For proposed system, 350 watt of 72pic Mono crystalline Silicon SKT350P-12 PV array has been considered. By optimizing, we have got 7 kW PV arrays according to HOMER software. The capital cost, replacement costs and lifetime of PV arrays have been considered $700, $600 and 25 years respectively which is shown in Table VII.

D. Biogas model

Biogas generally is the combination of different gases like (CH_4, CO_2, H_2S) that composed from farming waste, manure, municipal waste, forest deposits, plant material, sewage, green waste or food waste by the breakdown of organic matter in the absence of oxygen [10]. Biogas is a kind of renewable energy source because of it's created from atmospheric carbon by plants grown within current growing seasons. Generated gas is used as fuel that can be used as heating purpose or cooking purpose. This gas can also be used to generate electricity through gas engine. General specification of biogas generator which is shown in Table V. Fig.4 indicates monthly average biomass fuel in ton.

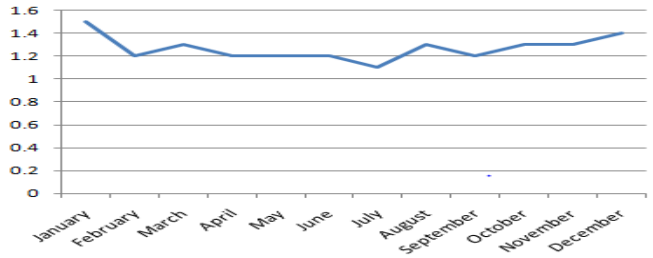


Fig. 4 Monthly average biomass fuel in ton.

TABLE V. GENERAL SPECIFICATIONS OF BIOGAS GENERATOR [11]

Specifications	Details
Model	YDNY-10
Output	AC 3 phase, 50 Hz
Rated Power	10kW/12.5Kva
Rated Current	18 A
Fuel	Biogas
Gross Weight	510 kg
Rated Speed	1500 rpm
Gas Consumption	6.5 m^3/h
No. of Cylinder	4

The capital cost, replacement cost and operation time of biogas generator is considered $1100, $1000 and 20,000 hours which is shown in Table VII.

E. Generator model

A diesel generator is a machine that connects with a diesel engine with an electric generator to produce electricity by rotating the alternator [12]. The working principle of diesel engine depends on Carnot cycle. Again, the mechanism of electric generator depends on principle of electromagnetic induction. General specification of diesel generator which is shown in Table VI.

TABLE VI. GENERAL SPECIFICATIONS OF DIESEL GENERATOR [13]

Specifications	Details
Model	POWERHOUSE, PH3826
Output power	AC, 3 phase, 50 Hz
Rated Power	4.8kW
Fuel	Diesel
Fuel volume	16 L
Rated Speed	3400 rpm
Oil Consumption	1.9 L/h

The capital and replacement cost of diesel generator is considered $150 and $130 for 1 kW which is shown in Table VII. The lifetime of diesel generator is considered 30,000 hours.

F. Battery model

Battery is used to stock extra energy from renewable sources that will be used when the system is not capable to give required power. The lifetime of battery depends on charge or discharge rate per day. By optimizing input data, we have got the number of battery having capacity 11 of 6CS25P model 6 volt 820Ah 6.94kWh. The capital and replacement cost of the battery is considered $1100 and the replacement cost $1000 for five years [14].

G. Converter model

Converter is used as both rectifier (AC to DC) and inverter (DC to AC). By simulation, the program finds an optimum result 11kW for converter. The capital cost and replacement cost of converter is considered $180 and $150 per kW for 15 years which shown in Table VII [14].

H. Load profiles

TABLE VII. THE INPUT HYBRID SYSTEM COMPONENTS.

Quantity	Size (kW)	Capital ($) for 1 kW	Replacement ($) for 1 kW	O & M ($) per year	Life Time
PV arrays	7	700	600	10	25 year
Biogas generator	10	1100	1000	1539	20000 hour
Diesels Generator	4	150	130	233	30000 hour
Battery	6.94kWh	1100	1000	8	5 year
Converter	10	180	150	5	15 year

The target of load demand is 220 units per day for selected zone. Daily load profile is presented in Fig. 05 (summer season) and Fig. 06 (winter season). Day to day load variation has been considered 16% and time to time load variation has been considered 22%.

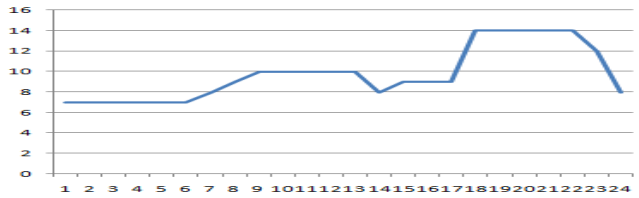


Fig. 5 proposed daily load profile for March to November in kW.

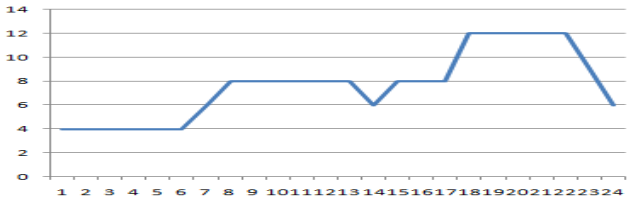


Fig. 6 proposed daily load profile for December to February in kW.

This proposed system is designed for four or five family member with 150 to 170 families. We consider typical electrical appliances such as CFL bulbs, electric Fan, TV, electronics charges in each family with standard voltage (220V AC, 50Hz). Table VIII presents the electric load characteristics of each house.

TABLE VIII. ELECTRIC APPLIANCES OF HOUSE.

Domestic Appliances	Number of Appliances	Power in watt	Summer (Avg. hr./day)	Winter (Avg. hr./day)	Approximate Total (kWh/day)	
					Summer	Winter
CFL Bulbs	3	20	9/10	10/11	1.508 kWh/day	.738 kWh/day
Ceiling Fan	2	35	10/11	0		
TV	1	18	4/5	4/5		
Mobile Charger	2	6	3/4	3/4		

III. OPTIMIZATION MODEL USING HOMER

Total Architectural view of research has advanced with the load profile, solar & biomass resources, system component as input of HOMER. Net present cost, optimum result, cash flow, and total capital cost, energy saving cost, renewable energy fraction, and excess energy fraction% as output by Homer. The configurations of proposed system's component are PV array, biogas generator, diesel generator, battery and power converter. PV array and Biogas generator are main renewable sources of electricity. Generated current from PV

array is DC. Again we used AC biogas generator. So, inverter is used to change DC to AC and vice versa to adjust overall system. The simulation diagram of proposed system is shown in Fig. 7.

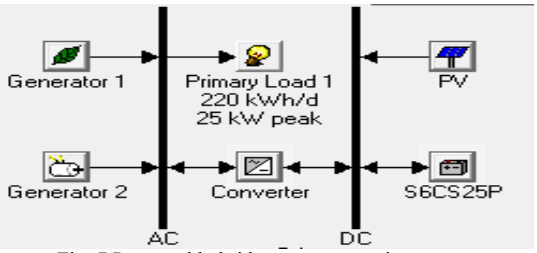


Fig. 7 Proposed hybrid energy generation system.

IV. POWER MANAGEMENT MODEL USING FLC

A fuzzy inference system (FIS) is a kind of artificial intelligence system that is created on a fuzzy set and IF-THEN rules [4]. Fuzzy logic can make ideas that not only "true" or "false" but also incomplete truth in where the truth rate may range between absolutely true and absolutely false [15]. It has four main parts i) Fuzzification which changes a actual value into a fuzzy set, ii)Knowledge base depends on IF-THEN rules, iii)Inference Engine depends on human analysis that create fuzzy inference between inputs and IF-THEN rules, iv) Defuzzification converts to real value as output from the fuzzy set which obtained from inference engine [15]. Fuzzy logic method is capable to apply in knowledgeable control system, nuclear fusion in order to increase efficiency and simplicity of the design procedure. Proposed management system has been developed based on FL technique using a Mamdani interface system.

The system consists of five inputs and four outputs shown in Fig. 8. The five inputs are PV (have four triangular membership function zero, low, medium, high shown in Figure 9), biogas generator, diesel generator, Load demand (LD) and State of Charge (SOC). Biogas generator and diesel generator have two triangular membership functions representing available and not available shown in Fig. 10 & 11. Last two input of load demand (LD) and state of charge (SOC) have three triangular membership functions of low, medium and high representing for 24 hours shown in Fig. 12 & 13. Four output membership function are PV-OFF-ON, BIOGAS-OFF-ON, GENERATOR–OFF-ON and Battery_NU_C_DC(B). PV-OFF-ON, BIOGAS-OFF-ON and GENERATOR–OFF-ON have two switches as OFF (S1) and ON (S2). Battery_NU_C_DC(B) has three switch as not use (S1) charge (S2) and discharge (S3

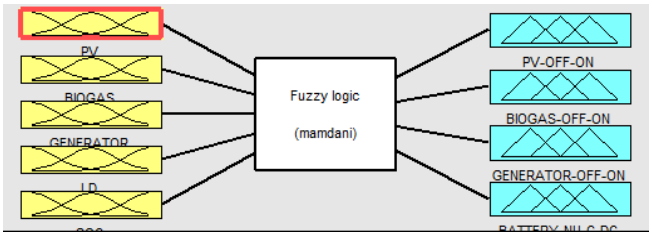


Fig. 8 Block diagram of proposed fuzzy logic controller

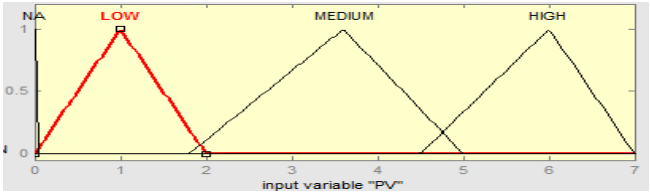


Fig. 9 Membership function of PV array

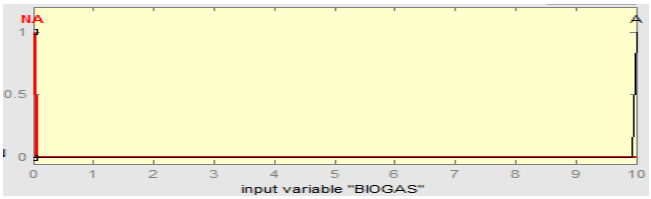


Fig.10 Membership function of Biogas Generator

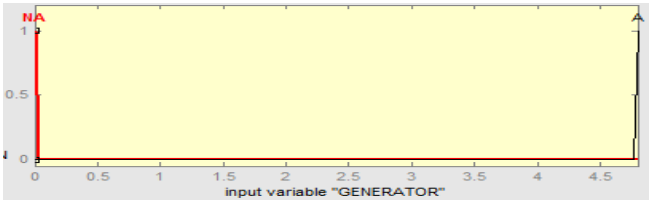


Fig.11 Membership function of Diesel Generator

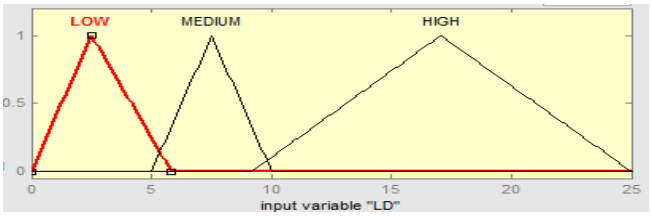


Fig.12 Membership function of Load demand

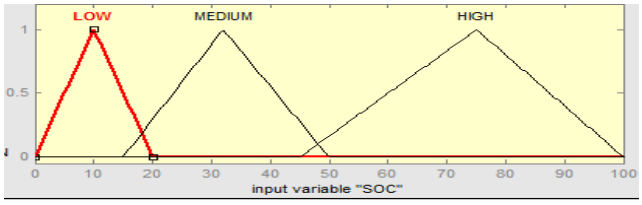


Fig.13 Membership function of SOC

FLC is related to inputs and outputs segment with a list of IF-THEN statements which is known as fuzzy rules. The IF known as input parts and THEN known as outputs part. In this research, we have established 24 rules to develop effective energy management system for different operational conditions which is shown in Table IX.

TABLE IX. CONSIDERING FUZZY RULES

		IF					THEN			
RULES NO.	TIME	PV	BIOGAS GENERATOR	DIESEL GENERATOR	LOAD DEMAND (LD)	STATE OF CHARGE (SOC)	PV-OFF-ON	BIOGAS-OFF-ON	GENERATOR-OFF-ON	BATTERY_NU_C_DC(B)
SUMMER & WINTER										
1-2	00-6	NA/NA	NA/NA	NA/NA	L/L	M/H	S1	S1	S1	S3
3-5	6-10	L/L/L	A/A/A	NA/NA/NA	M/M/M	L/M/H	S2/S2/S2	S2/S2/S2	S1/S1/S1	S2/S2/S1
6-8	10-13	H/H/H	A/A/A	NA/NA/NA	M/M/M	L/M/H	S2/S2/S2	S2/S2/S2	S1/S1/S1	S2/S2/S1
9-10	13-14	H/H	NA/NA	NA/NA	M/M	M/H	S2	S1	S1	S3
11-13	14-16	M/M/M	NA/NA/NA	A/A/A	M/M/M	L/M/H	S2/S2/S2	S1/S1/S1	S2/S2/S2	S1/S3/S3
14-15	16-17	L/L	NA/NA	A/A	M/M	M/H	S2	S1	S2	S3
16-17	17-21	NA/NA	NA/NA	A/A	H/H	M/H	S1	S1	S2	S3
18	21-22	NA	NA	NA	H	H	S1	S1	S1	S3
19-20	22-24	NA/NA	NA/NA	NA/NA	M/M	M/H	S1/S1	S1/S1	S1/S1	S3/S3
EXTRA RULE FOR SUMMER										
21-22	15-16	M/M	NA/NA	A/A	H/H	M/H	S2	S1	S2	S3
23-24	16-17	L/L	NA/NA	A/A	H/H	M/H	S2	S1	S2	S3

V. RESULT & ANALYSIS

On the basis of HOMER software requirement as input data of solar & biomass rescores, oil consumption and load demand, we have got some optimum & one sensitive result by simulation which is shown in Fig. 14.

	PV (kW)	Label (kW)	Label (kW)	S6CS25P	Conv. (kW)	Initial Capital	Operating Cost ($/yr)	Total NPC	COE ($/kWh)	Ren. Frac.	Diesel (L)	Biomass (t)	Label (hrs)	Label (hrs)	Batt. Lf. (yr)
	7	10	4	11	10	$ 30,300	11,500	$ 177,303	0.173	0.90	2,700	198	7,528	2,850	12.0
		10	6	7	9	$ 21,130	12,807	$ 184,841	0.180	0.86	3,554	218	8,083	2,961	12.0

Fig. 14 Sensitivity analysis by Homer

The annual electric energy production and consumption parameters from system which is shown in Table X. Average monthly generated electricity in the system which is shown in Fig. 15.

TABLE X. ELECTRIC ENERGY PRODUCTION & CONSUMPTION PARAMETERS.

Production	KWh/yr	Fraction in%	Consumption	KWh/yr
PV array	10547	12	AC load	80268
Generator 1	65199	77	Excess electricity	123
Generator 2	8692	10		
Total	84438	100	Total	80268

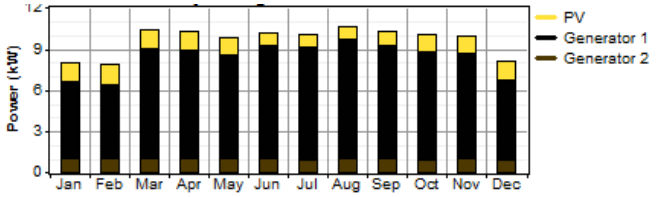


Fig.15 Monthly average electricity production in the system.

Fig. 16 indicates operation of PV array. Yearly electrical production is 10,547 kWh/yr. Maximum electrical output is 6.94 kW. Mean electrical output is 28.9 kWh/day. Capacity factor is 17.2%. Total operation hours is 4,375 hr/yr .

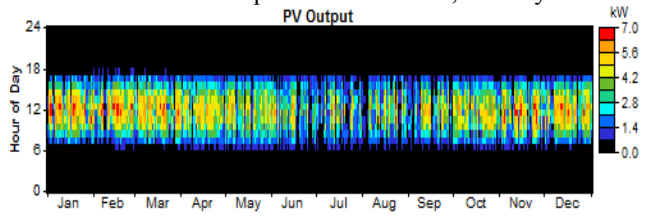


Fig.16 PV operation in the system.

Fig. 17 indicates operation of biogas generator. Yearly electrical production is 65,199 kWh/yr. Maximum and minimum output is 10 kW and 3 kW. Mean electrical output is 8.66 kW. Fuel consumption is 198 ton/yr. Specific fuel consumption is 2.124 kg/kWh and efficiency has been considered is 35% . Capacity factor is 74.4%. Total operation hours is 7,528 hr/yr .

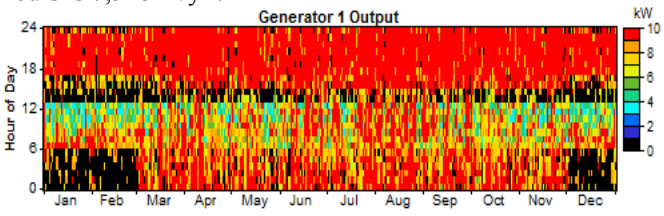


Fig.17 Biogas generator operation in the system

Fig. 18 indicates operation of diesel generator. Yearly electrical production is 8,692 kWh/yr. Maximum and minimum output is 4 kW and 1.20 kW. Mean electrical output is 3.05 kW. Fuel consumption is 2,700 L/yr. Specific fuel consumption is 0.311 L/kWh and efficiency has been considered is 32.7%. Capacity factor is 24.8% . Total operation hours is 2,850 hr/yr .

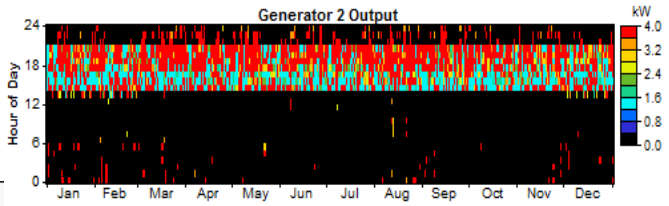


Fig.18 Diesel generator operation in the system

Fig. 19 indicates operation of battery. Energy in and out through battery is 8,795 kWh/yr and 7,044 kWh/yr . Total loss is 1,743 kWh/yr in rectifier.

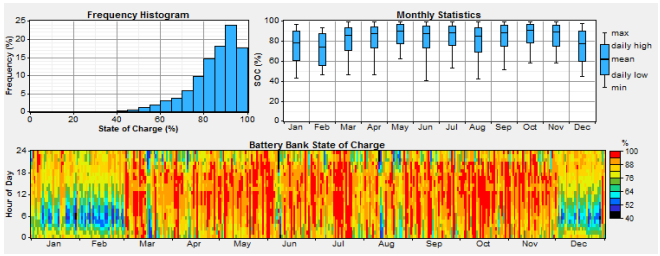
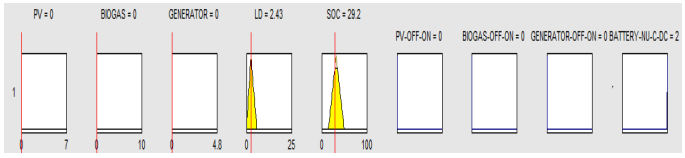


Fig.19 Battery status in the system.

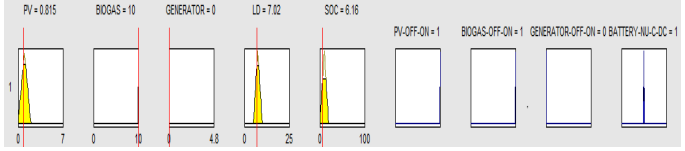
VI. MANAGEMENT STRATEGY EVALUATIONS

By optimizing input data of load demand, solar-biomass resources and system component, we have got 7 kW PV array, 10 kW biogas generator and 5 kW diesel generator. So the range of input membership function of PV array, biogas generator and diesel generator are 7kW, 10kW and 5kW. Fuzzy logic controller is designed with 24 rules for winter and summer season. We have generated only effective rules based on human knowledge that must minimize power loss. Optimum power is obtained through the tuning of fuzzy rules that depends on our thinking and analysis. At a time we have used different rules for meeting any condition within 24 hours, then shifted to another rules by depending on generated power and load demand.

CASE STUDY-1, FOR FUZZY RULE 1(TIME,00-6)
At night, when PV is not available, Biogas is not available, the fuel of diesel generator is not available, load demand low and state of charge is medium then PV-Biogas generator-diesel generator are OFF,Load demand is fullfiled by battery.

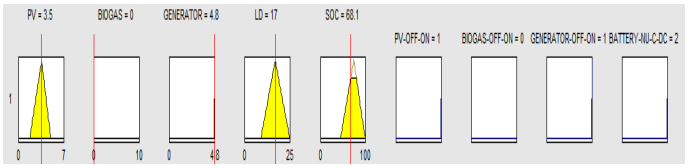


CASE STUDY-2, FOR FUZZY RULE 3(TIME,6-10)
At morning, when PV is low, Biogas is available, the fuel of Diesel generator is not available, load demand medium and state of charge is low then PV is ON, Biogas generator is ON, Diesel generator is OFF, Load demand is fullfiled by PV and Biogas. At a time extra power is used for bettery charge.



CASE STUDY-3, FOR FUZZY RULE 22(TIME,15-16)
At afternoon, when PV is Medium, Biogas is not available, the fuel of diesel generator is available, load demand is high and state of charge is high then PV is ON, Biogas generator is

OFF, Diesel generator is ON, Load demand is fullfiled by PV and diesel generator. But diesel generator and PV are not capable to fulfill the total load demand . So, bettery is discharge.



VII. COST SUMMARIES

Cash flow summary and net Present cost summary in terms of component are shown in Fig. 20 & 21 respectively. The system cost is defined as equation.

$$C_{SYSTEM} = C_{PV} + C_B + C_G + C_{BA} + C_{CON} \quad \text{.....(3)}$$

Where, Cpv indicate PV array cost, C_B indicate biogas generator cost, C_G indicate diesel generator cost, C_{BA} indicate battery cost and C_{CON} indicate converter cost [15]. Total net present cost is \$177,303, levelized cost of energy is \$ 0.173/kWh and operating cost is \$ 11,500/yr.

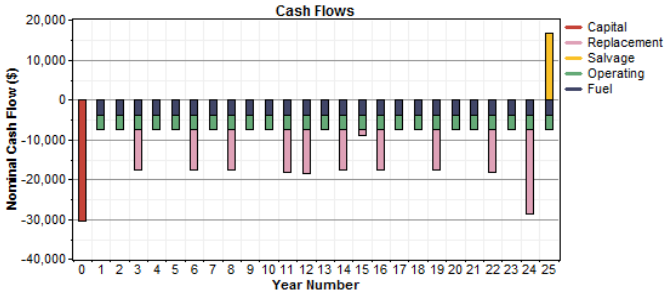


Fig. 20 Cash flow summary based on the selected components.

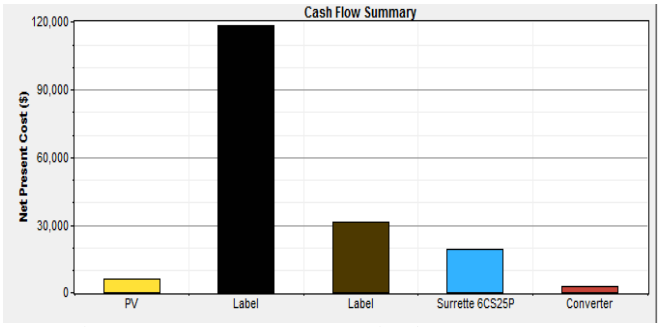


Fig. 21 Net present cost summary of each component.

VIII. CONCLUSIONS

The application of renewable energy is growing rapidly in all fields of the engineering to minimize CO_2 gas emissions in the system. This hybrid model is more effective than any other conventional systems that can easily applicable in everywhere in Bangladesh specially hill track areas, remote areas and Rohinga camps in where electricity is not available. By applying this proposed system, we can easily manage our domestic wastage and forest residue. HOMER Software is used to carry out optimal power based on load profiles and available renewable resources. Fuzzy Logic Controller (FLC) is used to manage the energy of the system properly. Component sizes calculation of PV arrays, biogas generator, diesel generator, battery and converter have been selected by using HOMER software and those are used as input of FLC for effective management system to reduce energy loss. Cost summary, cash flow summary and electrical production are the result of optimizing input data from HOMER. The solar energy, biogas energy and diesel generator contribute 12%, 77% and 10% respectively to generate power. Besides, the system is atmosphere approachable and favourable. The proposed system decreases CO_2 emissions and delivers clean energy that reduces our power crisis.

REFERENCES

[1] Bangladesh economic review 2017. Available:https://mof.portal.gov.bd/site/page/946aee4d-78bd-43b5-be2a-ed29342acba6, 2018.08.12

[2] Power Division Ministry of Power, Energy and Mineral Resources Government of the People's Republic of Bangladesh.2016

[3] A. S. N. Huda, S. Mekhilef , A. Ahsan "Biomass energy in Bangladesh: Current status and prospects" Renewable and Sustainable Energy Reviews 2014

[4] M. F. Ansari, S. Chatterji and A. Iqbal" A fuzzy logic control scheme for a solar photovoltaic system for a maximum power point tracker" International Journal of Sustainable Energy Vol. 29, No. 4, December 2010, pp.245–255.

[5] S. D. Saranya, S. Sathyamoorth and R. Gandhiraj" A fuzzy logic based energy management system for a microgrid" ARPN Journal of Engineering and Applied Sciences, Pp.2663, April 2015.

[6] Chittagong Hill track area. Available:https://en.wikipedia.org/wiki/Chittagong_Hill_Tracts , 2018. 9.29

[7] K. M. Kabir, M. A. M. Bhuiyan, Md. S. Uddin Chowdhury, R. Chowdhury, "Design and Analysis of Grid Connected Biomass System Utilizing Chittagong Municipal Food Waste "3rd International Conference on Electrical Information and Communication Technology (EICT), 7-9 December 2017, Khulna, Bangladesh.

[8] M. A. Islam, M. Maitra, A. B. Khan Majlis, S. Rahman" Spatial Changes of Land Use/Land Cover of Moheshkhali Island, Bangladesh: A Fact Finding Approach by Remote Sensing Analysis". Available:https://www.researchgate.net/publication/258837230, 2018.09.

[9] M. M. Sediqi, M. E. Lotfy, A. Yona and T. Senjyu" Optimal Economical Sizing of Grid-Connected Hybrid Renewable Energy System" Journal of Energy and Power Engineering, April 30, 2017.

[10] H. Tian, F. Mancilla-David, K. Ellis, and P. Jenkins, E. Muljadi" A Detailed Performance Model for PV Systems": Available: https://www.nrel.gov/docs/fy12osti/54601.pdf

[11] Salehin S. , Islam S., Hoque R., Rahman M., Hoque A. and Manna E. "Optimized model of a solar PV-biogas-diesel hybrid energy system for Adorsho Char Island, Bangladesh" in International Conference on Developments in Renewable Energy Technology (ICDRET), Dhaka, vol. 17, pp. 1-6, May 2014.

[12] Bala B. and Siddique S. "Optimal design of a PV-diesel hybrid system for electrification of an isolated island – Sandwip in Bangladesh using genetic algorithm", in Energy for Sustainable Development, vol. 13, pp. 137-142, November 2009.

[13] Diesel generator: Available:https://www.alibaba.com/product-detail/4-8kw-silent-dieselgeneratorportable_60501108654.html?spm=a2700.7735675.2017115.211.LUq1v6&s=p, 2018. 8. 17

[14] Biogas generator: Available:https://www.alibaba.com/product-detail/10kw-biogas-generator_60770277824.html?spm=a2700.7724857.normalList.2.65934a8efywXBp&s=p, 2018. 8.22

[15] J. M. Yang, K. W. E. Cheng, J. Wu, P. Dong, and B. Wang" The Study of the Energy Management System Based-On Fuzzy Control for Distributed Hybrid Wind-Solar Power System.

EYE ASSISTANT

Using mobile application to help the visually impaired.

Md. Amanat Khan Shishir, Shahariar Rashid Fahim, Fairuz Maesha Habib, Tanjila Farah
Department of Electrical and Computer Engineering, North South University, Bangladesh
{Amanat.khan, Shahariar.fahim, maesha.habib, tanjila.farah}@northsouth.edu

***Abstract*- Mobile devices or portable smart gadgets are largely used in our day-to-day life now. These devices can be very useful in helping people who are visually impaired and assist them to make their life easier. This paper presents an android application that recognizes objects using real time object and text detection by scanning them. The application can run on the device without any remote server. Our solution uses a voice feedback to tell the user about the detected object. The app doesn't need any photograph to detect the object. To enable robust recognition, we first segment the object from the background using tensorflow machine learning API based on iterative diagram cuts. We then formulate the recognition problem as an instance retrieval task and the user gets to know the object through a text to speech method. The system is designed to help a visually impaired person know about the environment around them. This app can be run on any low-end smartphone devices.**

Keywords- visually impaired; android application; scanning; TensorFlow; text to speech; image recognition; text recognition.

.

I. INTRODUCTION

Mobile phone has become a very significant device for out day to day life. Information of the whole world is in our hands due to this device. Nowadays we cannot imagine living without a mobile phone for even a day. From food to deliver service to calling taxi services, we are all dependent on mobile phones and it's applications. The use of mobile phone is growing every day. This is because it makes our life easier and makes various services easily accessible. According to GSMA data, which is an online mobile reviewer plateform, it is expected that the number of mobile phone user will pass around 5 billion [1]. In 2016, 56 percent of population in worldwide has owned mobile phone already and especially the use of smartphone is rapidly growing [2]. By 2014, it is estimated that around 70 percent of population were smartphone user [3]. By 2019, it is expected that the number of smartphone user will grow by 2.7 billion. The statistic can be seen in Figure 1.

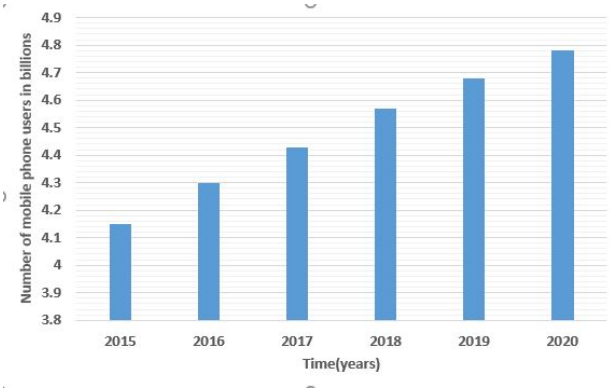


Figure 1. Number of mobile phone users between 2015 to 2020 (in billions)[4].

Visual impairment does not always mean complete loss of eyesight. According to the World Health Organization there were more than 285 million visually impaired living in the world in 2012 [5]. Even a slight visionary inability can affect a person's day to day life hugely. As represented in Figure 2 the number of visually impaired people have increased all over the world.

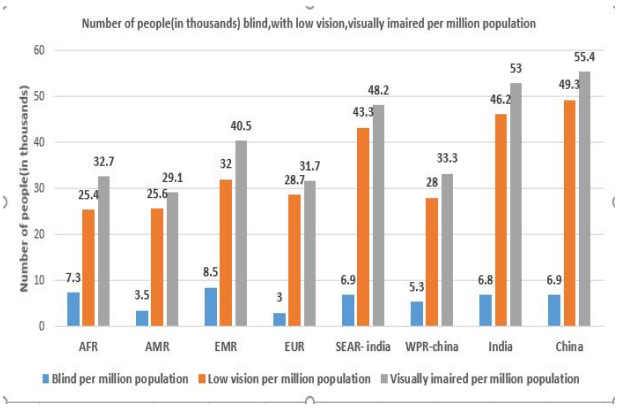


Figure 2. Number of people (in thousands) blind, with low vision & visually impaired per million population [6].

The inability to see the surrounding makes it difficult to move around or do anything on one's own. A lot of objects have the same texture or shape which makes it even harder for a visually impaired person to sense what objects are around or they holding. Nearly one-third of people with vision loss suffer from clinical depression, a rate that is twice as high as the general population of the

same age [7]. Visually impaired people suffer in many ways. The proposed system is an android application that detects objects by scanning them and replying with a voice feedback that can assist any visually impaired person. As usage of mobiles have increased and price has reduced in the past few years drastically, this vastly used and easily accessible device seemed better choice for this system.

II. SMARTPHONE APPLICATIONS

There are quite a few applications out there in the market that is similar to ours. Some of them are costly to use. This application is open to all. Existing applications are mostly used to detect currency bills of different countries. A few of them are there that detects objects around them. But nearly all of them requires taking a picture of the object or the bill and save it in memory before it can detect what object it is. What makes our application different from all the others is that our application does not need to take a picture of the particular object, but instead it scans the object and recognizes the object using real time object detection. It is easy to use for blind people. That way the visually impaired does not have to worry about taking a proper picture nor does he have to worry about.

A. How does it work

We have used two features for the application. A user can recognize object and can hear any text over voice. So, for these two features we have used an open source machine-learning framework, which helps to detect any object and google text to speech API to recognize text.

I. Object Detection

In this case, when a user wants to detect what object is it or what is the name of a particular object the user needs to open camera. For this, we used google camera API that can detect camera hardware information and eventually open the camera. For displaying any object texture View class is used to take frame. Later by using TensorFlow machine learning framework we have trained our dataset and then imported the trained dataset. For detecting object our algorithm take input by using camera and feed the input to the algorithm then communicate with trained model and eventually make the prediction. Figure 3 represents how algorithm actually works to detect object.

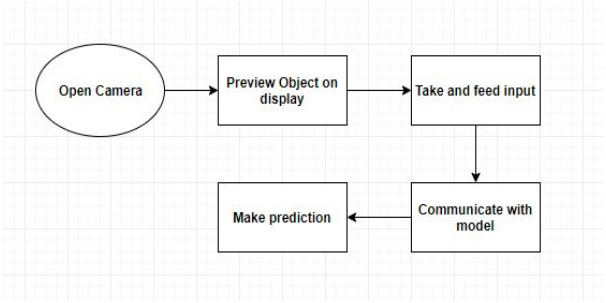


Figure 3. Flow chart showing object detection algorithm works.

II. Recognize Text

Optical Character Recognition (OCR) gives a machine power to read any text. OCR is the best way to recognize any text which make visually impaired user able to hear text out loud. Camera takes individual frame using detect methods and TextRecognizer. To read any text straight from camera, when any texts become available we have implemented processor in OcrDetectorProcessor which handle all the detections. From TextRecognizer ReceiveDetections will receive text blocks when text becomes available. For this, Text Recognizer is used to detect image and process the image that determines what text appears within the image and finally make an audio synthetic voice using text to speech API to give output. Figure 4 illustrates how the text to speech API works in that application.

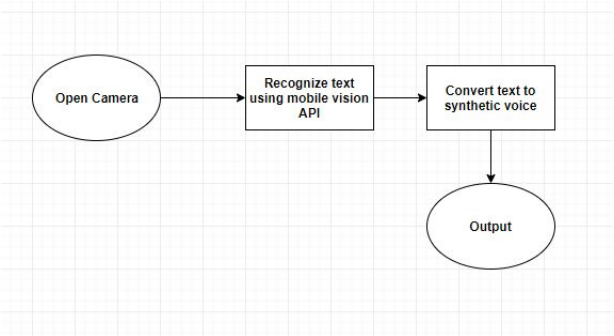


Figure 4. Flow chart of text to speech algorithm.

B. How to use

Visually impaired have difficulty being aware of their surrounding and recognize objects. In our app, at first they need to open the app and it will start working without any other command. The mobile camera will scan the object and the real time data will be processed evaluated with the system dataset and finally it will give the result. A user can just open the app using Google assistant just using voice. Later the user is asked whether the user wants to recognize object or recognize text. Figure 5 illustrate how to use the application.

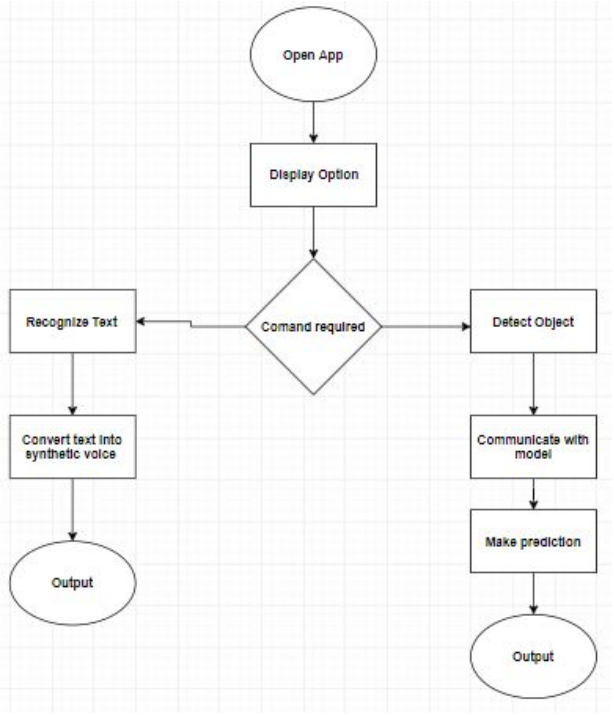


Figure 5. Flow chart showing the work of application.

III. TESTING THE APPLICATION

The whole system testing is done using a Huawei phone. In this system a user can open the application by using Google assistant and user can choose if user wants to recognize object or read any text. The system is very easy to use and user friendly. Figure 6 shows the interface. If the user chooses the "Detect Text" option then the camera opens and if you hold it against any text the app can read it out to the user. If the user chooses the "Detect Object" option then again the camera opens and if it is held against any object then the app tells the user what object it is.

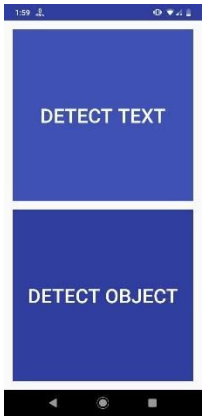


Figure 6. A user can detect object and detect text

For detection we have trained some data that are available around us. We have used tensorflow machine learning API, which is developed by Google brain team. So a user who has a poor vision can recognize any object. After testing the application we saw that the recognition accuracy for every object is more than 80 percent. So we can fairly say that if we train more data and develop the application then it will be able to detect as much objects as possible and with better accuracy rate. Figure 7 shows how a user recognize object.



Figure 7. Object recognition test.

We tested our app for recognizing text and we saw that our app can recognize all the text that we tested. As we have trained our app in English so our app can detect all texts written in English. So if the user holds the camera against a text the app gives out a clear output of the text and reads it out to the user. As our target users are visually impaired or blind people, this text recognition can help them in many ways to identify anything written around them or on any object. Figure 8 shows how text to speech works in real time.



Figure 8. Real time text recognition

IV. FUTURE WORKS

The app we have built so far is suitable to be used by people with poor vision or visually impaired. In future, we intend to add voice command so that fully blind people will be able to use the app easily. Currently the data in our app is limited data. We have used CIFAR 10 data set where there are 600 images of 10 objects, which we have trained. Our plan for the future is to work with many more data sets and to train them so that a blind or visually impaired user can use the app to recognize objects of their day-to-day life easily.

IV. CONCLUSION

Our client targeted is people who are visually impaired and we have noticed that their life is quite difficult. We created the app with features, which can help them make their life easier by knowing what's around them by using our app. We have seen that all the applications similar to ours have been made before but those app need to capture image and then return what object it is. But our application does not need to take an image and it only scan object and then tells the user what object it is. That way the visually impaired user doesn't have to delete images from memory every now and then.

V. REFERENCES

[1] 5 billion people now have a mobile phone connection,"Online .[Available] https://venturebeat.com/2017/06/13/5-billion-people-now-have-a-mobile-phone-connection-according-to-gsma-data/ [Accessed 30th April 2019].

[2] World wide internet and mobile users for 2016 Online .[Available]https://www.emarketer.com/Report/Worldwide-Internet-Mobile-Users-eMarketers-Estimates-2016/2001798 [Accessed 30th April 2019].

[3]The number of smartphones users in the world is expected to reach a giant 6.1 billion by 2020," Online .[Available] https://www.emarketer.com/Article/Smartphone-Users-Worldwide-Will-Total-175-Billion-2014/1010536[Accessed 30th April 2019].

[4] Statista, "Number of mobile phone users worldwide from 2015 to 2020," Online. [Available] https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/ [Accessed 30th April 2019].

[5] World Health Organization, http://www.who.int/mediacentre/factsheets/fs282/en/, [Accessed 19 May 2013].

[6] World wide blind statistics," Online. [Available] https://www.sitepoint.com/how-many-users-need-accessible-websites/ [Accessed 30th April 2019].

[7] J.E. Javitt, Z. Zhou, R.J. Willke, "Association between vision loss and higher medical care costs in Medicare beneficiaries costs are greater for those with progressive vision loss," Ophthalmology, 114(2), 238-245, 2007.

An Approach for Improving Complexity of Longest Common Subsequence Problems using Queue and Divide-and-Conquer Method

Ratul Bhowmick, Md Ibrahim Sadek Bhuiyan, Md. Sabir Hossain, Muhammad Kamal Hossen
Department of Computer Science and Engineering
Chittagong University of Engineering and Technology (CUET)
Chittagong-4349, Bangladesh
Email: u1604038@student.cuet.ac.bd, u1604036@student.cuet.ac.bd, sabir.cse@cuet.ac.bd, kamalcsecuet@gmail.com

Ahsan Sadee Tanim
Department of Computer Science and Engineering
The International University of Scholars
Dhaka-1212, Bangladesh
Email: ahsansadeecuet@gmail.com

***Abstract*—The general algorithms which are followed to solve the Longest Common Subsequence (LCS) problems have both time complexity and space complexity of $O(m * n)$. To reduce this complexity, two ways are proposed in this work. One is the use of Divide and Conquer approach, and another one is changing the data structure from two-dimensional array to queue. By using these approaches, an algorithm has been developed and implemented for the long length string in this paper. The time complexity after using these approaches becomes a function of logarithm $O(n * \log(n))$ which has been shown with proper explanation and the space complexity after using queue data structure becomes linear $O(n)$.**

Keywords—Longest Common Subsequence, Queue, Divide and conquer, Dynamic Programming, Recursion

I. INTRODUCTION

Longest Common Subsequence (LCS) is a popular algorithm to find the longest subsequence which is a common group of sequences. Normally, two sequences are given and from that, LCS has been found out. A subsequence is gained from another sequence by removing some elements but can't change the series of the other elements. There are many other algorithms to find common subsequence. Most of them follow the greedy approach. But LCS follows the dynamic approach. It can be decided in polynomial time.

The general LCS algorithm takes O(m*n) time. In this paper, there has been proposed an algorithm which finds the solution by following an optimal infrastructure. The problem which is given is broken into subproblems which is smaller and simple, which will be broken into yet further smaller subproblems until we find the final answer. For this approach, time and space complexity is optimal. The time complexity becomes $O(n * logn)$ for the best case and $O(n * \log(n)^2)$ for the average case. But the space complexity becomes linear by using data structure queue instead of two-dimensional array and it is $O(n)$in both best and worst cases. Although this algorithm is not unique, it can become really complex for a very bigger length. But for most of the cases of finding common subsequence, it finds the solution in optimum time.

Many researchers have researched about LCS over a period of time. They have given many results. But still it is very complex and there is a huge area left to research on this topic. The time and space needed to solve a problem by using LCS are still high. The main objectives of this research are:

1. To reduce the time complexity and find an optimum a bigger length string.
2. To reduce the space complexity by changing the data structure.
3. To use a new approach (Divide and Conquer) and to see the effects of using this approach.

The remainder of this paper is organized as follows. Section II provides a brief review of related work. In section III, we discuss in detail the proposed algorithm. Section IV contains experimental results and analysis. Finally, we conclude and sketch future research directions in Section V.

II. RELATED RESEARCH

J. Liu and S. Wu show that it is possible to find out the Longest Common Subsequence in linear time using a 2D array [1]. The result lies between in the diagonal line of the two-dimensional array. They used for them theorems provided in [2], [3]. Paper [3] has also shown that the space complexity can be linear in the longest common subsequence. This paper merges two algorithms to reduce the space complexity.

K. Kwarciak and P. Formanowicz propose a DNA sequencing technique using a greedy approach. In their algorithm, the greedy method is applied to two different sequences at the same time. That's why this greedy algorithm is much faster than the dynamic approach. This approach can be used as the longest common subsequence. J. F. Myoupo and D. Semé used a parallel algorithm to minimize time complexity [5]. This algorithm runs in O(logN*logM) which is much smaller than the existing LCS algorithm. In [6], this problem is solved in O(log m*log n) with O(n*m/log m) processors. The CRCW bound is O(log(n(log*log m)2)) time where the processors are O(m*n/log* log m).

In paper [7], X. Tang et al. used a block system by which the process become faster than before. The idea of a B+ tree and O tree is used in this algorithm which has less run time than other algorithms. For that reason, they get a result which takes only 36.5 mm^s. Paper [8] proposes a data-structure that helps to improve the competence of suffix array in the algorithm by storing in external memory model. The authors of this research paper also propose and implement two parallel algorithms (Index Partitioning and

Data Partitioning) on a PC cluster. Here the ASM problem is solved in DNA sequences with the help of suffix array algorithms.

R. L. Goldfeder et al. proposes a system where the DNA sequence/String is divided into some small pieces, and it will be reordered programmatically such that it can enable fast and efficient sequencing for human genomes or for a string [9]. In this paper, they divided the DNA or String sequence in a small portion and get all data and then merged.

Paper [10] presents real-time sequencing and single molecule data that is obtained from a DNA polymerase. DNA polymerase is performed in a continuous template directed fusion using four different fluorescently categorized dNTPs. The data is reported directly used to reveal the distinct polymerization states corresponding to DNA structure. The paper shows 90% accuracy in medical science. F. Chin and C.K. Poon demanded that there has L-matrix on the dynamic approach in LCS [11]. They showed the time complexity becomes O(ns+min(ds, lm)) and the space complexity becomes O(ns+d). They used different data structures to obtain variations of the basic algorithm.

III. The Proposed Algorithm

In the proposed algorithm, the existing Longest Common Subsequence algorithm has just modified by using the queue data structure and the divide-and-conquer method. The existing general LCS algorithm works on $O(MN)$ time. It can be made O(1) if the size of those strings is minimized. Here this idea is used. Let, two string A(A1, A2,….., An) and B(B1, B2,….Bn) are given and C(C1, C2,….., C k) is the sequence of two string. If A >= B, then this algorithm will divide the A string into small parts until the size of the string becomes less than 4. Then these small string and the B string will use in the second process. From the algorithm in Fig. 2the largest string is divided into several small strings (as shown in Fig.3). The small string and the second string are applied in the second algorithm which is given in Fig. 3. The whole algorithm runs in O(n*log(n)) time in the best case. This algorithm takes linear space because of using the queue instead of a two-dimensional array.

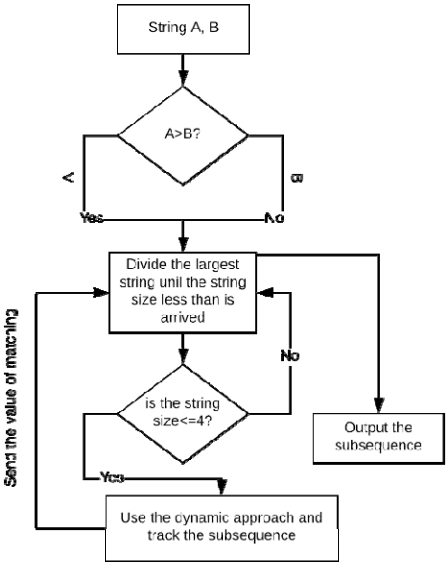


Fig.1 The proposed methodology

From Fig.1, it is clear that there have two steps. The first step is the divide-and-conquer method and the second step is DP.

A. Algorithm

In this algorithm, we have two phases. In the first phase, the large string will be made into a small string by the divide and conquer method. In the algorithm in Fig. 2, the algorithm of the first phase is given.

```
Function making_a_short_string( start , end)
1.  If(start – end <= 4)
2.      return LCS( A[ start … end] , B [ 0 …m])
3.  end
4.  mid = (start+end)/2
5.      return making_a_short_string(start, mid) +
    making_a_short_string(mid+1, end);
End
```

Fig.2 Pseudo code describing the divide-and-conquer method to short the large string

From string A, when the base case arrives in the first phase, it will execute the second algorithm which is given in Fig. 3. From this algorithm, we can get the sequence.

```
Function LCS(s1[ 0…n] , s2[0 …m])
                Bool Array[0…….m)
1.  Declare a queue data structure Q;
2.  matching_number <- 0
3.  for i:= 0 to n
4.    for j:= temporary to m
5.      if i == 0 or j == 0
6.        if i == 0
7.    Q<- 0
8.        End
9.        if j == 0
10.           Q <- 0
11.       End
12.     End
13.     Else
14.       if s1[i-1] == s2[j-1]
15.          diag <- Q.front and pop the front value
16.          previous_value <- diag+1
17.          Q <- previous_value
18.             If prevous_value> matching_number
19.                temporary <- j
20.                Array[j]<- true
21.                Matching_number = temporary
22.                If j == n-1 return Matching_number
23.             End
24.       End
25.       Else
26.     diag <- Q.front
27.          Pop the front value of Q
28.          Vertical_value <- Q.front
29.          Temp = max( previous_value,
    Vertical_value)
30.          Previous_value <- temp
31.          Q <- previous_value
32.       End
```

```
33.        If j == m
34.          Pop the front value of the queue
35.        end
36.        end
37.     end
38.    end
39.    Return matching_number
40.  End
```

Fig. 3 Pseudo Code describing the algorithm of LCS

From this algorithm, we can get an array "Array" which track the sequence. If we iterate through the "Array", we will able to get the longest common subsequence C [0….m] which is "Array".

B. Example

Let Two string A= "abcadbcadadcaaddaddcdcddcdadccdaa .....aaddaada" and B = "abdcad" and C is the sequence where C is equal the size of the min (|A|, |B|). Here B is small so the size of |C| = |B|Here |A|>|B|, so we will divide A into many small strings by algorithm which is shown in the Fig. 4. From the best case of the string, the Strings will go to the next algorithm, where the small string and the new string will go through the algorithm in Fig. 3.

Let the best case string from the algorithm is "abca" and the B string is "abdcad". Then the process will get an LCS between those two strings where the two strings are small so it will run in constant time. Then the subsequence can get by Fig. 3, when i is zero means at the first row of the algorithm, then the queue will be 0 0 0 0 0 0, then the second loop the first element will be zero then the queue will be 0 0 0 0 0 0 0, then if the i-th element and the j-th element is similar then the front value of queue will be incremented and it will be popped out and pushed in the queue, then the queue will be 0 0 0 0 0 1. Then if the i-th and j-th are not similar then it will get from the previous value and the vertical value.

To get the vertical value firstly popped out the first value of the queue then take the new front value. Then compare which is maximum and push it to the queue. Then the queue will be, 0 0 0 0 1 1 by continuing this process after the second-row finishes then the queue will look like 0 1 1 1 1. Here the queue is using, again and again, the first value is popped and the new value is inserted. Lastly, the queue will be 0 1 2 3 4. Then it will return the value 4. During this method, a temporary "Array" is declared which will track the elements which are the subsequence of both strings.

Strings will go to the next algorithm, where the small string and the new string will go through the algorithm in Fig. 3. Let the best case string from the algorithm is "abca" and the B string is "abdcad". Then the process will get an LCS between those two strings where the two strings are small so it will run in constant time. When i is zero means at the first row of the algorithm then the queue will be 0 0 0 0 0 0, then the second loop the first element will be zero then the queue will be 0 0 0 0 0 0 0, then if the i-th element and the j-th element is similar then the front value of queue will be incremented and it will be popped out and pushed in the queue, then the queue will be 0 0 0 0 0 1. Then if the i-th and j-th are not similar then it will get from the previous value and the vertical value. To get the vertical value firstly popped out the first value of the queue then take the new front value. Then compare which is maximum and push it to the queue. Then the queue will be 0 0 0 0 1 1. By continuing this process, after the second-row finish then the queue will look like 0 1 1 1 1.

Here the queue is using again and again, the first value is popped and the new value is inserted. Lastly, the queue will be 0 1 2 3 4. Then it will return the value 4. During this method, a temporary "Array" is declared which will track the elements which are the subsequence of both strings.

IV. Experimental Result & Discussion

The time complexity of this proposed method is n*log(n) and the space complexity is O(n). The process is given in the next sections.

A. Time Complexity

Let, two strings are A[0,….k] and B [0,…..m] and n is the size of min(|A|, |B|). The algorithm has two parts. One is for the divide and conquers method which is recursive and one is a dynamic approach method. Let the time complexity of the dynamic approach is T1(c) and the time complexity of the total algorithm is T(n).

Then, T(n) = T1(c) + 2T(n/2)
= T1(c) + 2{ 2T(n/4) + T1(n/2) }
= 2T1(c) + 4T(n/4)
= 2T1(c) + 4{ 2T(n/8) + T1(n/4) }
= 3T1(c) + 8T(n/8)

= kT1(c) + 2^k T(n/2^k) -------- (i)

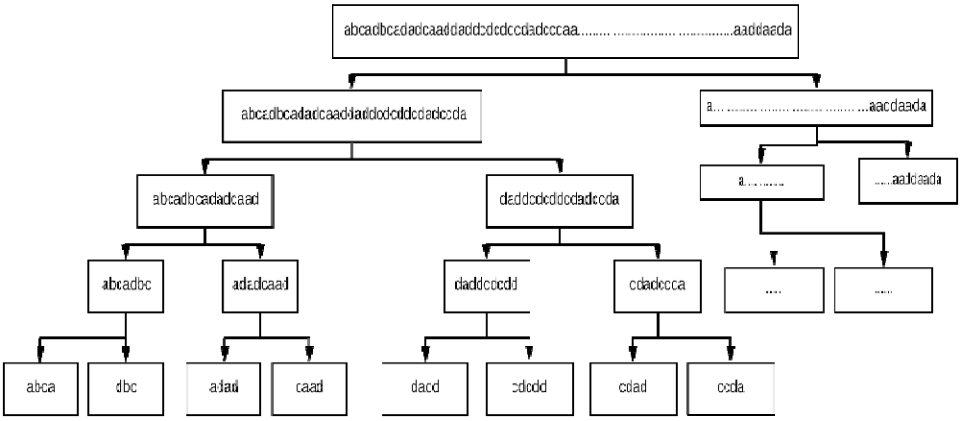


Fig. 4 Divide-and-Conquer method of the largest string

Let, n/2^k = 1

⇨ n = 2^k

⇨ log(n) = k [log base 2]

Substituting k in (i),

T(n) = log(n)T1(c) + n*log(n) -------- (ii)

So, the general time complexity is log(n)T1(c) + n*log(n).The best, worst and the average case is depends on the T1(n) term.

B. Best Case Analysis

From section VI A, the general time complexity is log(n)T1(c) + n*log(n). Let the new sting D and the second string is B. If the new string is the proper subset of the second string (D ⊂ B), then the T1(c) will the constant time. Because all the elements of D are within in B, and the size of the D is less than 4, then the second algorithm will run on constant time. So, T1(c) will be 1.

From equation (ii),T(n) = log(n)*1 + n*log(n)

So, the best case of this algorithm is n*log(n) where n is the size of min (|A|, |B|).

C. Average Case Analysis

The Average case of this algorithm is similar to the best case of the algorithm. Let the new sting D and the second string is B. In the average case, D is the subset of B, means that D has some elements which are in the B. So every recursive method the size of B is reducing, for that reason the algorithm has to go through the size of the B string. The second algorithm will run on linear time. So, T1(c) will be m.

From equation (ii),T(n) = log(n)*m + n*log(n)

So, the best case of this algorithm is k*log(n) where n is the size of min(|A|, |B|) and k is the size of max(|A|, |B|)

D. Worse Case Analysis

The general time complexity is log(n)T1(n) + n*log(n). Let the new sting D and the second string is B. If D' is the proper subset of B means that there has no element in D which is in B. So, it will iterate through all the elements in D so the algorithm will run on O(k*m) where the size of the D is k and the B is m. So, the worst case time complexity will be T(n) = log(n)* k*m + n*log(n) where k is log(n)/2.

Then, T(n) = m*log(n)^2 + n*log(n). So, the worst case of this algorithm is m*log(n)^2, where n is the size of min(|A|, |B|) and m, is the size of max(|B|,|A|).

E. Space Complexity

The space complexity of the first algorithm is O(n/2+1)~ O(n/2) and the space complexity of the second algorithm is O(n) because instead of a 2D array, there is used a data structure queue. The space complexity of queue for the best and worst case is O(n). So the space complexity of the full algorithm is O(n) means that the program is running on a linear space algorithm.

A comparison between our proposed method and the proposed framework of other authors is given in Table 1. From Table 1, it is clearly seen that our algorithm's space complexity is better than any other of these mentioned algorithms and the time complexity becomes a logarithmic function which is optimal than some of the other algorithms.

TABLE I Comparison of time and space complexity

Algorithm	Best time complexity	Average time complexity	Worst time complexity	Space complexity
LCS	Ω(m*n)	Θ(m*n)	O(m*n)	O(m*n)
Ref. [1]	Ω(n*n)	Θ(n*n)	O(n*n)	O(m+n)
Ref. [2]	Ω(n(m-p))	Θ(n(m-p))	O(n*m)	O(m+n)
Ref. [3]	Ω(m*n)	Θ(m*n)	O(m*n)	O(n)
Ref. [11]	Ω(ns+min(ds,lm))	Θ(ns+min(ds,lm))	O(ns+min(ds,lm))	O(n*m)
Our Algorithm	Ω(n*log(n)), n=max(|A|, |B|)	Θ(k*log(n), k=max(|A|, |B|)	m*log(n)^2 + n*log(n), m=max(|A|, |B|), n=max(|A|, |B|)	O(n)

TABLE II Comparison of time with General LCS algorithm

Length of two string (n and m)	LCS(sec)	Our Algorithm (sec)
n<100 && m<100	0.932	0.812
n<10000 && m<10000	3.003	2.855

In this paper, the space complexity of LCS is highlighted. The space complexity of this algorithm is linear. The best space complexity of queue is O(n) as well as the space complexity is O(n). Divide-conquer approach has been used here. For this approach, this will take constant time to solve a problem. Later the data structure has been changed from a two-dimensional array to queue. For this, space will be reduced. In Table II it has shown time comparison between general LCS and our algorithm.

Generally, for LCS, the worst time complexity becomes exponential and it becomes very complex to find a subsequence from a very large string. But in the proposed case of this paper, the worst time complexity is O(m*((log n)^2)) for which it becomes very optimal to get the final result from a very large string. But not for only the worst case, In the time complexity for best case in the proposed LCS is O(n*log(n)). It is very optimum and time-consuming to get the final result.

V. CONCLUSION

The LCS algorithm is studied and researched by a lot of researchers. Here, we have developed an algorithm by using a divide and conquer approach and a queue. By using this approach we have made the space complexity linear O(n). The time complexity becomes a function of the logarithm. For the best case, it becomes O(n*log(n)) where n is the length of the smallest string between the two string. For average time complexity, it becomes O(k*log(n)) where k is the length of the largest string between the two string. For the worst case, the time complexity is O(m*log^2(n)) where m is the length of the bigger string. Normally the time complexity for the general LCS algorithm over the world is

O (n^2). We are able to show that our algorithm is better which gives results faster and use linear space than the used algorithm.

REFERENCES

[1] J. Liu and S. Wu, "Research on longest common subsequence fast algorithm," in *2011 International Conference on Consumer Electronics, Communications, and Networks, CECNet 2011 - Proceedings*, 2011.

[2] N. Nakatsu, Y. Kambayashi, and S. Yajima, "A longest common subsequence algorithm suitable for similar text strings," *Acta Inform.*, 1982.

[3] D. S. Hirschberg, "Algorithms for the Longest Common Subsequence Problem," *J. ACM*, 1977.

[4] K. Kwarciak and P. Formanowicz, "A greedy algorithm for the DNA sequencing by hybridization with positive and negative errors and information about repetitions," *Bull. POLISH Acad. Sci. Tech. Sci.*, vol. 59, no. 1, 2011.

[5] J. F. Myoupo and D. Semé, "Time-Efficient Parallel Algorithms for the Longest Common Subsequence and Related Problems," *J. Parallel Distrib. Comput.*, 1999.

[6] A. Apostolico, M. Attalah, L. Larmore, and S. Mcfaddin, Efficient parallel algorithms for string editing and related problems, SIAM J. Comput. 19 (1990), 968?988.

[7] X. Tang, R. Tian, and D. F. Wong, "Fast evaluation of sequencepair in block placement by longest common subsequence computation," *IEEE Trans. Comput. Des.Integr. Circuits Syst.*, 2001.

[8] L. L. Cheng, D. W. Cheung, and S. M. Yiu, "Approximate string matching in DNA sequences," in *Proceedings - 8th International Conference on Database Systems for Advanced Applications, DASFAA 2003*, 2003.

[9] R. L. Goldfeder, D. P. Wall, M. J. Khoury, J. P. A. Ioannidis, and E. A. Ashley, "Human Genome Sequencing at the Population Scale: A Primer on High-Throughput DNA Sequencing and Analysis," *Am. J. Epidemiol.*, 2017.

[10] J. Eid *et al.*, "Real-time DNA sequencing from single polymerase molecules," *Science (80-.).*, 2009.

[11] F. Chin, and C.K. Poon, Chung, "Fast algorithm for computing longest common subsequences of small alphabet size," *Journal of Information Processing. 13. 463-469, 1991.*

Proficiency Assessment of Adaptive Neuro-Fuzzy Inference System to Predict Wind Power: A Case Study of Malaysia

Md. Rasel Sarkar
Dept. of Mechanical Engineering
University of Malaya,
Kuala Lumpur, Malaysia
raselbdeee@gmail.com

Mst. Jesmin Nahar
Dept. of Mathematics
Begum Rokeya University
Rangpur, Bangladesh
mjnahar901@brur.ac.bd

Awatif Nadia
Dept. of EEE
Ahsanullah University of Science & Technology
Dhaka, Bangladesh
nadiahossain6@gmail.com

Md. Abdul Halim
Dept. of EEE
World University of Bangladesh
Dhaka, Bangladesh
halimabdul552@gmail.com

S. M. Sajjad Hossain Rafin
Dept. of EEE
World University of Bangladesh
Dhaka, Bangladesh
rafin.sobhan@gmail.com

Md. Mamunur Rahman
Dept. of EEE
World University of Bangladesh
Dhaka, Bangladesh
mamunur.mail@gmail.com

***Abstract*— Wind power prediction is necessary for proficiency assessment of a potential wind site and can improve the technical and economic integration of wind energy into the existing electricity grid. The aim of the study is to compare the analysis of the wind power prediction between Kuala Lumpur and Melaka in Malaysia for potential place selection of wind power generation. Field data are collected from two places in Malaysia over a period of three years each place from December 2013 to December 2015. The model has been developed by an Adaptive Neuro-Fuzzy Interface System (ANFIS). In addition, the effectiveness of soft computing ANFIS is applied to the wind power prediction model. It involves four input variable-temperature, humidity, pressure and wind speed and output variable wind power has done using MATLAB. The predicted wind power by these sites are compared with the measured data of Meteorological Department of Malaysia. The proposed predicted model of ANFIS is clearly indicated high capability and reliability based on error values of estimated wind power. Therefore, ANFIS model has been shown excellent results in the analysis the root means square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and coefficient of determination (R^2).**

Keywords— Adaptive Neuro-Fuzzy Interface System (ANFIS), Wind power and prediction

I. INTRODUCTION

For the past few decades, the demands of energy have been gradually increasing, especially for electrical power and environmental issues and this has become a challenging issue for the world. Furthermore, pollution is growing parallel with the energy demand while sources of conventional energy such as fossil fuels are rapidly depleting. Over the last few decades, researchers have been conducting studies to improve the energy efficiency [1] [2]. Wind power has been used for nearly three thousand years. At the early twentieth century, the power of wind was used to deliver mechanical power to pump water and electricity [3]. Wind turbine is a device that produces the potential winds and converted to electrical energy [4]. The wind power plant is established in desolate areas whereas weather condition is harsh [5]. Basically, there are various types of wind turbines exist in this world but horizontal axis wind turbine (HAWT) and vertical axis wind turbine (VAWT) are the most common and widely implemented wind turbines

Economic development of a county depends on the energy. Fossil fuels cannot be contributed to fill up world energy demand. Some of developing countries are still under stress for energy shortage for fossil fuel has limitation for usage [6]. In addition, it has more negate environment impact to compare renewable energy. Alternatively, renewable energies are playing a role model to achieve energy demand in the World which has less environment impact.

Wind energy is a renewable energy which is produced by wind blowing by captured wind turbine. Wind turbine converted to mechanical energy into electrical energy. Some of countries namely Denmark, Germany, Spain, United Sates and India are producing electrical from wind energy [7]. Site selection for wind turbine installation is essential to produce maximum wind energy. In addition, the accurate wind power and wind speed prediction would be profitable for operation. There are several mathematical models suggested for wind power prediction. They have some difficulties namely higher calculation time required. On the other hand, soft computing techniques are provided better advantages including no required knowledge of internal systems parameters, compact solution for multi-variable problems and fast calculation. Recently, some researchers have successfully worked in the wind energy.

The crisis of energy is the changing for the researcher. For the overcome this crisis, researchers are concentrated to the renewable energy such as wind, solar, rain, tides, waves, and geothermal heat energy for power generation as alternate natural sources. Wind power can be used for power generation where availability of wind speed is existing. There are various European countries to start power generation using wind speed. The nature of wind speed is fluctuating so the small change of wind speed can be affected to wind power generation. The site selection is very crucial issue for wind power generations. Before wind turbine installation in wind farm, the wind speed prediction is very important criteria for site selection [8, 9].

Wind power prediction can be classified input variables and time series. In this study, input variables are considered for prediction. During operation condition of wind turbine can be controlled with millisecond to second by wind speed prediction [10, 11]. Mainly there are two types of methods those are used for prediction purpose namely physical and

statistical. The outcome of the statistical method is to build up a relationship with wind power forecast and measured data of historical time series which can be helped for site selection of wind farm installation [12]. On the other hand, physical method tries to find wind power output by considering some variable namely environment temperature, local terrain and numerical weather [13, 14].

There are several statistical methods such as Autoregressive Moving Average (ARMA), persistence model and Kalman filter [15]. In Addition, for historical time series employed by the artificial intelligences namely radial basis function neural network (RBFNN), backpropagation neural network, support vector machine (SVM), nonlinear autoregressive neural networks (NARNN) and adaptive neuro-fuzzy inference (ANFIS) [16] [17]. In this study, ANFIS employed for wind power prediction two different area namely Kuala Lampur and Melaka in Malaysia. The proposed ANFIS is tested on a case study real data from metrology department in Malaysia.

The layout of the paper is as follows. Power of wind turbine is presented in section II. ANFIS is described in section III. Results and discussion are displayed in section IV. Finally, the conclusion of this paper is presented in section V.

II. Power Of Wind Turbine

The wind power has achieved advanced efficiency due to the modern technology. Since the 1980, the structural dynamics, aerodynamics and micrometeorology have contributed to an increase about 5% yearly [18]. The heavy weight and noise of wind turbines are accountable for the low power output. The higher power generation of the wind turbine not only depends on an electric generator and high-capacity machine but also on faster computer base techniques [19]. For HAWTs, most of the researchers make an analysis on the HAWT performance with respect to aerodynamic characteristics of the blades for the power generating. Authors are not concentrated on the self-stating of HAWT because it depends solely on the wind turbine speed [4]. The conversion of wind power system is included Wind speed is the major factor for wind power production. Wind power could be generated from wind turbine by capturing wind speed though blades of wind turbine. Slightly change of wind speed which is directly influence to the wind power in order to the relation between wind power is directly proportional to the cube of wind speed. The wind power captured by rotor blade of wind turbine which can be expressed by [8].

$$P = 0.5\rho A C_p v^3 \tag{1}$$

Where, ρ is the density which is constant at hub height, swept area of the blades A, v is the wind speed and C_p is the power coefficient which means the percentages of wind power captured by wind turbine blades which depends on the pitch angle β of the blades and tip-speed ratio λ .

III. Adaptive Neuro-fuzzy Inference System

ANFIS is a system that is used for different purposes to classify, and to identify the chaotic time series. Besides that, ANFIS also works to generate input –output pairs as the basis of fuzzy 'if-then' rules construction by membership function and it also takes FIS system for tuning with the help of back propagation algorithm. The FIS has three components: rule base, database and reasoning mechanism. These systems show best performance like human expertise to recognize a specific domain [20]. In ANFIS, recognition of pattern and implementation in environment can done through neural network. To observe and verification of the model, half of the training date have to be used while others half would be used for testing purposes [21]. Therefore, ANIFS can accurately determine wind power prediction with the help of training phenomena as well filtered data. The first layer, which have two input variables, input 1 and input 2 for the purpose to give input value for the next layer. Every node in this layer has an adaptive node. The 'IF THEN' rules of Takgi and Sugeno's class of fuzzy logic with two inputs for the first order Sugeno are employed for the purpose of this study [22, 23].

If x is P and y is R, then

$$f_1 = p_1 x + q_1 y + r_1 \tag{2}$$

Each MFs weight can be checked through the second layer. In this layer, every node is non-adaptive which would multiply the input signals. The MFs layer refers as multiplier to receive to signal and send out the outcome in $w_i = \mu_{PQ}(x) * \mu_{PQ}(y)$ form

In the next layer, third one, is called rule layer. In the layer pre-condition of the fuzzy rules matching can be done by each node. Fuzzy rules and layer numbers are in equal number. Outcome of this layer is generally called normalized firing strengths.

$$(x) = tri(x; p_i, q_i, r_i) = \begin{Bmatrix} 0, & x \le p_i \\ \frac{x-p_i}{q_i-p_i} & p_i \le x \le p_i \\ \frac{r_i-x}{r_i-q_i}, & q_i \le x \le r_i \\ 0, & r_i \le\le x \end{Bmatrix} \tag{3}$$

The layer, defuzzification, can provides output values by inference of rules. In this layer, every node is an adaptive node function $0_i^4 = w_i^* xf = w_i^*(p_i x + q_i y + r_i)$ where the input variables are $\{p_i, q_i, r_i\}$.

$$0_i^5 = \sum_i w_i^* xf = \frac{\sum_i w_i f}{\sum_i w_i} \tag{4}$$

The last layer is called output layer that can to sum all the incoming inputs (coming from fourth layer) and to transform them. In addition, the fuzzy classification outputs converted by this layer. The structure of ANFIS is shown in Fig. 1.

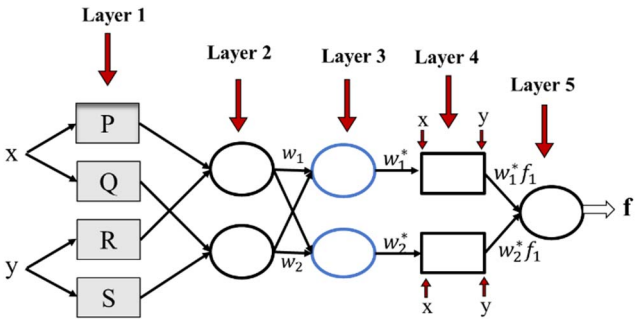
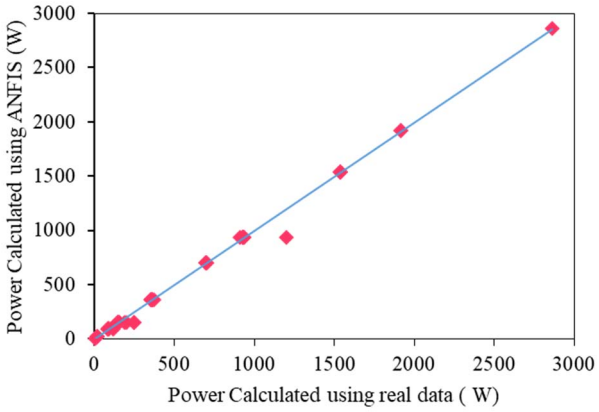


Fig. 1: The structure of ANFIS.

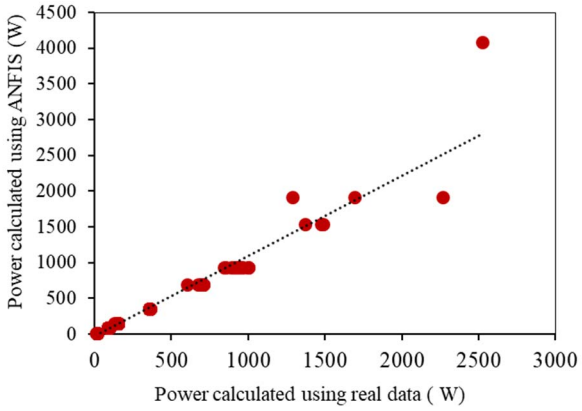
IV. Results and Discussion

To predict yearly wind power a model is developed by ANFIS methodology which is shown in this study. To develop the proficiency of ANFIS model; the comparison of wind power prediction at Kuala Lumpur and Melaka approaches have compared. As input parameters the temperature, wind speed, humidity and pressure data collected from the Metrology Department of Malaysia the wind power is used for first step analysis. These wind power values were calculated using the Equation (1). Those wind power values which were attained from both methods; were gathered and defined as training and testing data. To achieve this aim, firstly, three years averaged measured wind power data were used and divided to training phase (two years) and testing phase (one year). To estimate the wind power, at first the ANFIS network was trained with the extracted wind power data. The amount of wind power should estimate by the ANFIS network. To fuzzily the ANFIS inputs; three bell-shaped member ship functions were used. When the training period is finished then the remaining data set is used for testing the ANFIS network.

Fig. 2 shows scatterplot wind power values via the predicted wind power values against ANFIS model used measured data of Kuala Lumpur. Fig. 2(a) showed the results for the training data phase of wind power, whereas Fig. 2(b) stipulated the results for the testing phase. Evidently, most of the points fall along the diagonal line in training phase as well as in testing phase with slightly distinction; which concluded a suitable level of agreement between the estimated values by ANFIS technique and those obtained using measured data. According to the Fig. 2(a) and (b), as expected, the dispersion degree of the testing phase is slighter compared to the training phase. Similarly, in Fig. 3, scatterplots were showed the predicted wind power values via the determined wind power values against ANFIS model used measured data of Melaka. Moreover, Fig. 3(a) presented the results for the training phase, whereas Fig. 3(b) indicated the results for the testing phase of Melaka respectively.

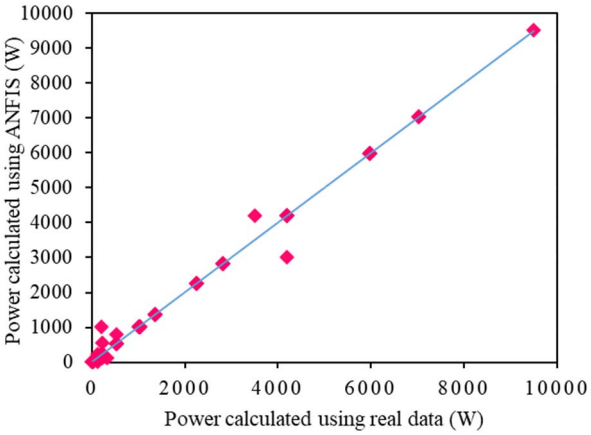


(a)

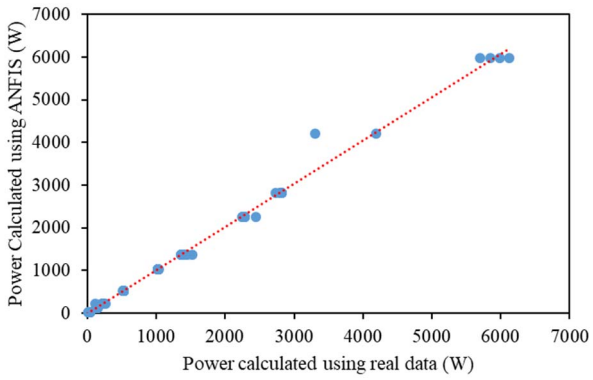


(b)

Fig. 2. Scatter plot of the wind power values based on measured data versus estimated ones via ANFIS for:(a) training data set and (b) testing data set at Kuala Lumpur



(a)



(b)

Fig. 3. Scatter plot of the wind power values based on measured data versus estimated ones via ANFIS for:(a) training data set and (b) testing data set at Melaka

Fig. 4 and Fig 5 display the measured and predicted wind power data of Kuala Lumpur and Melaka respectively. From the Fig. 4, the predicted is almost near to the measured wind power. Moreover, based on the indicators, the predicted wind power data is very close to measured data those predicted and measured data set evaluated by ANFIS. It can be concluded by above discussion, the performance of ANFIS shows the better result at Kuala Lumpur area to compare the Melaka area to based on the Table 1.

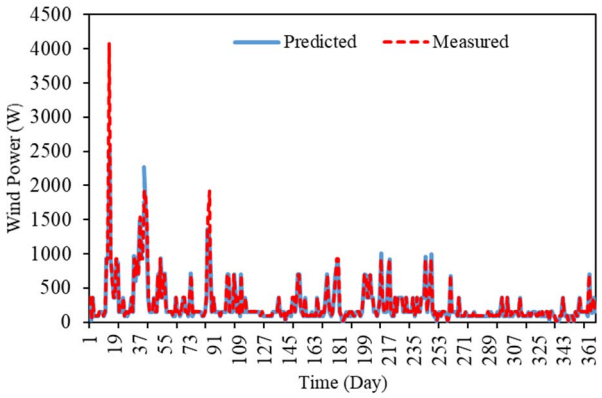


Fig. 4: Predicted data from testing ANFIS and real data of Kuala Lumpur.

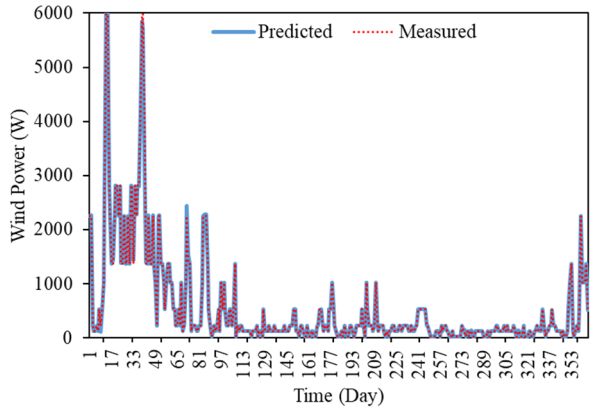


Fig. 5: Predicted data from testing ANFIS and Measured data of Melaka

TABLE I: PERFORMANCE OF PROPOSED ANFIS MODEL FOR WIND POWER PREDICTION USING FOUR INDICATORS

Model	Location	Training			
		RMSE	MAE	MAPE	R^2
ANFIS	KL	2.68	2.13	11.36	0.985
	Melaka	2.92	2.35	13.95	0.973
		Testing			
		RMSE	MAE	MAPE	R^2
	KL	1.65	1.87	12.35	0.968
	Melaka	1.82	1.98	13.23	0.944

Table 1 illustrations the four performance indicators which are evaluated the performance of ANFIS model for wind power prediction in Kuala Lumpur and Melaka at both training and testing stage. At training, performance of ANFIS model for Kuala Lumpur (RMSE=2.68, MAE=2.13, MAPE 11.36 and $R^2 = 0.985$) shows the slightly better than Melaka (RMSE=2.92, MAE=2.35, MAPE=12.95 and $R^2 = 0.973$). In addition, at testing stage, the accuracy of ANFIS model for wind power prediction in Kuala Lumpur shows satisfactory results comparison to the Melaka based on the performance indicators in shown in Table 1. According to the Table and above-mentioned discussion, it can be concluded that the ANFIS model enjoys high capability to predict daily wind power in Kuala Lumpur.

V. CONCLUTION

The reliable and accurate wind resources assessment is play and important role for harnessing wind power. In addition, the precise of wind power prediction is key parameter for wind turbine or wind farm installation. In this research Soft computing ANFIS model predict monthly wind power at Kuala Lumpur and Melaka areas in Malaysia. ANFIS model simulated in MATLAB. The outcome of this study shows that ANFIS model is more efficient at Kuala Lumpur to compare Melaka based on the performance indicator namely RMSE MAE, MAPE and R^2 from Table 1. The achieve results obeyed that the ANFIS model is the more capable the predicting wind power with favorable reliability and accuracy.

REFERENCES

[1] G. BoroumandJazi, R. Saidur, B. Rismanchi, and S. Mekhilef, "A review on the relation between the energy and exergy efficiency analysis and the technical characteristic of the renewable energy systems," *Renewable and Sustainable Energy Reviews,* vol. 16, pp. 3131-3135, 2012.

[2] M. R. Sarkar, S. Julai, C. W. Tong, O. Z. Chao, and M. Rahman, "MATHEMATICAL MODELLING AND SIMULATION OF INDUCTION GENERATOR BASED WIND TURBINE IN MATLAB/SIMULINK," *power,* vol. 2, p. 7, 2015.

[3] T. Ackermann and L. Söder, "An overview of wind energy-status 2002," *Renewable and sustainable energy reviews,* vol. 6, pp. 67-127, 2002.

[4] A. Mohamad, M. Amin, N. Amri, H. T. Toh, A. Majid, M. Shukry, *et al.*, "Review of Analysis on Vertical and Horizontal Axis Wind Turbines," *Applied Mechanics & Materials,* 2014.

[5] P. J. Schubel and R. J. Crossley, "Wind turbine blade design," *Energies,* vol. 5, pp. 3425-3449, 2012.

[6] A. Ucar and F. Balo, "Assessment of wind power potential for turbine installation in coastal areas of Turkey," *Renewable and Sustainable Energy Reviews,* vol. 14, pp. 1901-1912, 2010.

[7] A. U. Haque, P. Mandal, M. E. Kaye, J. Meng, L. Chang, and T. Senjyu, "A new strategy for predicting short-term wind

speed using soft computing models," *Renewable and sustainable energy reviews,* vol. 16, pp. 4563-4573, 2012.
[8] J. Yan and T. Ouyang, "Advanced wind power prediction based on data-driven error correction," *Energy Conversion and Management,* vol. 180, pp. 302-311, 2019.
[9] N. Bokde, A. Feijóo, D. Villanueva, and K. Kulat, "A Review on Hybrid Empirical Mode Decomposition Models for Wind Speed and Wind Power Prediction," *Energies,* vol. 12, p. 254, 2019.
[10] S. Fan, J. R. Liao, R. Yokoyama, L. Chen, and W.-J. Lee, "Forecasting the wind generation using a two-stage network based on meteorological information," *IEEE Transactions on Energy Conversion,* vol. 24, pp. 474-482, 2009.
[11] J. Liu, X. Wang, and Y. Lu, "A novel hybrid methodology for short-term wind power forecasting based on adaptive neuro-fuzzy inference system," *Renewable energy,* vol. 103, pp. 620-629, 2017.
[12] G. Osório, J. Matias, and J. Catalão, "Short-term wind power forecasting using adaptive neuro-fuzzy inference system combined with evolutionary particle swarm optimization, wavelet transform and mutual information," *Renewable Energy,* vol. 75, pp. 301-307, 2015.
[13] A. Tascikaraoglu and M. Uzunoglu, "A review of combined approaches for prediction of short-term wind speed and power," *Renewable and Sustainable Energy Reviews,* vol. 34, pp. 243-254, 2014.
[14] F. Cassola and M. Burlando, "Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output," *Applied energy,* vol. 99, pp. 154-166, 2012.
[15] J. L. Torres, A. Garcia, M. De Blas, and A. De Francisco, "Forecast of hourly average wind speed with ARMA models in Navarre (Spain)," *Solar Energy,* vol. 79, pp. 65-77, 2005.
[16] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renewable and Sustainable Energy Reviews,* vol. 13, pp. 915-920, 2009.
[17] R. Sarkar, S. Julai, S. Hossain, W. T. Chong, and M. Rahman, "A Comparative Study of Activation Functions of NAR and NARX Neural Network for Long-Term Wind Speed Forecasting in Malaysia," *Mathematical Problems in Engineering,* vol. 2019, 2019.
[18] O. Barambones, J. M. G. de Durana, and E. Kremers, "A neural network based wind speed estimator for a wind turbine control," in *MELECON 2010-2010 15th IEEE Mediterranean Electrotechnical Conference*, 2010, pp. 1383-1388.
[19] M. Islam, N. Rahim, K. Solangi, and R. Saidur, "Assessing wind energy potentiality for selected sites in Malaysia," *Energy Education Science and Technology Part A-Energy Science and Research,* vol. 29, pp. 611-626, 2012.
[20] M. Sarkar, S. Julai, C. Wen Tong, and S. F. Toha, "Effectiveness of Nature-Inspired Algorithms using ANFIS for Blade Design Optimization and Wind Turbine Efficiency," *Symmetry,* vol. 11, p. 456, 2019.
[21] D. Petković and S. Shamshirband, "Soft methodology selection of wind turbine parameters to large affect wind energy conversion," *International Journal of Electrical Power & Energy Systems,* vol. 69, pp. 98-103, 2015.
[22] A. Al-Hmouz, J. Shen, R. Al-Hmouz, and J. Yan, "Modeling and simulation of an adaptive neuro-fuzzy inference system (ANFIS) for mobile learning," *IEEE Transactions on Learning Technologies,* vol. 5, pp. 226-237, 2012.
[23] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International journal of man-machine studies,* vol. 7, pp. 1-13, 1975.

Cascaded H-Bridge Multilevel Inverter Using SVPWM Modulation Technique

Robin Ray*, Mashrur Ahmed Shadh**, Md. Shamim Reza
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
*email: robinray57@gmail.com
**email:mashrurahmed095@gmail.com

***Abstract*— In grid connected renewable energy generation system, using converter is a must. For this purpose, we proposed a multilevel H-bridge inverter that will contain lower third harmonic distortion. To achieve this target, higher number of output voltage level is necessary. Space Vector Pulse Wide Modulation (SVPWM) technique is very much efficient and relevant to our need because of its good utilization of dc bus voltage and reduced switching frequency here we implemented this technique. Using MATLAB/Simulink we achieved seventeen, twenty five and forty seven level of output voltage. We compared THD among them and found that THD was decreasing when output voltage level was increased. Finally THD below five percent could be achieved.**

Keywords- Multilevel Inverter, Dwell time, Cascaded H-Bridge Inverter, Space Vector Pulse Width Modulation (SVPWM), Third Harmonic Distortion(THD).

I. Introduction

Recently, there is a huge development and enhancement in the field of renewable energy[1]. Now-a-days in the developed countries renewable energy is one of the most expanding sector for power generation. All around the world lots of renewable energy plant has been established. Renewable energy requires dealing with DC-AC conversion. DC-AC conversion requires an inverter. With the development of semiconductor devices in recent times, we can develop highly efficient multilevel inverter[2][3] for conversion of DC into AC. There are a number of topologies and modulation techniques[4] which are being used in an inverter. The most commonly used topology is H-bridge topology and the most commonly used modulation techniques are Pulse Width Modulation (PWM) and Space Vector Modulation (SVM)[2][3][4]/ Space Vector Pulse Width Modulation(SVPWM)[5].In our proposal we used H-bridge as topology and modulation technique is Space Vector Pulse Width Modulation(SVPWM).SVM is one of the most complex method in both mathematical analysis and to implement in control circuit as well. The entire modulation process depends on three factors which are- voltage space vectors, transformation of a three phase time dependent system into a two co-ordinate time invariant system and effective pulse width modulation (PWM) which results in gate pulse generation. In spite of these complexity SVM is very much efficient and also gives better output.

II. Proposed Inverter

A. H-Bridge

In the field of research of power electronics multilevel inverters are of greater interest. Multilevel inverters are the most important technology for controlling high voltage and it also controls power distribution and power quality. The general structure of the multilevel converter is to synthesize a near sinusoidal which is actually a multilevel stair case like structure attained from a dc voltage to convert it into a perfect sinusoidal wave which is the required AC output. As more steps are added to the waveform, total harmonic distortion of the output wave decreases and it approaches zero as the number of levels increases. A cascaded multilevel inverter consists of a series of H-bridge inverter units. A H-Bridge inverter is available for both single phase and three phase conversion. In our proposal we used three phase topology for DC-AC conversion.

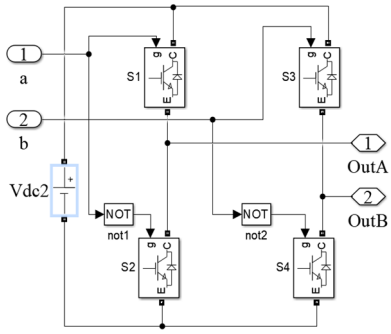


Figure1: H-Bridge circuit

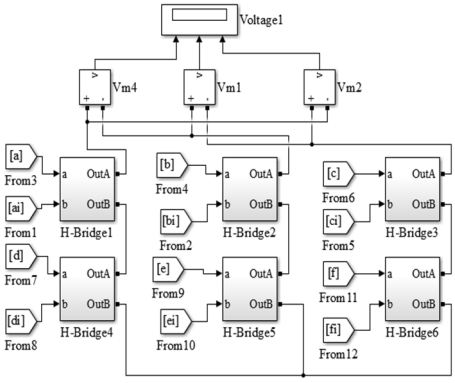


Figure2: Multilevel H-Bridge for 5 level output

As semiconductor device, we used MOSFET for switching with respect to gate impulse given by the controller circuit[6]. For multilevel output we cascaded[7] a number of H-bridge circuit For increasing the level we need to add H-bridges with the circuit. The equation is quite like number of levels = (2H+1); where H= number of bridges. Higher the levels higher the efficiency in output and higher is the number of H-bridges needed to be cascaded.

B. Modulation Technique

In electronics and telecommunication, modulation is a process which changes or modifies a particular property of the signal with help of a modulating signal. But in power electronics this modulation means the way or the process of controlling the on/off of gate devices to control the voltage output of the inverter as well as the harmonics of the inverter output as well. There are many types of modulation and some of them are highly efficient and advantageous in particular sectors of converter technology.The main aim of all these modulation techniques in inverter technology is to shape the output approximately similar to a sine wave. The quality of the modulation techniques are justified by the amplitude of the output, the harmonic contents, switching loss, effect of harmonics on the source, controllability etc. In case of hardware implementation the semiconductor devices and oscillators play a vital role. In our proposal we used Space Vector Modulation (SVM)/Space Vector Pulse Width Modulation (SVPWM) because of its efficient use of dc supply voltage, high number of level generation, low harmonic content and lower total harmonic distortion(THD).

C. SVM/SVPWM

The space vector PWM (SVM)[8] method is an advanced computation-intensive PWM method and is possibly the best method among the all PWM techniques for variable frequency drive application. Because of its superior performance characteristics, it has been finding wide spread application in recent years. Again this process has wide applications in motors [12], three-phase AC machines [13], electrical vehicles[14],solar PV cells[15]etc.In the Space Vector Pulse Width Modulation(SVPWM) technique, the duty cycles are computed rather than derived through comparison as in Sine Pulse Width Modulation. Each sector covers an area of 60 degree. In this process the sinusoidal voltage is treated as a constant amplitude vector rotating at constant frequency. Then the three phases are transformed into two axis or in other words it can be said that three-phase voltage vector is transformed into a vector in the stationary d-q coordinate frame which represents the spatial vector sum of the three-phase voltage. This PWM technique approximates thereference voltage V_{ref} by a combination of the eights witching patterns ($V_0$to V_7)[9]. The vectors ($V_1$to V_6) divide the plane into six sectors. The reference voltage is synthesized by using the nearest two neighboring non-zero vectors and zero vectors. The choice of the non-zero vectors depends upon the sector number in which the reference is located. Once the reference voltage is located, the vectors to beused for the SVM implementation to be identified. The

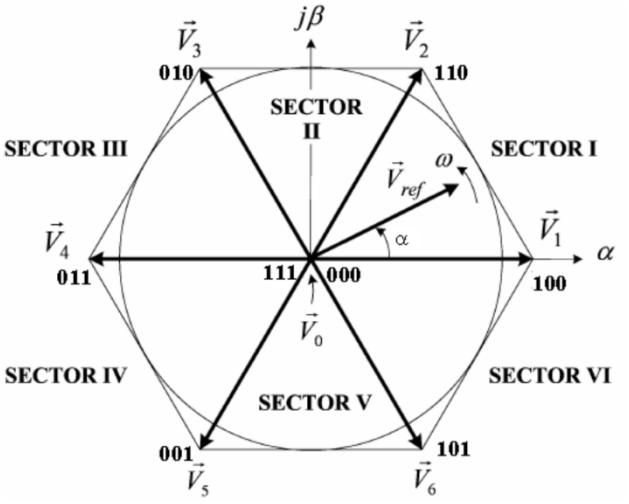


Figure 3: Space Vector Diagram

next task is to find the time of application of each vector which is called the dwell time. The output voltage frequency of the inverter is the same as that of the speed of the reference voltage and the output voltage magnitude is the same as the reference voltage.

D. Working Mechanism

In this paper, classical 2-level space vector technique has been used for calculation of on-time.Space vector diagram is a hexagon, consists of six sectors.Let's consider 3-phase sinusoidal voltage components are [10]

$$V_a=A\sin\omega t \quad \ldots\ldots\ldots\

If, $180 < \theta \leq 240$, sector =4

If, $240 < \theta \leq 300$, sector =5

If, $300 < \theta \leq 360$, sector =6

Then for each sector, we introduce a new angle γ

If, sector=1, $\gamma = \theta$

If, sector=2, $\gamma = \theta - \frac{pi}{3}$

If, sector=3, $\gamma = \theta - \frac{2pi}{3}$

If, sector=4, $\gamma = \theta + pi$

If, sector=5, $\gamma = \theta + \frac{2pi}{3}$

If, sector=6, $\gamma = \theta + \frac{pi}{3}$

Here let us consider, sampling period T_s and modulation index is denoted by a. So we find dwell time,

$$T_1 = T_s a \sin\left(\frac{pi}{3} - \gamma\right) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (7)$$

$$T_2 = T_s a \sin(\gamma) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (8)$$

$$T_0 = T_s - T_1 - T_2 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (9)$$

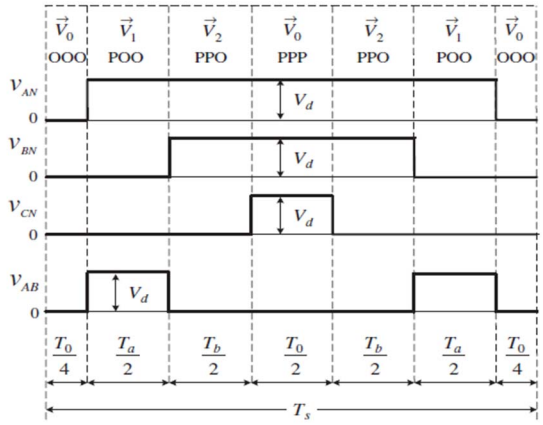


Figure 4: Segment switching sequence for $\overrightarrow{V_{ref}}$ in sector 1.

Table-1: Seven segment switching sequence for six sector

Sector	Switching Segments						
	1	2	3	4	5	6	7
I.	$\vec{V}_0$	$\vec{V}_1$	$\vec{V}_2$	$\vec{V}_0$	$\vec{V}_2$	$\vec{V}_1$	$\vec{V}_0$
	000	100	110	111	110	100	000
II.	$\vec{V}_0$	$\vec{V}_3$	$\vec{V}_2$	$\vec{V}_0$	$\vec{V}_2$	$\vec{V}_3$	$\vec{V}_0$
	000	010	110	111	110	010	000
III.	$\vec{V}_0$	$\vec{V}_3$	$\vec{V}_4$	$\vec{V}_0$	$\vec{V}_4$	$\vec{V}_3$	$\vec{V}_0$
	000	010	011	111	011	010	000
IV.	$\vec{V}_0$	$\vec{V}_5$	$\vec{V}_4$	$\vec{V}_0$	$\vec{V}_4$	$\vec{V}_5$	$\vec{V}_0$
	000	001	011	111	011	001	000
V.	$\vec{V}_0$	$\vec{V}_5$	$\vec{V}_6$	$\vec{V}_0$	$\vec{V}_4$	$\vec{V}_6$	$\vec{V}_0$
	000	001	101	111	101	001	000
VI.	$\vec{V}_0$	$\vec{V}_1$	$\vec{V}_6$	$\vec{V}_0$	$\vec{V}_6$	$\vec{V}_1$	$\vec{V}_0$
	000	100	101	111	101	100	000

After calculating dwell time, next step is to arrange switching sequence. Switching sequence must satisfy following two requirement to minimize device switching frequency.

1.The transition from one switching state to the next involves only two switches in the same H-bridge, one being switched on,and the other switched off

2. The transition of $\overrightarrow{V_{ref}}$ moving from one sector to another requires no or minimum number of switching.[11]

Here, even segment switching sequence for (V_{ref}) in sector 1 is shown in Figure 4. For all 6 sectors of space vector diagram, we find the switching sequence as figure 5. Using these switching sequences, 3 switching signals are found. Then these switching signals are considered as modulating signal and 8 carrier signals are taken .The carrier signals are triangular signal. A level shifted PWM [16][17] was performed between switching signal and carrier signals

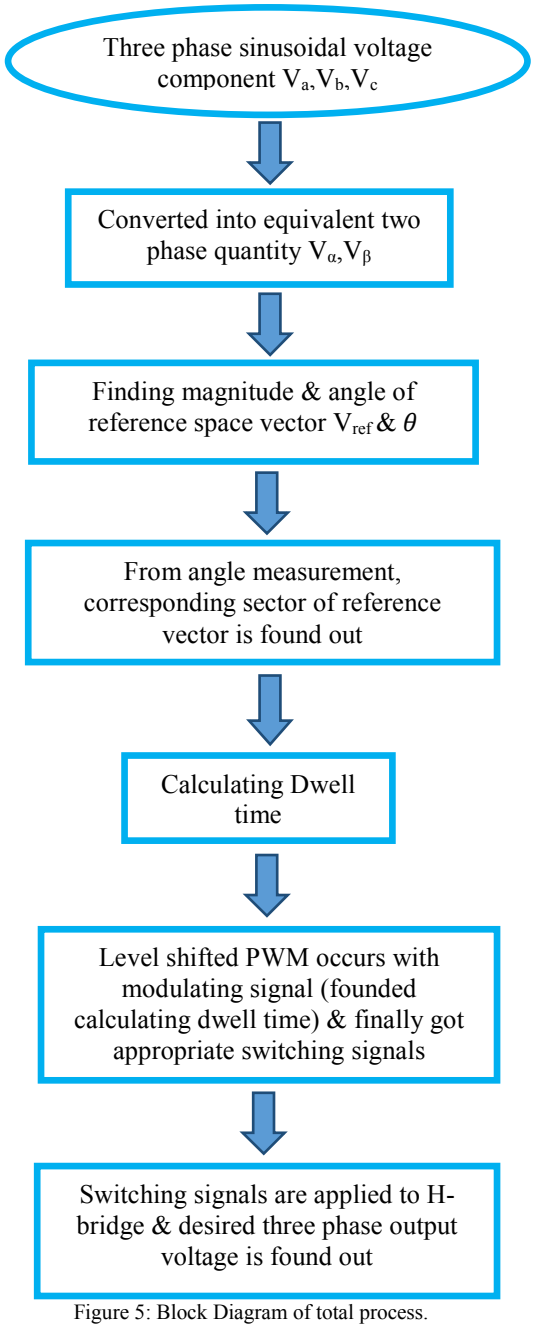


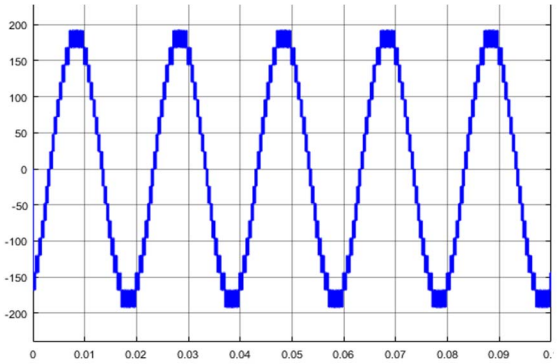
Figure 5: Block Diagram of total process.

and final output switching signals were found out .These final switching signals are applied to H-bridge MOSFET's gate.

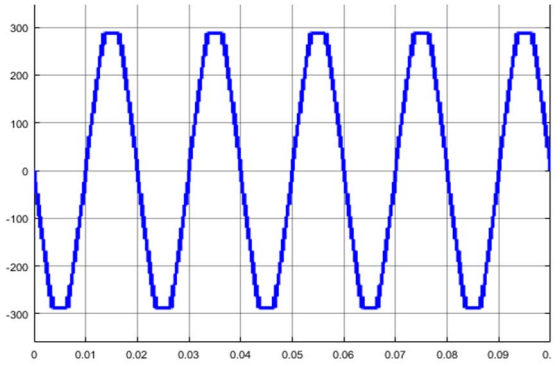
In cascaded H-Bridge Inverter, number of voltage level is denoted as m=2H+1,where H is number of H Bridge .As we cascaded 8, 12 & 23 H-bridge in each phase, so respectively. 17,25 & 47 level phase to phase output voltage could be achieved.

III. Simulation Result

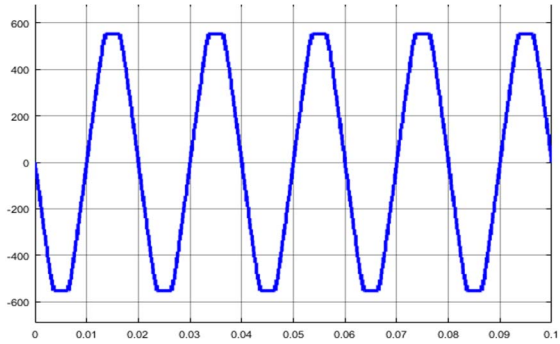
Simulation was performed using MATLAB/Simulink[18] to achieve 17,25 & 47 levels output voltage. In fig 6,output voltage curve and in fig 7,THD spectra has been shown. In table 1, comparison of THD between these 3 output has been shown. It has been seen that THD decreases with increase of number of output voltage levels.



(a)

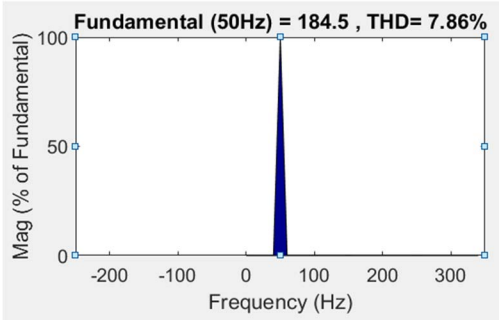


(b)

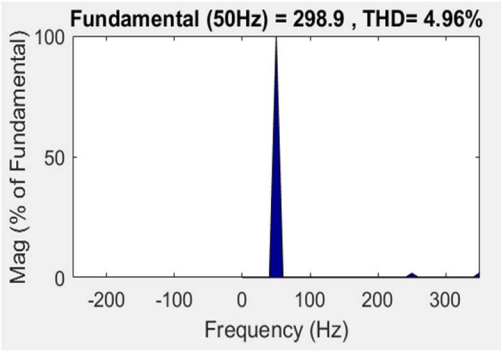


(c)

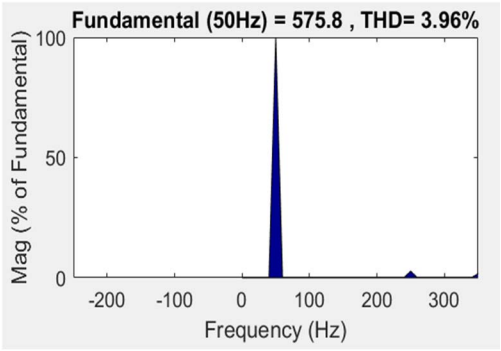
Figure 6: Phase to phase output voltage (a) output voltage for 17 level (b) output voltage for 25 levels (c) output voltage for 47 levels



(a)



(b)



(c)

Figure 7: Spectrum Analysis Graph of Cascaded H Bridge invertor (a) spectral analysis for 17 level output (b) spectral analysis for 25 level output (c) spectral analysis for 47 level output.

Table-2: Third Harmonic Distortion for different output level

Number of Levels	THD
17	7.86%
25	4.96%
47	3.96%

Here, it is seen that for increasing number of levels we get lower total harmonic distortion. That is as the number of levels are increased output becomes more efficient and perfect as expected.

IV. Conclusion

Here we proposed several designs including 17, 25 and 47 level Multilevel H-bridge inverter using SVPWM.We analyzed third harmonic distortion (THD) and found that it decreases with number of level increases. Though we can simulate higher level in Simulink but in hardware implementation we may face some complexity. Due to switching losses increases with increase in level, desired output may be distorted. Renewable energy is getting popular day by day. As here we find desired output voltage and low THD as well as higher efficiency. So in case of renewable energy, we will be much more benefited.

References

[1] F. Sevilmi and H. Karaca, "An Experimental Study of Grid Tied Inverter for Renewable Energy Systems," *2018 Int. Conf. Smart Energy Syst. Technol.*, pp. 1–6, 2018.

[2] I. Ahmed *et al.*, "Simplified Space Vector Modulation Techniques for Multilevel Inverters," vol. 8993, no. c, 2016.

[3] H. M. Inverter, T. H. Distortion, S. Pwm, and T. Svpwm, "Chapter 4," no. 2006, pp. 61–111,

[4] A. K. Gupta, S. Member, A. M. Khambadkone, and S. Member, "A Space Vector PWM Scheme for Multilevel Inverters Based on Two-Level Space Vector PWM," vol. 53, no. 5, pp. 1631–1639, 2006.

[5] A. Patil, A. Khot, C. Awagathe, S. Pillai, and P. Kumar, "Control Techniques for Multilevel H–Bridge Inverters," vol. 010, no. April 2015.

[6] I. S. Techniques, "Chapter 4 Control Algorithm For Proposed H-Bridge," no. 2002, pp. 65–94.

[7] T. H-, "Chapter 3 Cascaded H-Bridge Multilevel Inverter," vol. 1, pp. 39–60.

[8] S. Vector, M. Three, P. Voltage, and S. Inverter, "Space vector modulated three phase voltage source inverter 4.1.," pp. 45–68.

[9] C.Science, "Simulation And Implementation Of Two- Level And Three-Level Inverters By Matlab And Rt-Lab," 2011.

[10] José Rodríguez, Jih-Sheng Lai, Fang Zheng Peng, "Multilevel Inverters: A Survey of Topologies, Controls, and Applications", IEEE Transactions on Industrial Electronics, Vol. 49, No. 4, Aug, 2002.

[11] I. Ahmed and V. B. Borghate, "Simplified space vector modulation technique for seven-level cascaded H-bridge inverter," *IET Power Electron.*, vol. 7, no. 3, pp. 604–613, 2014.

[12] D. Based and I. Motor, "Application of Cascaded H -Bridge Multilevel Inverter in DTC-SVM Based Induction Motor Drive," 2012.

[13] N. P. Quang and J.-A. Dittrich, Vector Control of Three-Phase AC Machines, vol. 2. 2015.

[14] K. Abdelfatah *et al.*, "Space Vector PWM Control Synthesis for a H-Bridge Drive in Electric Vehicles To cite this version : Space Vector PWM Control Synthesis for a H-Bridge Drive in Electric Vehicles," no.c, 2014.

[15] M. Tawhidha, A. R. Begum, S. Balakrishnan, S. E. Peter, A. S. Priyadharshini, and K. Vinoth, "Design and Simulation of Seven Level Cascaded H-Bridge Quasi Z Source Multilevel Inverter for Solar PV System," vol. 2, no. 5, pp. 1–6, 2015.

[16] M. Angulo, P. Lezana, S. Kouro, J. Rodríguez, and B. Wu, "Level-shifted PWM for cascaded multilevel inverters with even power distribution," *PESC Rec. - IEEE Annu. Power Electron. Spec. Conf.*, pp. 2373–2378, 2007.

[17] I. Ahmed and V. B. Borghate, "Simplified space vector modulation technique for seven-level cascaded H-bridge inverter," *IET Power Electron.*, vol. 7, no. 3, pp. 604–613, 2014.

[18] C. Paper, "Matlab / Simulink Model of Space Vector PWM for Three-Phase Voltage Source Inverter," vol. 2, no. OCTOBER 2006, pp. 1096–1100, 2015.

A Modified Naïve Bayesian-based Spam Filter using Support Vector Machine

Md. Sabir Hossain
Department of Computer Science and Engineering
Chittagong University of Engineering and Technology (CUET), Chittagong-4349, Bangladesh
sabir.cse@cuet.ac.bd

Md. Zubair
Department of Computer Science and Engineering
Chittagong University of Engineering and Technology (CUET), Chittagong-4349, Bangladesh
zubairhossain773@gmail.com

Mohammad Obaidur Rahman
Department of Computer Science and Engineering
Chittagong University of Engineering and Technology (CUET), Chittagong-4349, Bangladesh
obaidur_91@gmail.com

Muhammad Kamrul Hossain Patwary
Institute of Information and Communication Technology
Chittagong University of Engineering and Technology (CUET), Chittagong-4349, Bangladesh
muhammadkamrulhossain@gmail.com

Md. Golam Sarwar Rajib
Department of Computer Science and Engineering
Chittagong University of Engineering and Technology (CUET), Chittagong-4349, Bangladesh
gsr.rajib@gmail.com

Abstract*— The ever-growing problem which is threatening the current mailing system is spam. Spam is nothing but an unsolicited bulk e-mail frequently sent in a financial nature which generates the need for creating an anti-spam filter. Amongst many spam filtering techniques, the most advanced method "Naïve Bayesian filtering" using the Support Vector Machine (SVM) have been implemented. Spammers are very careful about the filtering techniques. For that very reason, dynamic filtering is needed and the proposed method meets the demand. The algorithm splits the received email into tokens and uses Bayes' theorem of probability to calculate the probability of spam for each token to determine the total spam probability of the mail. Implementation of SVM instead of corpora is one of the added features of the algorithm. The most challenging feature was to take the words as well as whole sentences as input in the SVM as tokens and feature vectors. The inclusion of sentences in the dataset training has increased the accuracy of detecting spam and ham. Natural Language Tool Kit (NLTK) has been used as a useful language processing tool to tokenize the sentences and also to understand the meaning of the same types of sentences to some extent. As a test mail is being compared by word to word and also sentence to sentence from the training datasets to determine if the mail is spam or not, it will improve the performance of the filter. With some simple modifications, the filter can be used in both server and client end. The efficiency increases gradually with the increased number of email it processes.

Keywords- Spam, Bayesian Approach, SVM, Tokenization, Spamicity, Dataset.

I. INTRODUCTION

The internet and E-mail become major communication media in every aspect of our life, they have revolutionized the way of doing business and socializing. The Internet has made information gathering and publishing easy and provided the convenience of online shopping and financial management. However, these new technologies have created new issues and problems at the same time. Among the issues, the so-called 'spam' is the major puzzle.

The word "Spam" as applied to email means Unsolicited Bulk Email (UBE). It can be said in another way that the recipient has not any granted way to detect the sender of mail containing a large collection of messages. Some people may include the e-mail produced by mass-mailing viruses or Trojan horses as spam.

The manifestation of spam has become a fundamental problem for internet user in general. The recipients of bulk emails get puzzled while differentiating the important mail from spam. It also reduces the efficiency of a company and somewhat more. Apart from direct losses, indirect losses are made as well as consumption of storage, internet bandwidth and so on. Some folks even predict that the arrival of spam will bring about the end of e-mail entirely.

Dealing with the issues, numbers of spam detection and filtering methods have been developed. These techniques are effective in some degrees but not limitation free. Signature-based filtering is accurate and fast but applicable for particular emails. On the contrary, the Learning-based system is slower. Above techniques are based on pattern matching rules. Most often it requires tuning for each user's messages which is time-consuming and difficult. Furthermore, the spammers are very aware of the spam detection techniques and try to overlook the techniques variously. An automated learning system that would able to separate the legitimate message is highly suggested.

This paper examines and implements the cutting edge anti-spam technique 'Naïve Bayesian filtering' using Support Vector Machine (SVM) which is learning based dynamic solution. It can adapt to the new techniques of spammers by enriching the trained dataset.

II. RELATED WORK

The naïve Bayesian spam filter is a member of the Learning-based spam filter family. At the very beginning,

Pantel and Lin developed a spam filtering technique based on text classification [1]. Afterward, many scientists and researchers worked on Bayesian spam filtering [2]. Paul Graham proposed a modified algorithm with a large dataset in 2003 and be able to extend the accuracy rate and minimize the probability of false positive significantly [3]. The experiment of Pantel and Lin showed their success rate is 92% with 1.16% false positive. On the other hand, Paul's algorithm was able to detect 99.5% spam with 0.03% false positive. Paul claimed that the main reason behind it was working with a large number of dataset.

Biju Issac and Wendy J. Jap developed a system using Porter's stemmer algorithm along with Bayes theorem. They used porter's stemmer algorithm to split every keyword to its stem which improved the efficiency. Because it reduced the keywords to be searched and developed the margin of accuracy. [4]

Lin Li & Chi Li also implemented a system based on the Naïve Bayesian spam filter. They use the TF-IDF method for more accuracy [5].

Spam detection with K-means clustering algorithm can also be used to detect spams. It also provides a promising result [6].

Naïve Bayes is a strong tool for spam filtering. There are several forms of Naïve Bayes which gives different accuracy [7].

In some cases, unsupervised learning provides more accurate result and it is best suited [8]. Another most recent technique is deep learning. Researchers are also trying to deploy a technique based on deep learning in spam detection [9].

Although all of the solutions discussed and implemented above are effective to some degree, each of the processes has limitations too. None of the solutions can provide 100% accuracy, and there's always a percentage of false positive. No single method is enough to detect and filter spam because spammers are very much aware of the modern spam detection techniques and they are smart enough to circumvent a spam filter. For this reason, a "cocktail" approach is used to detect spam successfully. Combination of the strength of each system may provide a robust solution in case of spam detection.

III. PRELIMINARIES

A. Various Spam Filtering Techniques

Spammers are very conscious of spam filtering techniques. They try their best to overlook existing spam filtering techniques. For meeting up the demand, over the years many filtering techniques have been introduced for filtering spam emails. Signature-based, rule-based and learning-based filtering are some of them.

- In the signature-based filtering, the hash code of known spam is stored in the database. If an incoming email's signature is matched with the stored signature, then the incoming mail is said to be spam. The process is extremely accurate, but it can't detect a new technique of spam.
- The rule-based filter uses a list of keywords; these keywords are the spam keywords. If an incoming mail contains these keywords, then it is considered as spam. The keywords are stored in the database. These filters have high false positive and negative rates. It's because of specific words.
- In learning-based filtering, the filter is trained by itself from time to time according to its input. This approach is a dynamic one. As the spammers are being more conscious of the conventional spam filtering techniques, learning based filtering is the best solution for it.

Through the experiment Bayesian spam filter is used which is one of the learning based spam filters. This system tokenizes the received mail and compares them with the tokens stored in the database. It requires an initial training period. Throughout the training period, the Bayesian filter supervises the incoming and outgoing email and started to create a dataset for the mail flow. System administrators may also tweak the database to better train the Bayesian engine.

B. Bayes' Theorem in Spam Filtering

Computing conditional probabilities, Bayes' formula is one of the famous formulae. Generalized Bayes' formula:

$$P(B_1|A) = \frac{P(B_1 \cap A)}{P(A)} = \frac{P(B_1)P(A|B_1)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)} \quad (1)$$

Where,

- $P(B_1|A)$ is a conditional probability that indicates occurring $P(B_1)$ given that P(A) is true.
- $P(A|B_1)$is also a conditional probability that indicates the probability of occurring P(A)given that $P(B_1)$ is true.

B. Naïve Bayes' Formula in Spam Filtering

Bayesian email filters are based on Bayes' theorem (1). The theorem is most often used in the field of spam detection.

Naïve Bayes' theorem is given below which is used to calculate the spam probability of the system.

$$P_r(S|W) = \frac{P_r(W|S).P_r(S)}{P_r(W|S).P_r(S) + P_r(W|H).P_r(H)} \quad (2)$$

Where,

- Pr(S|W) denotes the probability of spam message when the spam detecting word is in it.
- Pr(S) denotes the overall probability of spam message.
- Pr(W|S) denotes the probability of spam word in the spam messages.
- Pr(H) denotes the overall probability of spam message
- Pr(W|H) denotes the probability of spam word in the ham messages.

Initially when the mail received it is assumed that the mail is neither spam nor ham i.e., the probability of the mail to be spam or ham is equal, so Pr(S) = Pr(H) = 0.5.

a. Calculating individual probability

The overall probability of an email being spam, regarding all of the tokens or a set of tokens of the mail can be calculated using the following formula**:**

$$P = \frac{P_1 P_2 \ldots\ldots\ldots P_N}{P_1 P_2 \ldots\ldots\ldots P_N + (1-P_1)(1-P_2)\ldots(1-P_N)} \quad (3)$$

Here,

- P is the probability spam of a fishy message.
- P_1 is equal to the probability of $P_r(S|W_1)$.
- P_N is the probability $P_r(S|W_n)$

The output of the formula p is usually compared to a defined point for deciding a spam or ham message. If the probability is greater than the specified value, the message is considered spam. Otherwise, it is not a spam message.

b. Overall View of Naïve Bayes' Formula:

Naïve Bayes is easy, but effective classifier.it is totally based on the Bayes theorem. This feature defines feasibility related to a class based on the probability of each attribute value. Abstractly, NB is a conditional probability model: given a problem instance to be classified to any class {Cj }1<j<n, represented by a vector of features (independent variables) X € {x1…..xn} it can be formally shown as:

$$C_{MAP} = argmax_{c \epsilon C} P(x_1, \ldots\ldots\ldots, x_n | c) P(c) \quad (4)$$

$$C_{NB} = argmax_{c \epsilon C} P(c_j) \prod_{x \in X} P(x|c) \quad (5)$$

Reducing the number of parameters required for modeling the important advantages of the NB algorithm in technical concerns. Exemplary, due to the independence assumption, we need 2n parameters to model P (X|Y) instead of the original 2(2n – 1). These quality procedures guarantee simplicity and speed.

C. Support Vector Machine (SVM)

SVM is associated with statistical learning theory. In 1992(Boser, Guyon & Vapnik), it was first introduced properly [10]. SVM algorithm became popular after getting success in handwriting digit recognition.

Training data consists of input and output functionality is known as supervised learning in machine learning. The support vector machine (SVM) is one of the supervised machine learning algorithms [11]. It is mostly used for solving two-group classification problems. In this system, it is used for differentiating between spam and ham.

An SVM training algorithm builds a model that enrolls new examples to one of the two categories which makes the non-probabilistic binary linear classifier. If the new example which will be classified falls into one side of the hyperplane, it is defined in the category according to the class of that side.

Beside linear classification, different types of method are introduced like kernel trick, etc. for non-linear classification in SVM.

D. Natural Language Tool Kit (NLTK)

As the first challenge to create a spam filter was to build an SVM that will take input an email and extract the words as well as sentences of the message body. So that there will be word to word and sentence to sentence comparison simultaneously, to calculate spam probability. For this approach, the Natural Language Tool Kit (NLTK) is used in python programming.

The NLTK is introduced for helping the system with Natural Language Processing (NLP) technique. The NLKT helps the system by splitting sentences from paragraphs, splitting up words, recognizing the part of speech of those words and so on. However, the primary purpose of using NLTK is the tokenization of sentences and words.

IV. PROPOSED METHODOLOGY

A. Algorithm for Building the filtering system

The total process is given below sequentially.

1) Retrieving spam and ham email messages

A lot of sample ham and spam messages are needed to train the SVM. Almost 16,000 ham messages and 3,000 spam messages have been collected from the machine learning repository of Enron dataset [12]. The system retrieves all the messages one by one and takes it into the machine.

2) Tokenization

Tokenization is done into two steps. Each email is split into two parts, words and sentences are separately taken. Words are directly listed into a database which is called data dictionary. Using NLTK, the sentences are converted into a vector and then stored to the data dictionary.

3) Feature Extraction

Feature extraction measures the frequency of spam and ham emails. Generally, each of the words is labeled 0 if it came from the ham email and labeled 1 if it is from spam email. In this method, the vectors generated from sentences are also labeled 0 and 1 in accordance with spam and ham messages. In this way, SVM calculates the frequency of the tokens.

4) Shuffle the data dictionary

After labeling, all the tokens are stored in the data dictionary in the format of the 2D array in a random approach. The spam and ham email messages are stored randomly in the data dictionary.

5) Generate. sav file

After completion of all the steps above, SVM generates a .sav file which contains the training datasets of the data we have collected to train the system.

6) *SVM. Seeding database*

The data dictionary has been seeded with spam and ham messages as many as possible. The more data added, the more accuracy of a mail being ham or spam will be achieved.

7) *Test method (Input message)*

After training the SVM with enough datasets of spam and ham emails, the testing session has been taken place. A sample email body has been taken to test.

8) *Tokenization of test message*

The whole message is split into two portions. One portion takes the word and keeps them in an array and another portion converts every sentence into a feature vector and keeps it in an array.

9) *Frequency calculation*

After tokenizing the test message, the frequency of each token and vectors in the token list is calculated.

10) *Retrieve Spam and Ham frequency*

In this step, the spam and ham frequency is retrieved from the training datasets of our SVM model.

11) *Retrieve Spam and Ham count*

Spam and ham count are done from the datasets.

12) *Calculate Spam probability*

Calculate the spam probability of each token that is found in the database using the following formula:

Spam probability, P_s

$$P_S = \frac{\text{spam frequency in database}}{\text{number of spam mail in database}} \quad (6)$$

If the probability is greater than 1, set it to 1.

13) *Calculate Ham Probability:*

Calculate the spam probability of each token that is found in the database using the following formula:

Ham probability P_H,

$$P_H = \frac{Ham\ frequency\ in\ database}{number\ of\ Ham\ mail\ in\ database} \quad (7)$$

If the probability is greater than 1, set it to 1.

14) *Calculate Spamicity*

Calculate the spamicity of each token that is found in the database output of "(6)" & "(7)" using the following formula:

$$\text{Spamicity} = \frac{P_S}{P_S + P_H} \quad (8)$$

15) *Total spam probability:*

Calculate the total spam probability using the output of "(8)". It can be found as:

$$\text{Total probability} = (S_1 * S_2 * .... * S_n) + ((1-S_1) * \ldots * (1-S_n)) \quad (9)$$

Where,

- S_1 is the spamicity of 1st token.
- S_2 is the spamicity of 2nd token.
- S_n is the spamicity of nth token.

16) *Decision*

We examined the proposed system with different threshold values. Which shows the results in Table I.

TABLE I EXPERIMENTAL ACCURACY WITH DIFFERENT THRESHOLD VALUES

Value	.2	.3	.4	.5	.6	.7	.8
Accuracy(%)	39.11	60	80	90	85	70.21	43.55

When the threshold value is less than 0.5, the system considers some spam emails as ham. On the other hand, some ham email is considered as spam with the threshold value greater than 0.5. So, we considered 0.5 as an optimal threshold value.

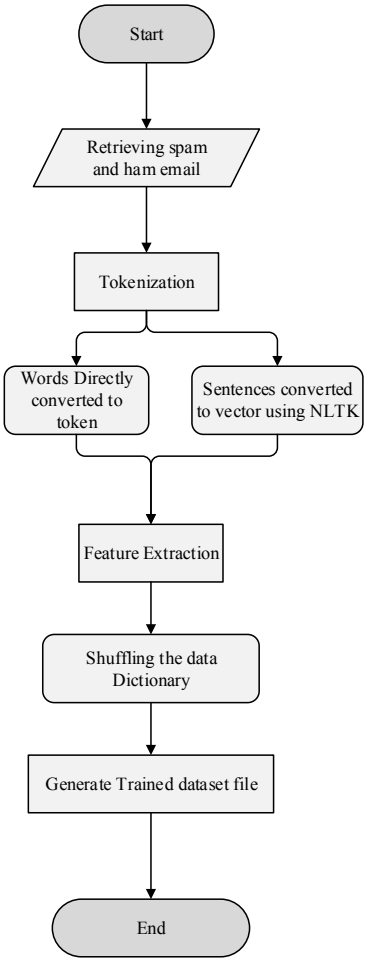


Fig.1 Representation for building Support Vector Machine (SVM)

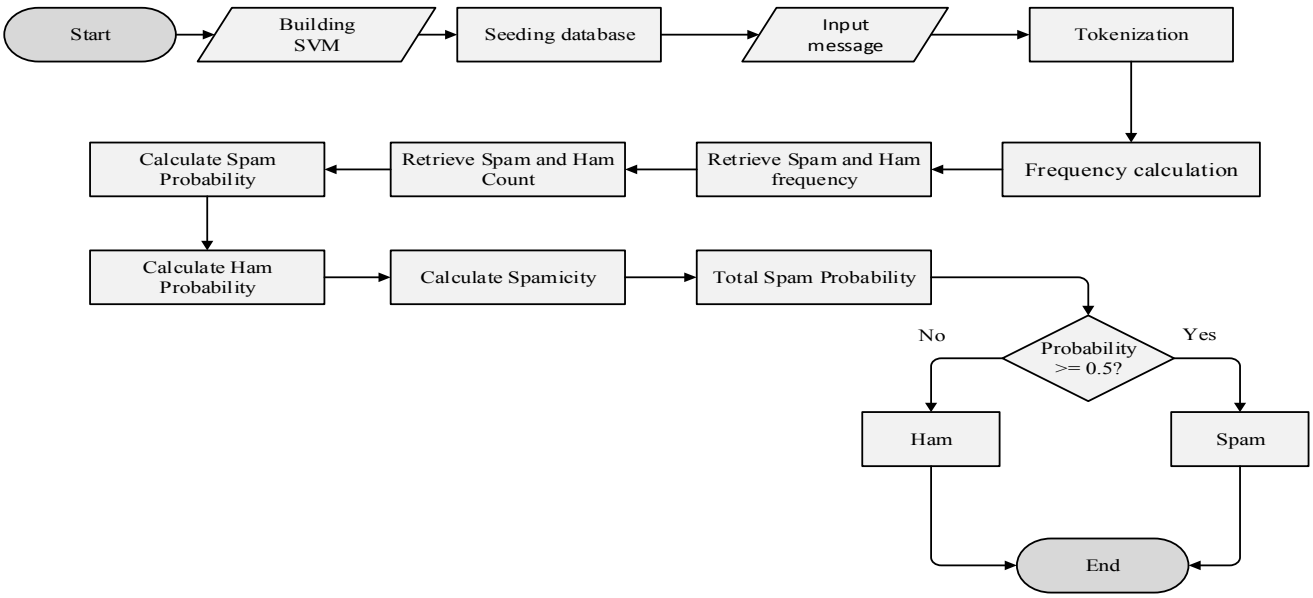


Fig. 2 Representation of Naïve Bayesian Spam Filter using SVM

B. Flowchart of the Naïve Bayesian spam filter using SVM

1) Building SVM

First of all, we have built an SVM for training datasets.

The flowchart of building SVM is in fig.1.

2) The complete procedure of Spam Filter

After successfully building the SVM, we prepared the spam filter for testing new mail. The whole process is shown in fig. 2.

V. Implementation of BayesiaN Filter

A. Sample spam message

The result of the algorithm after receiving a spam email and processing it with the system is shown using a well-known "Enron datasets" [12] spam message. The message is given below:

From: BUMA SARO WIWA <bsarowiwa@incamail.com>
To: imukaviva@incamail.com
Subject: Dobmeos with hgh my energy level has gone up! stukum introducing doctor formulated hgh.

-increased muscle strength. - loss in body fat. - increased bone density. - lower blood pressure. - quickens wound healing. - reduces cellulite. - improved vision. - wrinkle disappearance. -increased skin thickness texture.	- increased energy levels. - improved sleep and emotional stability. - improved memory and mental alertness. - increased sexual potency. - resistance to common illness. - strengthened heart muscle. - controlled cholesterol. - controlled mood swings. -new hair growth and color restore.

It is referred to in medical science as the master hormone. it is very plentiful when we are young, but near the age of twenty - one our bodies begin to produce less of it. by the time we are forty nearly everyone is deficient in hgh , and at eighty our production has normally diminished at least 90 - 95 % advantages of hgh :

Read more at this website www.hmboe.com.

Subscribe today.

B. Processing steps

1) Tokenization: After receiving the mail, the Bayesian spam filter first split the message into tokens. For the concern message tokens are shown in "table II". At the same time, sentences are separated too by counting the comma (,), full stop (.), exclamatory sign (!), interrogative sign (?), etc. punctuation marks. It is shown in "table III".

2) *Frequency Calculation*: Using the stop words list, we eliminate the unnecessary words that are not effective in spam detection. Then the frequency of single tokens is calculated. After that, with the help of NLTK, the sentences are converted into feature vectors, and then their frequency is also calculated.

TABLE II TOKEN LIST (WORDS)

Dobmeos	with	hgh	My	energy
level	has	Gone	up	stukm
human	growth	hormone	also	Called
hgh	Is	Referred	To	in
medical	Master	hormone	It	Is
very	plentiful	When	we	are
Young	but	near	The	Age
Of	Twenty	one	our	Bodies
Begin	to	Produce	less	of
It	by	the	Time	we
are	Forty	nearly	everyone	Is
deficient	of	In	hgh	and
At	eighty	our	Production	Has
normally	Diminished	at	least	90-95%
advantages	of	Hgh	Increased	muscle
Strength	loss	In	Body	fat
increased	Bone	density	lower	Blood
Pressure	quickens	Wound	healing	reduces
Cellulite	improved	vision	Wrinkle	disappearance
increased	Skin	thickness	texture	Increased
energy	levels			

3) Retrieve Spam and Ham Frequency: Retrieve Spam and Ham Frequency: From the SVM, the data dictionary has retrieved where the spam and ham frequency is stored.

TABLE III TOKEN LIST (SENTENCES)

Sentence
Dobmeos with hgh my energy level has gone up
Introducing doctor – formulated hgh
human growth hormone - also called hgh
It is referred to in medical science as the master hormone
it is very plentiful when we are young
but near the age of twenty - one our bodies begin to produce less of it
by the time we are forty nearly everyone is deficient in hgh
and at eighty our production has normally diminished at least 90 - 95 %. advantages of hgh
increased muscle strength
loss in body fat
increased bone density
lower blood pressure
quickens wound healing
reduces cellulite
improved vision
wrinkle disappearance
increased skin thickness texture
increased energy levels
improved sleep and emotional stability
improved memory and mental alertness
increased sexual potency
resistance to common illness
strengthened heart muscle
controlled cholesterol
controlled mood swings
improved memory and mental alertness
Read more at this website www
Subscribe today

4) Retrieve spam and ham count: After getting the spam and ham frequency, we now go through the spam and ham count. Here numbers of spam and ham messages from the data dictionary are found.

5) Calculate spam and ham probability: In this step, the spam and ham probability of each token, and vectors generated from sentences are calculated using the formula discussed above. If either spam or ham probability is found greater than 1, and then it is set to 1.

6) Spamicity Calculation: The spamicity of each token is calculated by using spam and ham probability using the formula discussed above.

7) Total spam probability: In our test mail, there are 122 tokens for words, and 27 vectors generated from sentences. It is really a lengthy calculation, so it would take some time in the compiler to show the result.

Using Naïve Bayes formula, the total probability of the message to be a spam can be calculated as:

Term 1: (0.99911) (0.654) (0.756) (0.889) (0.776) (0.445) (0.667) (0.334) (0.554) (0.233) (0.556)……. (0.912) (0.814) (0.823)

Term 2: (1-0.99911) (1-0.654) (1-0.756) (1-0.889) (1-0.776) (1-0.445) (1-0.667) (1-0.334) (10.554) (1-0.233) (1-0.556) ……. (1-0.912) (1-0.814) (1-0.823)

$$Total\ spam\ probability = \frac{Term\ 1}{Term\ 1 + term\ 2} = 0.74 \quad (10)$$

We can see that the final probability of the message to be spam is 74%. In our methodology, we have already stated that our probability measurement threshold value is 0.5. If the probability is greater or equal to 0.5, the message will be considered spam, and if less, we will consider it ham.

VI. ANALYSIS OF THE PROPOSED SPAM FILTER

A. Experimental Result

We have tested the performance of the spam filter using the corpus of the SpamAssassin project of Apache Foundation [13]. For some sample input email, the output of the Naïve Bayesian spam filter is given in "table IV" & "table V".

TABLE IV EXPERIMENTAL RESULT OF NÄIVE BAYESIAN FILTER

Steps	Total no. of Email (input)	No. of Spam (input)	No. of Ham (input)	No. of Spam (output)	No. of Ham (output)
initially	120	45	75	-	-
1	30	10	20	9	18
2	60	30	30	28	27
3	90	38	52	34	45
4	120	45	75	39	68

Experimental result of false positive, false negative and accuracy of "table IV" is shown in "table V".

TABLE V EXPERIMENTAL RESULT OF NÄIVE BAYESIAN FILTER

Steps	False Positive (Output)	False Negative (Output)	Accuracy (%)
initially	-	-	-
1	1	2	90
2	2	3	91.67
3	4	7	88
4	6	7	87

We considered a total of 120 emails among them 45 spam and 75 ham/nonspam. The input emails are provided in four steps. At step-1 and step-2, step-3 and step-4 total 30, 60, 90, 120 emails are used respectively.

During the experiment for step-1 and step-2, we used one spam and one ham alternatively. This is done so that there is a balance between the spam and ham database. For the next two steps, no such order was maintained.

B. Comparative Analysis

Naïve Bayesian spam filters are adequate to detect spam than other types of filters such as K-means clustering [6], deep learning [9]. But supervised learning is better than unsupervised learning in this case, and deep learning is a complex process to classify spam and ham. So, we can say that our filtering technique using a Naïve Bayesian spam filter is legitimate.

Another reason for the great performance of Naïve Bayesian spam filters is that most of them that are available and personally trained by individual users to better filter their emails using own set of the token. But other filters don't support this feature such as signature-based and list-

based filters are used in the server side and rule-based filters support client-side filtering but in small scale. Comparative analysis is shown in "table VI".

The learning system is slower than other approaches because it works with lots of tokens and database. It takes time to retrieve, insert and update the database each time. But other systems do a little business with the database than learning-based systems. Learning systems are reliable because it works with the content of the email rather than the sender or sending network of the message. As a result, the spammer cannot circumvent the filter by sending it from a new address or network. This increases the reliability of learning systems while rule-base and list-based are less reliable.

TABLE VI COMPARATIVE ANALYSIS OF FILTERING SYSTEM

Filter Type	Fast	Reliable	Dynamic	Application Area
Signature-based	Yes	Yes	No	Server
Rule-based	Yes	Sometimes	No	Both server and client
Our proposed filter	No	Yes	Yes	Both server and client
List based	Yes	Sometimes	No	Server

It is also worth mentioning that the Naïve Bayesian spam filter is not our new invention. We added some more techniques to enhance reliability, accuracy, and dynamicity. Fig. 3 depicts the comparison between the conventional Naïve Bayesian spam filter [14] and our proposed model. We calculate average accuracy, recall, precision, F1-measure.

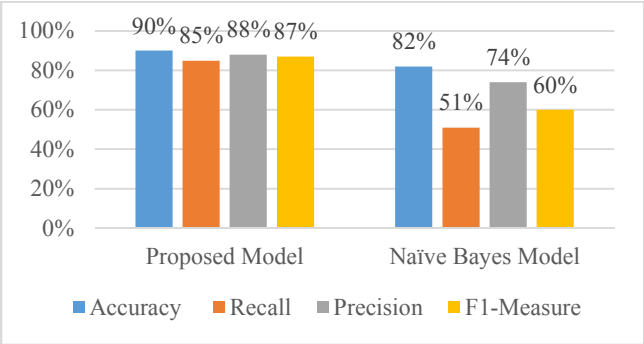


Fig.3 Comparison with the Naïve Bayes Model

The comparison clearly says that our proposed model is best suited in all the purposes.

VII. Conclusion

No spam filter can detect spam with 100% accuracy. So, the proposed spam filter is not above all weakness but we can certainly improve the performance of the filter by incorporating various features into the filter. The accuracy of the filter is shown based on some real-time data during the experiment phase of the filter, and we have gained the accuracy of spam detecting at a standard level of 88%. The accuracy can be increased by tokenizing the incoming email more efficiently and suitably building the spam and ham seeding database with proper tokens. The accuracy also depends on the order at which emails are received during the training period of the database. Lastly, to defeat spam, it is more important to analyze the content of the mail than other techniques, because the content is everything for the spammer to deliver to the people. Spammers are getting smarter day by day. But with proper analysis of content and a large amount of data training, it is possible to develop an efficient spam filter. Hence, there is a lot of scope for future improvement. Modern email system also contains HTML data, images, and other languages besides English and so on. None of these things are considered in the system. One can work with these fields for detecting spam email more effectively.

References

[1] Pantel, Patrick, and Dekang Lin. "Spamcop: A spam classification & organization program," in *Proceedings of AAAI-98 Workshop on Learning for Text Categorization*, 1998.

[2] Androutsopoulos, Ion, et al. "An evaluation of naive bayesian anti-spam filtering." arXiv preprint cs/0006013 (2000).

[3] P. Graham, "Better Bayesian Filtering", *Paulgraham.com*, 2019. [Online]. Available: http://www.paulgraham.com/better.html. [Accessed: 14- Apr- 2019].

[4] Issac, Biju, and Wendy J. Jap. "Implementing spam detection using Bayesian and Porter Stemmer keyword stripping approaches." *TENCON 2009-2009 IEEE Region 10 Conference*. IEEE, 2009.

[5] Li, Lin, and Chi Li. "Research and improvement of a spam filter based on Naive Bayes." *7th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Vol. 2*. IEEE, 2015.

[6] N. Omer Fadl Elssied and O. Ibrahim, "K-Means Clustering Scheme for Enhanced Spam Detection", *Research Journal of Applied Sciences, Engineering and Technology*, pp. 1940-1952, 2014.

[7] V. Metsis, I. Androutsopoulos and G. Paliouras, "Spam Filtering with Naive Bayes – Which Naive Bayes?", *The Third Conference on Email and Anti-Spam*, 2016.

[8] E. Tan, L. Guo, S. Chen, X. Zhang and Y. Zhao, "UNIK: unsupervised social network spam detection", *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management - CIKM '13*, 2013. Available: 10.1145/2505515.2505581

[9] Y. Ren and D. Ji, "Neural networks for deceptive opinion spam detection: An empirical study", *Information Sciences*, vol. 385-386, pp. 213-224, 2017. Available: 10.1016/j.ins.2017.01.015

[10] Boser, Bernhard E., Isabelle M. Guyon, and Vladimir N. Vapnik. "A training algorithm for optimal margin classifiers." In *Proceedings of the fifth annual workshop on Computational learning theory*., ACM, 1992.

[11] Cristianini, Nello, and John Shawe-Taylor. Cambridge University Press, 2000.

[12]Enron-Spam-dataset [Online] Available: http://nlp.cs.aueb.gr/software_and_datasets/Enron-Spam/index.html [Accessed: 14- Apr- 2019].

[13]"Apache SpamAssassin: Welcome", *Spamassassin.apache.org*, 2019. [Online]. Available: https://spamassassin.apache.org/. [Accessed: 14- Apr- 2019].

[14] N. Fitriah Rusland, N. Wahid, S. Kasim and H. Hafit, "Analysis of Naive Bayes Algorithm for Email Spam Filtering across Multiple Datasets", in *International Research and Innovation Summit (IRIS2017)*, 2017.

COMPARISON OF LINEAR DISPLACEMENT MEASUREMENTS BETWEEN A MEMS ACCELEROMETER AND HC-SR04 LOW-COST ULTRASONIC SENSOR

Azmyin Md. Kamal
Department of Mechanical and Production Engineering, Ahsanullah University of Science and Technology
Dhaka, Bangladesh
azmyin.mpe@aust.edu

Salman Habib Hemel
Department of Mechanical and Production Engineering, Ahsanullah University of Science and Technology
Dhaka, Bangladesh
salmanhabibhemeldrmc@gmail.com

Mohsan Uddin Ahmad
Department of Mechanical and Production Engineering, Ahsanullah University of Science and Technology
Dhaka, Bangladesh
mohsan004buet@gmail.com

Abstract – **In mobile robotics measuring displacement with high level of certainty is very critical to successful localization and mapping. However, sensors such as the depth cameras and LIDARS are often very expensive and difficult to obtain in Bangladesh. To circumvent this problem researchers in Bangladesh uses HC-SR04, a low-cost ultrasonic ranger finder for displacement measurement. MPU6050 a low-cost IMU contains a 3 axis accelerometer which enables it to calculate displacement by double integrating acceleration data. This paper compares the displacement measured using HC-SR04 to the displacement measured by MPU6050 in order to determine whether the displacement calculated from the accelerometer data alone is sufficiently accurate for robot localization. Single step trapezoidal integration method was used and the accelerometer was calibrated using 7000 data sample requiring 1604ms. A convolution type Moving Average Filter took acceleration data 64 times in each time instance for removing noise signals. All data were updated in 15ms intervals. A novel multiplier called Scale Factor (S.F) was introduced to compensate for the resolution issue without amplifying noise signals. Experimental result showed that displacement calculated from accelerometer data can estimate displacement in a mobile robot with reasonable accuracy, provided some assumptions are taken which are not representative of the real world scenarios.**

Keywords — Linear displacement, MPU6050, HC-SR04, accelerometers, moving average, double integration.

I. INTRODUCTION

Localization, the process of determining a robot's position in the world coordinate frame has received much attention in the last decade. A key part of this process is finding the displacement of the robot with high accuracy. This is of paramount importance as literature on the topic indicates that accurate localization leads to accurate environment mapping and an accurate map facilitates in better localization [1 2]. Robots which have been successful in localization (and consequently mapping) are often outfitted with high-performance LIDARs, research grade Inertial Measurement Units (IMU), depth perception cameras and powerful onboard computers. A couple of examples in this regard are the University of Waterloo's 2013 NASA Sample Return Robot [1] and CMU's Atlas with the Multisense SL sensor [2]. However, hardware used in such projects are often very expensive. For example, the IMU used in [1] costs around $1500 while the price of the LIDAR used was around $2000. Funds for such high-cost equipment and technical support from sensor manufacturers are very difficult to obtain by roboticists here in Bangladesh due to various socio-economic reasons.

Thus in order to implement robot localization, engineering students and scientists in Bangladesh often rely on low-cost hobby grade sensors such as the HC-SR04 Ultrasonic ranger for object detection and distance measurements. They also use Invensense MPU6050, a low-cost IMU for measuring motion and orientation in 3D space. Use of HC-SR04 and similar low cost ultrasonic sensors have been studied extensively in works such as object detection, analysis and classification [3, 4, 5], binaural ultrasonic sensor pod design for obstacle avoidance, in a quadruped robot [6], design of one-class classifiers for human detection [7], signal analysis metrics for low cost ultrasonic sensors [8] and using an array of HC-SR04s to demonstrate the usage of Kalman Filter for accurate obstacle detection in an UAV application [26]. From these studies, it can be inferred that HC-SR04 is capable of measuring distances from inanimate objects made of materials such as wood, plastic, brick wall with the high degree of accuracy provided they are placed perpendicularly to its view cone and located at a distance between 20 cm to 150 cm.

Invensense's MPU6050, a well-known commercial IMU is composed of 3 axis gyroscope and 3 axis accelerometer with an auxiliary port for connecting external magnetometer for full 3 axis orientation measurement. Madgwick el at [9] showed the capability of this low-cost IMU for accurately measuring the 3D orientation of an object in comparison with a propriety Kalman filter and ground truth data from an external optical measurement. His work showed that MPU6050 can work very well with an 8bit microcontroller which possesses very limited computing resources for robust orientation measurement. Since then the MPU6050 have seen use in numerous scientific studies such as Fabio Varesano's open source hardware IMU framework [10], smart glove to translate hand gestures into sign language [11], direction agnostic fall monitoring device for old people [12], pose determination for balancing robots [13 14] and so on.

A key benefit of having an onboard accelerometer on a mobile robot is that it is possible to determine displacement (or distance) from the accelerometer data directly using double integration. This has been thoroughly examined in [15, 16, 23]. Furthermore, from those studies, it was shown that a system which accelerated and decelerated relatively at a constant rate, its displacement can be estimated using the Trapezoidal numerical integration method whose result remained within an acceptable margin of error. Motivated by these findings, we present our study of comparing the displacement measured by HC-SR04 with the displacement calculated from accelerometer data of MPU6050. Our goal is to answer whether the data from an accelerometer alone is sufficient to determine the displacement of the robot with reasonable accuracy.

For this study we take the following assumptions:

i. The robot moves only in the positive X-axis direction of the IMU.

ii. Taking one axis data of the accelerometer is sufficient to remove the effects of normal forces exerted due to earth's gravity on the accelerometer.

iii. The distance measured by the HC-SR04 ultrasonic sensor is accurate and precise.

iv. The time interval between each reading is identical.

The rest of the paper is arranged as follows: Section 2 describes the experiment procedure, Section 3 describes our findings and Section 4 concludes the paper discussing some of the shortcomings and potential future studies.

II. METHODOLOGY

Fig. 1 shows our test robot for this experiment. It is a food serving robot built to study the interactions between the robot and the students in a dynamic cafeteria environment. At the time of writing this paper, it is a two-wheel drive mobile robot consisting of an Arduino Due embedded development board based on Atmel SAM3X8E, an ARM Cortex-M3 32bit microcontroller running at 86 MHz. It is outfitted with an expansion shield to interface a 16 by 2 Liquid Crystal display (shown in Fig. 2) and a GY-86 10DOF sensor board comprising of InvenSense MPU6050 a 6 axis inertial measurement unit (IMU). The robot is powered by a 3 cell 12.6V 36.6Wh Lipo battery. The Atmel SAM3X8E is programmed using the open source Arduino software framework [21]. The data from MPU6050 is read using I2CDevlib library written by Jeff Rowberg [19] and the ultrasonic sensor data was read using the NewPing library by Tim Eckel [20]. All accelerometer data were captured using CoolTerm serial monitor written by Roger Meier [24].

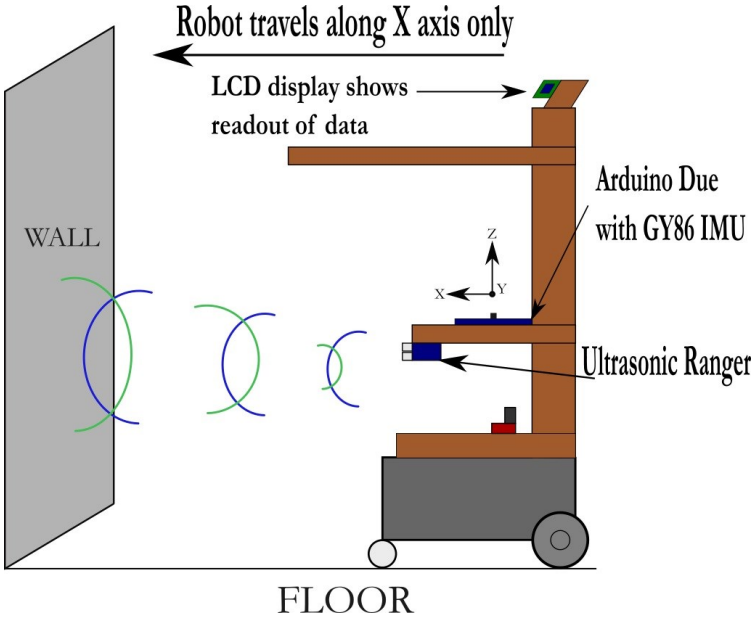


Fig. 1 Schematic diagram of the experiment process

From the datasheet of MPU6050 [22], the full-scale range of the accelerometer was set to ±2g and sensitivity scale factor was set to 16,384. This means that if the output from the accelerometer is 16,384 on a particular axis then the sensor is measuring 1g force on that axis. The accelerometer is then calibrated keeping the robot stationary and the zero reference values for all three axes were determined by averaging the values of 7000 samples. The HC-SR04 ultrasonic sensor was calibrated prior to its installation on the robot. We ascertained that our HC-SR04 unit can measure distances between 20cm to 150cm accurately with 1-1.5% of error when compared to physical measurements using a tape ruler. After calibration, the robot is placed 150mm away from the wall as shown in Fig. 1 for each run. The robot moves forward for 2s in each cycle while recording displacement measured by the ultrasonic sensor and accelerometer simultaneously. Deducting the distance measured by the ultrasonic sensor from the initial and final positions gives us the ground truth data for comparing the displacement calculated using the data obtained from the accelerometer.

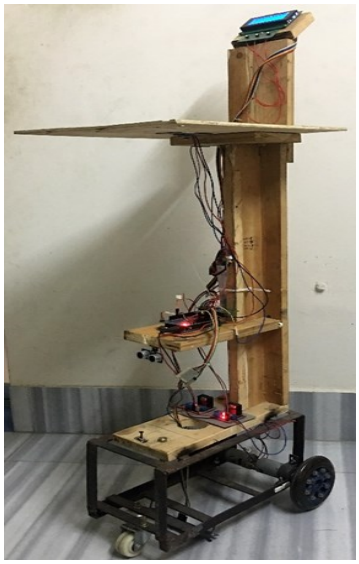


Fig. 2 Food serving robot showing all of its major components



Fig. 3 LCD showing calibration data

In each time increment, we log 64 readings from the MPU6050 and then fed the data through a Moving Average Filter, a Low Pass FIR filter programmed on basis of [17] and [18]. Our DSP filter is of convolution type, meaning each output point is produced by sampling a set of input points and then averaging the sum of those values. Mathematically this can be written as

$$y[i] = \frac{1}{M}\sum_{j=0}^{M-1} x[i+j] \tag{1}$$

The filtered value is then subtracted from the zero reference values for obtaining the body acceleration value at that instant. Even with two low pass filters implemented in software and hardware end (MPU6050 has a 250Hz low pass filter activated by default) respectively, noise signals were still present which registered erroneous displacements. To rectify this, we employed another software filter named Mechanical Noise Window (MNS) as discussed in [17]. Its low and high threshold values were determined experimentally. The acceleration values were then converted to cm/s^2 using the following equation

$$a_g = \frac{a_{LSB} * 9.81 * 100}{16834} cm/s^2 \tag{2}$$

This data is now suitable for double integration to determine the displacement of the robot in cm. We implemented the single step Trapezoidal rule for double integrating the accelerometer data using the following formulas adapted from [23].

$$\int_{t(0)}^{t(n)} a(t)\,dt \cong \sum_{i=1}^{n}\left(\frac{a(i-1)+a(i)}{2}\right)\Delta t \tag{3}$$

$$v_c(i) = S.F * v_c(i-1) + \left(\frac{a(i-1)+a(i)}{2}\right)\Delta t \tag{4}$$

$$d_c(i) = d_c(i-1) + \left(\frac{v(i-1)+v(i)}{2}\right)\Delta t \tag{5}$$

Here *dc (i)* is the instantaneous displacement and S.F is a "scale factor" which will be discussed in more details in Section III.

III. RESULTS AND DISCUSSION

To commence our study we required a suitable sample size for running the calibration routine. Table 1. shows the time required in milliseconds by the Atmel SAM3X8E for executing the calibration method for different sample sizes. It is clear from the data that once we move above 2000 sample size, the zero reference value for X-axis becomes stable. Since our experiment is designed to be implemented in short bursts, we opted to go for the 7000 sample size which requires 1604 ms to be executed.

TABLE I. EFFECT OF NUMBER OF SAMPLE POINTS ON CALIBRATION TIME

Number of samples	*Ax value (bits)*	*Time (ms)*
10	-124	2
50	-127	11
100	-99	22
150	-106	34
300	-120	68
500	-120	114
750	-114	172
1000	-117	229
1000	-119	231
2000	-123	458
3500	-123	803
5000	-122	1148
7000	-123	1604
7500	-122	1722
10,000	-122	2296

Table 2. shows the time required in milliseconds by the microcontroller to execute the moving average method in different configurations for removing noise signals from the accelerometer data. Iteration number here refers to the number of calibrated data the MCU will poll from the MPU6050 within the 15ms window to calculate the body acceleration value in LSB/g format [22]. The data showed that doubling iteration number doubles the time required to execute the Low Pass filter. However, for 256 and 512 iterations, the time required was exactly the same and this oddity held true after multiple calibration attempts.

TABLE II. Effect of the number of iterations on the loop time for executing moving average method

Iteration number	*Loop time (ms)*
8	2
16	4
32	8
64	15
128	30
256	59
512	59
1024	118

For our study, we chose 64 iteration number which is the sample size used in Freescale application note [17]. However, even with the Moving Average filter, there existed errors in accelerometer data caused by sensor noise. These errors were interpreted as a constant velocity by the microcontroller and consequently summed as a steady movement even though the robot was stationary. A second software filter dubbed the Mechanical Noise Window (MNS)

first introduced in [17] was used to cut off values within a certain band to eliminate this problem. Fig. 2 shows the raw and calibrated data from the accelerometer and Fig. 3 demonstrates the use of MNS to eliminate the erroneous displacement caused by the sensor noise shown in Fig 2. The error is linear in nature which was expected as in a standstill position, no other force was acting along the X-axis direction and on the basis of assumption 2, gravity vector which is assumed to be registered only along the Z-axis by the IMU in a stationary position was omitted from the start. On the basis of assumption 1, we only considered the absolute value of the calibrated data. As mentioned previously the motors are allowed to spin forward for only 2s. When the power is cut off the robot will decelerate but still move forward by some amount. Thus we are required to take the calibrated value because when decelerating, a negative displacement is computed which reduces the overall calculated displacement thereby giving false readings.

Table 3. shows the calculated displacement from the accelerometer data using (3 – 5). As shown in Fig. 1 the sonar, measures the robot's displacement from the wall by deducting initial and final position readings. The initial position for all runs is 150 cm. However, we noticed that the IMU data lagged behind irrespective of the number of iteration set for the moving average method. Using the CoolTerm serial monitor, we monitored the internal variables and deduced that with the default accelerometer sensitivity the IMU was not updating fast enough as its I2C bus is capped at 400Khz even though our microcontroller is capable of much faster data transfer.

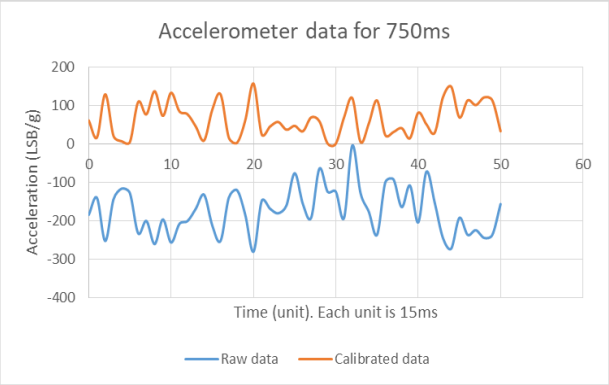


Fig. 4 Raw and Calibrated data

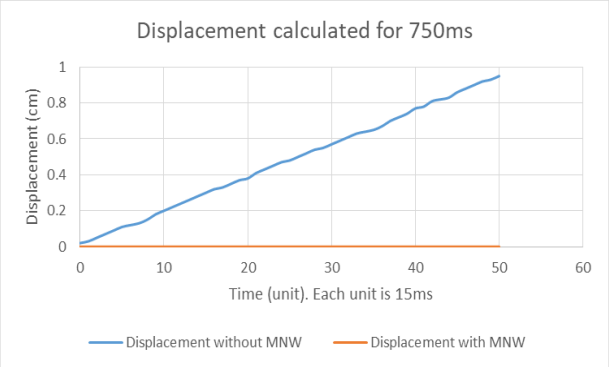


Fig. 5. Use of Mechanical Noise Window (MNW)

TABLE III. Calculated displacements with percent of error

No.	*Sonar final (cm)*	*Sonar displacement (cm)*	*Accelerometer displacement (cm)*	*Scale Factor (S.F)*	*Error (%)*
1	53	97	15	10	85
2	51	99	15	10	85
3	48	102	15	10	85
4	48	102	19	15	81
5	44	106	25	15	76
6	47	103	22	15	79
7	47	103	32	20	69
8	42	108	34	20	69
9	42	108	29	20	73
10	40	110	38	25	65
11	41	109	39	25	64
12	43	107	35	25	67
13	40	110	47	30	57
14	41	109	42	30	61
15	40	110	40	30	64
16	38	112	52	35	54
17	39	111	58	35	48
18	39	111	52	35	53
19	43	107	57	40	47
20	44	106	55	40	48
21	44	106	55	40	48
22	41	109	61	45	44
23	41	109	61	45	44
24	44	106	59	45	44
25	41	109	64	50	41
26	40	110	69	50	37
27	40	110	67	50	39
28	49	101	93	60	8
29	38	112	94	60	16
30	40	110	96	60	13
31	38	112	98	65	13
32	39	111	90	65	19
33	37	113	97	65	14
34	35	115	93	70	19
35	38	112	105	70	6
36	37	113	107	70	5
37	38	112	113	75	1
38	38	112	114	75	2
39	39	111	113	75	2
40	37	113	110	75	3

We decided not to change the sensitivity factor as most of the literature on MPU6050 used the default sensitivity scale (+/- 2g). This is due to the fact that when the sensitivity of the accelerometer is increased, it proportionally increases the effect of the noise source in the measured signal. Sources of

noise in MEMS accelerometer is explained in depth in [25]. For the reasons above, we introduced a new multiplication factor named Scale Factor (S.F). S.F is an integer used in (4) to increase the value of velocity to compensate for the sensitivity problem without amplifying signal noise.

From Table 3, we observed that for a S.F value of 70, the data between the Ultrasonic sensor and accelerometer starts to converge well. Table 4. shows the data for determining a S.F value for which the error will be minimum.

TABLE IV. Determination of S.F value

No.	*Sonar final (cm)*	*Sonar displacement (cm)*	*Displacement IMU (cm)*	*Scale Factor (S.F)*	*Error (%)*
1	30	120	128	70	7
2	25	125	104	70	17
3	24	126	113	70	10
4	22	128	117	71.5	9
5	22	128	115	71.5	10
6	21	129	112	71.5	13
7	23	127	125	72.5	2
8	24	126	131	72.5	4
9	25	125	127	75	2
10	**26**	**124**	**124**	**75**	**0**
11	**25**	**125**	**125**	**75**	**0**
12	**23**	**127**	**131**	**75**	**3**

From Table 4 it is evident that the S.F value for our experimental robot is 75, a value for which the error is consistently within 5% of error margin. Our experiment indicates that low-cost IMU can be used for measuring linear displacement with reasonable accuracy but requires a controlled environment. Thus we conclude that though IMUs are capable of displacement measurements, they should not be used as the only sensing device for this purpose and should be coupled with other sensors such as HCSR04, LIDARs and so on. We claim this notion from the fact that at the onset of this study we had taken four assumptions. These assumptions are not representative of most real-world applications. A byproduct of our study is that both HC-SR04 and MPU6050 have shown remarkable performance with the Arduino platform which indicates their potential for use in complex projects by roboticists here in Bangladesh.

IV. Conclusion

This paper presents a comparative study of the performance and reliability of using low-cost commercial MEMS accelerometer in measuring linear displacement. It was compared to the displacement measured by HC-SR04, a low-cost hobby grade ultrasonic ranger widely available in Bangladesh. Even though a good agreement between the measured values were seen, assumptions were imposed which are not representative of real-world scenarios. Most notability assumption about the accuracy of HC-SR04 is not always fixed due to the fact that HC-SR04's view-cone changes with respect to the sensor's distance from the object. Furthermore, the assumption about each loop running at exactly 15ms is not true. In our initial trials, we had observed that as the ultrasonic ranger came closer to the wall, it changed the overall loop time. Tim Eckel's library resolved this issue by utilizing an independent *timeout* argument which allowed receiving data for a fixed period of microseconds before returning 0. We suggest that linear displacement computed from accelerometer data should be used in conjunction with other sensors. In this regard, the use of the weighted average method is recommended for maximum accuracy as in perspective of a microcontroller application, this method is easy to code and requires low computing resources. Our future work will include the use of all three axes data, advanced numerical methods and will use the timer modules to independently count time increments rather than depending on Arduino's delay method which drifts with time.

V. Acknowledgment

The authors acknowledge the support of the Department of Mechanical and Production Engineering of Ahsanullah University of Science and Technology for providing workspace in their Simulation Lab. We are also thankful to the open source Arduino community for providing the necessary codes and helpful documentation on interfacing HC-SR04 and MPU6050 with Arduino Due.

VI. References

[1] A. Das, M. Diu, N. Mathew, C. Scharfenberger, J. Servos, A. Wong, *et al.*, "Mapping, Planning, and Sample Detection Strategies for Autonomous Exploration," *Journal of Field Robotics,* vol. 31, pp. 75-106, 2014.

[2] S. Kuindersma, R. Deits, M. Fallon, A. Valenzuela, H. Dai, F. Permenter, et al., "Optimization-Based Locomotion Planning, Estimation, and Control Design for The Atlas Humanoid Robot," *Autonomous Robots*, vol. 40, pp. 429-455, 2015.

[3] C. N. Yalung and C. S. Mathew Adolfo, "Analysis of Obstacle Detection Using Ultrasonic Sensor," *International Research Journal and Engineering Technology*, pp. 1015–1019, 2017.

[4] R. A. Dimitrov and D. Minchev, "Ultrasonic Sensor Explorer," in *19th Int. Symp. on Electrical Apparatus and Tech., SIELA 2016*, Borgus, Bulgaria, Aug. 2016, pp. 1-5.

[5] D. Vakula and Y. K. Kolli, "Low Cost Smart Parking System for Smart Cities," in *Proc. Int. Conf. Intell. Sustain. Syst.,* Dec. 2017, pp. 280–284.

[6] X. T. Le, S. B. Oh, W. S. Lee, C. S. No, and S. H. Han, "Design of Intelligent Mobile Robot System Based on Ultrasonic Sensors," in *Int. Conf. Control. Autom. Syst.,* Dec. 2007, pp. 72–76.

[7] Sonia, A. M. Tripathi, R. D. Baruah and S. B. Nair, "Ultrasonic Sensor-Based Human Detector Using One-Class Classifiers," in *IEEE Int. Conf. on Evolving and Adaptive Intell. Syst.,* Dec. 2015, pp. 1-6.

[8] G. Gibbs, H. Jia, and I. Madani, "Obstacle Detection with Ultrasonic Sensors and Signal Analysis Metrics," in *Transp. Res. Procedia*, vol. 28, pp. 173–182, 2017.

[9] Madgwick, S.O.H.; Harrison, A.J.L.; Vaidyanathan, A. "Estimation of IMU And MARG Orientation Using A Gradient Descent Algorithm," in *Proc. of IEEE Int. Conf. of Rehabilitation Robotics*, Zurich, Switzerland, 29 Jun.–1 Jul., 2011, pp. 1–7.

[10] F. Varesano, "FreeIMU: An Open Hardware Framework for Orientation and Motion Sensing," 2013, Accessed on: Dec. 2018 [Online] Available: arXiv:1303.4949v1 [cs.ET].

[11] A. Yudhana, J. Rahmawan, and C. U. P. Negara, "Flex Sensors and MPU6050 Sensors Responses on Smart Glove for Sign Language Translation," in *IOP Conf. Ser. Mater. Sci. Eng.* 2018, vol. 403, no. 1, doi:10.1088/1757-899X/403/1/012032.

[12] J. Luo, B. Zhong, and D. Lv, “Fall Monitoring Device for Old People Based on Tri- Axial Accelerometer,” Int. J. Adv. Comput. Sci. Appl., vol. 6, no. 5, pp. 34–39, 2015.

[13] M V Momot, E V Politsinskaia, A V Sushko and A Semerenko, “Justification of Filter Selection for Robot Balancing in Conditions of Limited Computational Resources,” *IOP Conf. Ser. Mater. Sci. Eng,,* 2016, Accessed on: Dec. 2018 [Online] Available: https://doi.org/10.1088/1757-899X/142/1/012108

[14] A.K.M Asfaque Shakil Shaon, S. Bhowmik, B. K. Bhawmick, P. Das and N.K. Das “Design And Implementation Of A Self-Balancing Robot” *in Proc. Int. Conf. on Mechanical Eng. And Renewable Energy*, Dec. 2017, Accessed on: Dec. 2018 [Online]

[15] T.J. Cheng et al., “Characterisation of Rollator Use Using Inertial Sensors,” *Healthc. Technol. Lett.*, vol. 3, no. 4, pp. 303–309, 2016.

[16] J. Hrisko, *Introduction To Capacitive MEMS Accelerometers and A Case Study On An Elevator,* Oct. 2017, Accessed on: Dec. 2018 [Online] Available: https://engineersportal.com/blog/2017/9/25/accelerometer-on-an-elevator

[17] K. Seifert and O. Camacho, *Implementing Positioning Algorithms Using Accelerometers,* 2007, Freescale Semiconductor, Accessed on: Dec. 2018 [Online] Available:https://www.nxp.com/docs/en/application-note/AN3397.pdf [Accessed: December 2018]

[18] S.W.Smith, “Moving Average Filters,” in *The Scientist and Engineer’s Guide to Digital Signal Processing*, San Diego, CA, California Technical Publishing, 1999. Accessed on: Dec. 2018 [Online] Available: http://www.dspguide.com/pdfbook.htm

[19] Jeff Rowberg, *I2CDevlib,* Sept. 2018, [Online] Available:https://github.com/jrowberg/i2cdevlib/tree/master/Arduino/MPU6050

[20] Tim Eckel, *Arduino New Ping,* Nov. 2018, [Online] Available:https://bitbucket.org/teckel12/arduino-new ping/wiki/Home#!history

[21] Arduino, *Arduino IDE 1.8.4,* Dec. 2018, Accessed on: Dec. 2018 [Online] Available: https://github.com/arduino/Arduino/

[22] InvenSense Inc. Technical Staff, “MPU-6000 and MPU-6050 Product Specification Revision 3.4”, p13, InvenSense Inc., Aug. 2013

[23] M. Arraigada and M. Partl, “Calculation of displacements of measured accelerations, analysis of two accelerometers and application in road engineering” in *Proc. 6th Swiss Transport Research Conf.,* Ascona, Switzerland, Mar. 2006*,* pp. 30-60

[24] Roger Meier, *CoolTerm*, Dec. 2018 [Online] Available: http://freeware.the-meiers.org/

[25] L. B. Pupo, “Characterization of errors and noises in MEMS inertial sensors using Allan variance method,” M.S thesis, UPC, Barcelona, Spain, 2016. [Online] Available: http://hdl.handle.net/2117/103849

[26] J.S.G.Guerrero, A.F.C.Gonzalez, J.I.H.Vega and L.A.N.Tovar, “Instrumentation of an Array of Ultrasonic Sensors and Data Processing for Unmanned Aerial Vehicle (UAV) for Teaching the Application of the Kalman Filter,” *Procedia Computer Science*, vol. 75, pp. 375-380, 2015.

Students Evaluation in AGILE after Completing Software Engineering Course

Sabbir M. Saleh [1], Supta Richard Philip [2]
Department of Computer Science and Engineering
City University
Dhaka, Bangladesh
[1]saleh.sabbir.aiub@gmail.com
[2]supta.philip@gmail.com

Abstract— **In the arena of Software Development, the software engineering philosophy Agile is becoming more popular juxtaposed to the long-established methods. Software organizations have a prerequisite from graduates with cutting-edge knowledge of software engineering methodologies. Due to Agile's popularity universities and institutions have already added agile in the curriculum of software engineering. They have started to teach the students Agile theoretically, practically in labs and also have some seminars where empirical knowledge is shared with students. This paper aims to acquit the determination of Agile to the perceptual experience of Software Engineering undergraduate students, after completing their projects in the lab by following Agile principles, where they were trying to adopt Agile.**

Keywords— Software Engineering Education, Agile Software Development, Software Project Management.

I. INTRODUCTION (*HEADING 1*)

As Agile frameworks are getting to be standard in the Software organizations [1] there has been an expansion in the interest for them be schooled in the classroom. Following Scrum, one of the most popular Agile frameworks has heightened in the last few years through the demonstrations of surveys. Considering "2015 state of Scrum Report" Scrum has been followed in organization for more than 50% of the time, and that 42% of responders follow Scrum entirely [2]. Regarding the VersionOne [3] State of Agile Survey, around 88% of responders are applying agile in the workspace. Apparently, it is hard to adopt the Agile principles as well as gain knowledge for most of the organizations. Agile is a philosophy to develop software with project management procedures, such as Scrum, eXtreme Programming (XP), Lean Software Development (LSD), Kanban, Adaptive Software Development (ASD), Feature Driven Development (FDD), Dynamic System Development Method (DSDM), Agile Modeling (AM), Crystal, Agile Unified Process (AUP) [4].

Agile Squads perform their jobs differently compared to other software development teams, where customers and team members have good continuous communication between them. They are delivering working software frequently. These types of duties need some essential accomplishments like communication skills and faster implementation potentiality. Advance, because agile squads are self-organized teams there is a demand to consume self-denial, a power to assume the obligation, to consecrate, plus take responsibility and tenure. So, this is an ecumenical swerve that pupils need to evolve not only fantabulous technological expertise but also capability of solving a problem, innovative skills of solutions, scheme skills, kinship attainments, to make a quick decision with ability in a fast surrounding [5, 6, 7].

A gap has been created between the knowledge and accomplishments instructed in universities and the one's prerequisite in software organizations, after the popularity of Agile use [8]. So that universities are now adding agile in their software engineering course curriculum, for a good outcome.

This paper intends to survey the relationship of tutoring Agile frameworks and the perceptual experience also may have called as an evaluation of undergraduate students of software engineering which are related to the Success Factors in Agile Software Projects (from Tsun Chow – 2007). We have to keep in mind that, this evaluation was done by those students who didn't know anything about Agile before studying the software engineering course. This work reports the leads of a first effectuation of agile learning in the circumstance of undergraduate Computer Science and Engineering (CSE) Education Department.

The theory of this research is finding students' team's performance to submit their software engineering project between a deadline fixed by their tutors. This paper can be a suggestion for other software project management education area.

The rest of the paper is written into following sections: Related Works in Section II, Description of class and lab environment in Section III, Research Methodology and Results Discussion are respectively in Section IV and V, Limitations, and Future Works in Section VI, and Conclusion in Section VII.

II. RELATED WORK

Although the agile framework has been nearby for more than a tenner, tutoring agile philosophy has just attracted substantial tending in research and educational fields in the endure years [10]. Many research workers have been admitted the demand to instruct the agile process in software engineering curricula [11, 12] as an intends to construct ethical and social skills in accession to technological [12]. There is arising consciousness that conventional theoretical speeches only unable to give students learn agile frameworks, instead students necessitate to essentially implement them for upskilling themselves and enriching the acquiring procedure [9, 13].

Be that as it may, revealing students to agile process and patterns in the context of university attaches to a set of disputes [10] counting, student aims, and motivation can be diverged [9, 14, 15], bounded availableness and backup from Product Owner/Client [14, 15], lack of counseling from experienced scrum master (Scrum) or a coach (XP) [15, 16, 18]. Another dispute reported let concisely length of classes with one-half of the course nonce consecrated to teaching constructs cutting time to projects work, students involved to

register multiple subjects in accession to their professional and personal dedications [14, 16]. As a result, now and then students seemingly benefit experience in exercising the agile frameworks, but they cannot be proficient to apply them acceptably [15, 16, 18].

More or less of the universities comprise agile patterns into the course through a dissimilar scaled (large, medium and small) sequence of projects [19] while others have taken the method of presenting agile philosophy in one semester followed by a final project in the following [17]. More or less allowed students to study agile approaches by functioning on software developments under the supervising of lecturers [21]. Not very many worked together with industry to increase the reasonable experience of teaming up with genuine clients through ventures [18]. Understudies taking a shot at existing industry frameworks is uncommon [20].

The curriculum of IS 2010 model indications "agile methods" in the subjects list for the IS 2010.6 Systems Analysis and Design class. Project management and Programming are two other subjects that may include agile approaches [22]. Schwalbe has combined agile into the textbook of project management [23, 24], though, the Project Management Institute (PMI) has approved agile notions in its outline, with proposing the Agile Certified Practitioner (PMI-ACP®) certificate [25].

Mahnic and Viljan Mahnic planned a software engineering course which is project-oriented, concentrated completely on Scrum by executing four sprints (three practical and one theoretical) in a semi--actual project. This type, of course, had prerequisite software engineering subjects where all the compulsory technical notions for basic software development were addressed consuming an 'old' software paradigm [26].

The selection of Agile philosophies has been demonstrated to encourage the administration of the development procedure, to lessen the additional time, and to decidedly effect on clients' satisfaction [26].

Structure upon this arising form of understanding, our paper exhibits the evaluation by students when applying agile on quasi-real projects, functioning in teamwork within university contexts and specific clients' requirements and description; and the adaptions made by student teams to agile patterns to work across the limitations. Moreover, perceived effectivity of these adaptions is talked about conducting to an understanding of the affect they made on students' learning based on success factors.

III. Our Class and Lab Environment

Software Engineering at our Computer Science and Engineering under the Faculty of Science and Engineering is a three semesters succession for specializations centering System Analysis & Design plus Software Engineering conceptions where the curriculum follows the textbook of Roger S. Pressman and Ian Somerville and the third one is Software Quality Assurance. Semesters have a closed theoretical class and sessional which copes with respectively 160 and 160 minutes hebdomadally. The requirement for the students are to attend classes and in labs with their required works. Tutors help students with troubles which are related to their assigned projects during lab as needed and discussed on theoretical software engineering in classes with historical data analysis and proper examples for each section of software engineering. Sizes of the course section are comparatively little, usually amongst 25 or 40 pupils.

Students were permitted to custom their very own teams of somewhere in the range of 3 and 4 each. The team measure considered randomness, more blend, and disallowed groups from complimenting excessively little because of drops or unlucky deficiencies. Companions would, in general, remain unitedly as did nonages. Certain students or the experts likewise would in general structure teams as did the less capable to complete a project. We experienced issues with this condition, such as teams with uneven aptitude levels. So after that, we have reformed some team members with adjusted abilities, and rouse every one of the students that all great in one part can't be a decent group.

From that point onward, each team has been appointed with software projects. The programming language for building up their software was freely picked by them. Each group was autonomous and had distinctive activities, yet being in a similar room, thoughts and ideas shared on a regular basis between teams. In the Fall 2018, the tasks were the genuine and real problem that the teacher of the class was keen on utilizing as a client. For practicing Agile principles, after each feedback teachers were changing the requirements.

Despite the fact that students at this dimension were not prepared to give real-time market software for genuine clients yet, we trusted they would profit by delivering something they knew was advantageous for a real "client," regardless of whether it was their educator.

This methodology additionally situated students for an imperative industriousness concern, the software developing that fulfilled clients as opposed to software that basically satisfied composed specification and features. The length of the task enabled the students to encounter three sprints and gave an adequate time span to take into consideration full of feeling development in group abilities.

Below are the assigned projects of Fall 2018:

- Bangladesh Travel Guide with interactive MAP feature – php Laravel 5.7
- Cha Wala (an android app for the Tea Stallers) - Java
- Boi-Mela (an online book selling market place) – php Laravel 5.7
- Nimontron (an event management system) – php Laravel 5.7
- Attendance System with Image Processing – Python
- Library Management System – Java
- Mini Militia (a game) – C# with Unity Game Engine

IV. Research Methodolgy

There were 24 students divided into 7 software projects following Agile for software engineering lab project. An online survey was created and sent to all students after completing their projects. The students' aspects and experience of abiding by Agile were accumulated through the contrived questionnaire based on Success Factors in Agile Software Projects. The inquiries were acquired and chose from the success factors [27].

The Organization factor dimension was excepted since the students are not expert in the real workspace. Consequently, in the questionnaire, there were questions in 4 dimensions namely People, Technical, Process, and Project and the questionnaire can be found here:

https://forms.gle/BAWKXquD7KB2V6ib9

Several success factors in a particular dimension are not incorporated due to their young programmers standing too narrow and cumbersome for the respondents. In order to assess the observations of respondents practicing Agile frameworks, closely answered inquiries with a Likert scale of 1 (strongly disagree) to 5 (strongly agree) were included.

Reason for choosing Agile: There are so many methodologies to develop software, and in a limited time of classroom projects are so small rather than developing in an organization. So, waterfall the traditional methodology of developing software can be used here but it's not followed by the students for their learning. Organizations are now following Agile frameworks to satisfy their customers with their requirements which are changing frequently after a sprint during the feedback session. Agile is followed here mainly to prepare the mind of students with its rules. Respective faculties have changed the requirement frequently as like customers. Students were bound to keep an update through using OrangeScrum, the open source project management, and collaboration tool helps to manage projects, team, tasks at one place. Students were given an update in their team's account frequently.

V. Results and Discussions

The statistical procedure was used to break down the data of 24 respondents, as well as qualitative data analysis for the replies through the open questions in the questionnaire. Considering the result of success factors, the majority of the students accounted for positive responses about following Agile in their semester project as demonstrated in TABLE I.

TABLE I. Survey Responses

Dimensions	Sl.	QUESTIONS	Series 1					Total
			Strongly Agree (%)	Agree (%)	Neutral (%)	Disagree (%)	Strongly Disagree (%)	
People	Q1	AGILE is a useful framework to develop a good kinship between practitioners and clients	66.67	25.00	0.00	8.33	0.00	24
	Q2	AGILE is a philosophy to motivate practitioners	25.00	70.83	0.00	4.17	0.00	24
	Q3	AGILE philosophies assist Knowledge Direction	29.17	41.67	25.00	4.17	0.00	24
Technical	Q1	AGILE uses a programming criterion to accomplish ends	29.17	58.33	12.50	0.00	0.00	24
	Q2	AGILE is capable to deliver working software frequently	45.83	41.67	4.17	8.33	0.00	24
	Q3	Students are willing to prioritize task befittingly after following AGILE	12.50	66.67	20.83	0.00	0.00	24
	Q4	AGILE helps the pupil to produce a suitable testing	29.17	58.33	0.00	12.50	0.00	24
	Q5	AGILE can ascertain a flaw in a team which assists the practitioners to explore a suitable resource like an experienced and skilled person	45.83	41.67	4.17	4.17	4.17	24
Process	Q1	AGILE assists the practitioners to infer the prerequisites perfectly	41.67	58.33	0.00	0.00	0.00	24
	Q2	Practitioners can use AGILE frameworks to track and monitor their works until the deadline	37.50	45.83	16.67	0.00	0.00	24
	Q3	Students are able to evaluate and predict the progress of the project submit within the deadline	29.17	50.00	20.83	0.00	0.00	24
	Q4	Students achieve more skills in communication by planning the project development and solve each problem when the requirements change frequently by Course Teacher	45.83	50.00	4.17	0.00	0.00	24
Project	Q1	AGILE philosophy is suited in little teams	25.00	50.00	16.67	8.33	0.00	24
	Q2	AGILE philosophy is suited in a project which is dynamic compared to a project which is static	41.67	45.83	8.33	4.17	0.00	24
	Q3	AGILE philosophy is appropriate for the projects with no several autonomous squads	8.33	50.00	20.83	20.83	0.00	24

At undifferentiated rate from the over factor, the following assuming purpose of Agile is characterizing the shortcoming viewpoint of a team. Usually to demonstrate that the speed of each sprint discharge undertaking relies upon individuals or the team. Without an exact leader, Agile team is performing [28], Scrum master defined as a part of a team to supervise, actuate and tracking another member when the team confronts a difficulty [29]. Any of these individuals in a group can be a Scrum Master to share an alternate quality and shortcoming perspective in a group. The scrum master can break down team issues and gather the frail focuses toward the finish of each sprint discharge or amid the daily scrum

meeting. The shortcoming focuses can be explained by sending the group to a fitting instructional class. In any case, there are some minority viewpoints about Technical components. As per the aftereffect of an assessment, it was demonstrated that the coding style is not really utilized by the members as unpracticed designers. The ability of coding relies upon the involvement of programming advancement. While the team was utilizing Agile for a software project, it is hard to keep the standard of each bit of assignments with individuals ailing in developing the experience of the software. This shortcoming side can influence task delays.

We have prepared graphs based on the percentages of the answers of students where each one denoted as Q1, Q2 series with the name of the dimensions and the graphical measurements were denoted as Series 1 which is mentioned on the table too. These graphs were prepared for the help of the readers and future researchers.

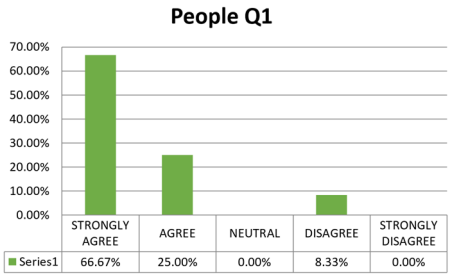


Fig. 1. Graphical view of People Q1

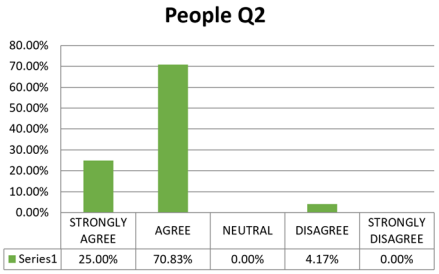


Fig. 2. Graphical view of People Q2

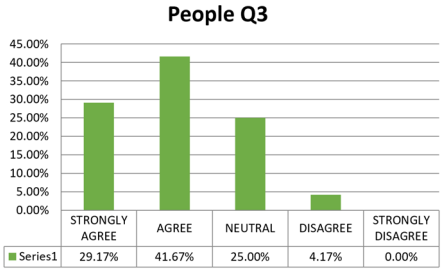


Fig. 3. Graphical view of People Q3

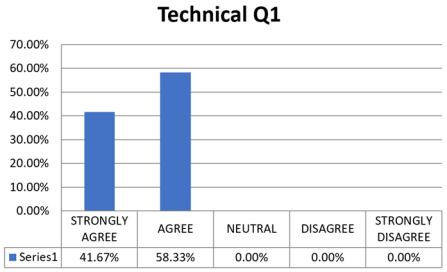


Fig. 4. Graphical view of Technical Q1

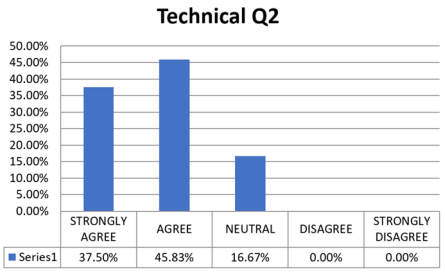


Fig. 5. Graphical view of Technical Q2

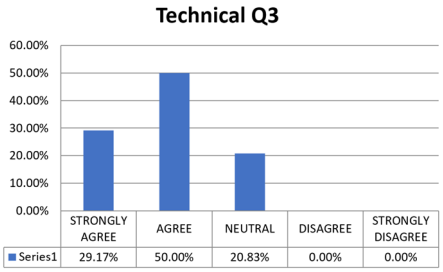


Fig. 6. Graphical view of Technical Q3

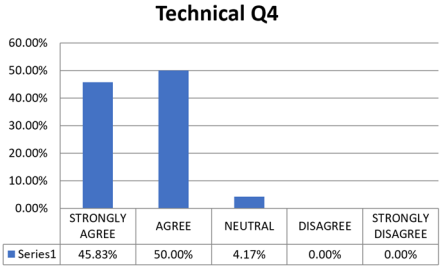


Fig. 7. Graphical view of Technical Q4

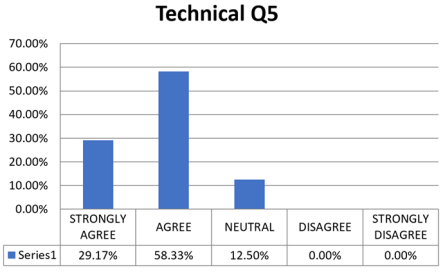


Fig. 8. Graphical view of Technical Q5

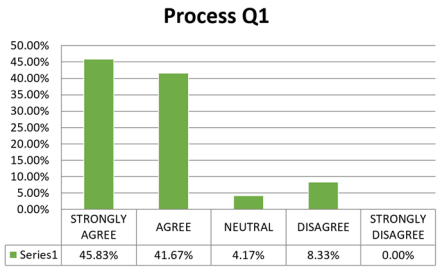


Fig. 9. Graphical view of Process Q1

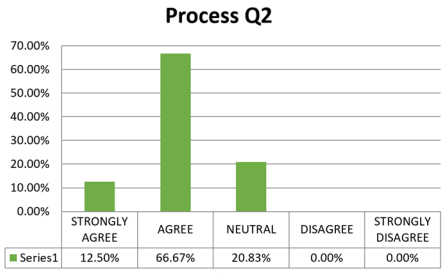


Fig. 10. Graphical view of Process Q2

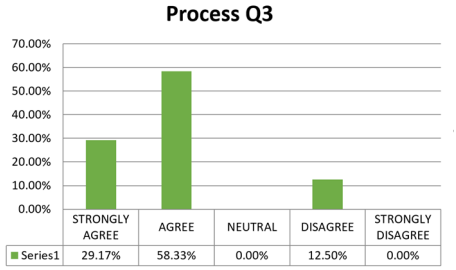


Fig. 11. Graphical view of Process Q3

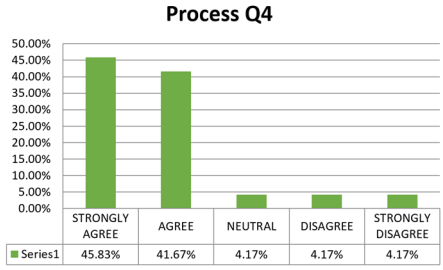


Fig. 12. Graphical view of Process Q4

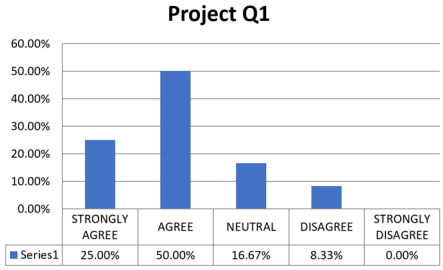


Fig. 13. Graphical view of Project Q1

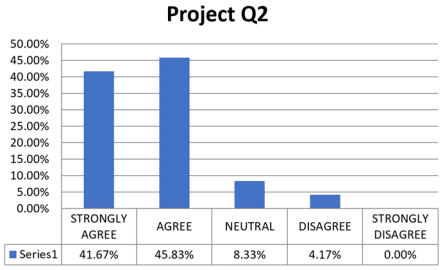


Fig. 14. Graphical view of Project Q2

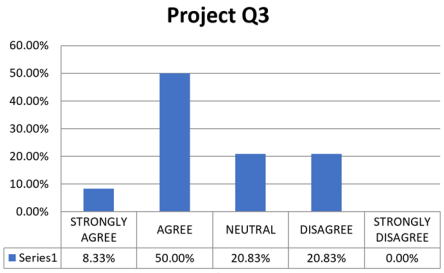


Fig. 15. Graphical view of Project Q3

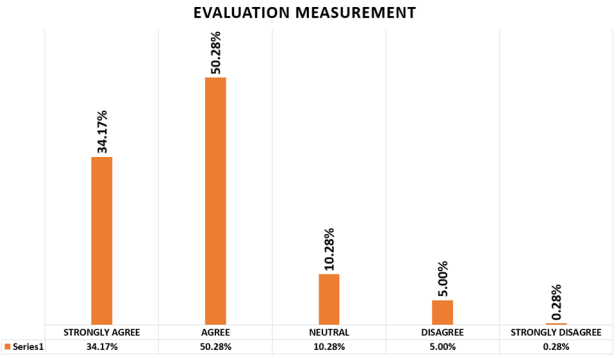


Fig. 16. Graphical view of the Evaluation Measurement of Students

VI. Limitations and Future Works

The authors think that the limitations of this paper and works are

- The limited size of responders,
- Bounded numbers of Success Factors in Agile Software Projects
- Not following a remarkable Agile Framework in software engineering Lab to coach students to be a companion with at least one of the Agile frameworks rather follow the Agile philosophy, that has been accepted for software developing project management in the software organizations.

The more works and researches are required to deal more questions, features and students' observations and evaluations of teaching the Agile framework at the students of undergraduate level in diverse sections and concepts.

There are also have some success factors and practices of the practitioners who are following and maintaining Agile frameworks named Scrum, Kanban, XP [30] which should be discussed in the classrooms and should be followed by the

students' in the lab through completing a project, basically it can be possible in the Final Year project under an active supervision of faculties. After the successful completion of those projects, we will evaluate these again with the same students those who were responded in this survey.

VII. CONCLUSIONS

The master influence of this learning is to review the effect of the success factors of projects using Agile to the insights and knowledge of the students. The intended questionnaire was sent out through google doc for collecting data from 24 undergraduate computer science and engineering pupils. The initial consequences showed that most related dimensions of success factors of Agile espousal were People, Technical, Process and Project respectively. The supreme selected success factors for each dimension is the convincing connection between the practitioners and end-users, prioritizing projects and enforcing the jobs first, supplying thoroughly and essential coaching, participating client engagement, and a few amounts of practitioners in a team for each project. But then, the to the lowest degree success factors to each one is cognition direction, unequivocal writing in code standard, in the nick of time for case-by-case work deliverance, active and accelerate the schedule of task sets. This investigation shows that the People dimension of following Agile frameworks was most bucked up in the education of software engineering at undergrad phase.

REFERENCES

[1] S. Nerur and V. Balijepally, "Theoretical reflections on agile development methodologies," Commun. ACM, vol. 50, pp. 79–83, Mar. 2007

[2] Scrum Alliance, 2015. *The 2015 State of Scrum Report*. https://www.scrumalliance.org/scrum/media/scrumalliancemedia/files%20and%20pdfs/state%20of%20scrum/scrum-alliance-state-of-scrum-2015.pdf

[3] VersionOne Inc. (2014). 8th Annual State of Agile Survey, Retrieved March 26, 2015 from http://www.versionone.com/pdf/2013-stateof-agile-survey.pdf.

[4] H. K. Flora, X. Wang, and S. Chande, "Adopting an Agile Approach for the Development of Mobile Applications," International Journal of Computer Applications, vol. 94, pp. 43-50, 2014.

[5] HAZZAN, O., AND DUBINSKY, Y. Why Software Engineering Programs Should Teach Agile Software Development, Software Engineering Notes 32, 2: 1–3.

[6] ABELSON H. AND GREENSPUN, P. 2001. Teaching Software Engineering. In 10th International World Wide Web Conference, Hong Kong.

[7] VINCENTIS, S. De. 2007. Agile education: Student-driven knowledge production. In ACEL/ASCD conference, New Imagery for Schools and Schooling Sydney.

[8] Fitsilis, Panos and Alex Lekatos. "Teaching software project management using agile paradigm." PCI (2017). DOI: 10.1145/3139367.3139413.

[9] Goldman, A., et al., Being extreme in the classroom: Experiences teaching XP. Journal of the Brazilian Computer Society, 2004. 10(2): p. 5-21.

[10] Kropp, M. and A. Meier, Teaching Agile Software Development at University Level: Values. Management, and Craftsmanship, CSEE&T, 2013.

[11] Shukla, A. and L. Williams. Adapting extreme programming for a core software engineering course. in Software Engineering Education and Training, 2002. (CSEE&T 2002). Proceedings. 15th Conference on. 2002. IEEE.

[12] Hazzan, O. and Y. Dubinsky, Why software engineering programs should teach agile software development. ACM SIGSOFT Software Engineering Notes, 2007. 32(2): p. 1-3.

[13] Bruegge, B., M. Reiss, and J. Schiller. Agile principles in academic education: A case study. in Information Technology: New Generations, 2009. ITNG'09. Sixth International Conference on. 2009. IEEE.

[14] Schneider, J.-G. and L. Johnston. eXtreme Programming at universities: an educational perspective. in Proceedings of the 25th international conference on Software engineering. 2003. IEEE Computer Society.

[15] Bunse, C., R.L. Feldmann, and J. Dörr. Agile methods in software engineering education. in International Conference on Extreme Programming and Agile Processes in Software Engineering. 2004. Springer.

[16] Wainer, M. Adaptations for teaching software development with extreme programming: An experience report. in XP/Agile Universe. 2003. Springer.

[17] Muller, M.M. and W.F. Tichy. Case study: extreme programming in a university environment. in Software Engineering, 2001. ICSE 2001. Proceedings of the 23rd International Conference on. 2001. IEEE.

[18] Schneider, J.-G. and R. Vasa. Agile practices in software development-experiences from student projects. in Software Engineering Conference, 2006. Australian. 2006. IEEE.

[19] Fenwick, J.B. Adapting xp to an academic environment by phasing-in practices. in XP/Agile Universe. 2003. Springer.

[20] Rico, D.F. and H.H. Sayani. Use of agile methods in software engineering education. in Agile Conference, 2009. AGILE'09. 2009. IEEE.

[21] Scharf, A. and A. Koch. Scrum in a software engineering course: An in-depth praxis report. in Software Engineering Education and Training (CSEE&T), 2013 IEEE 26th Conference on. 2013. IEEE.

[22] Topi, H., Valacich, J., Wright, R., Kaiser, K., Nunamaker, Jr., J., Sipior, J., and de Vreede, G. (2010). IS 2010: Curriculum Guidelines for Undergraduate Degree Programs in Information Systems. Communications of the Association for Information Systems, 26(18).

[23] Schwalbe, K. (2012). Managing a Project Using an Agile Approach and the PMBOK® Guide, In The Proceedings of the Information Systems Education Conference 2012, v 29 (New Orleans): §1985. ISSN: 2167-1435.

[24] Schwalbe, K. (2014). Information Technology Project Management, Seventh Edition, Cengage Learning, Boston.

[25] Project Management Institute (2011). PMI Agile Certified Practitioner Examination Content Outline, Retrieved March 3, 2015 from http://www.pmi.org/Certification/~/media/Files/PDF/Agile/PMI_Agile_Certification_Conte nt_Outline.ashx.

[26] Viljan Mahnic. A capstone course on agile software development using scrum. IEEE Transactions on Education, 55(1):99–106, 2012.

[27] Chow, T. & Cao, D. (2008). A survey study of critical success factors in Agile software projects. The Journal of Systems and Software, 81(6), 961–971. DOI:10.1016/j.jss.2007.08.020.

[28] Beck, K., et al. (2001). *The Agile manifesto*.

[29] The Scrumban [R] Evolution: Getting the Most Out of Agile, Scrum, and Lean Kanban 1st Addison-Wesley Professional ©2015.

[30] Saleh, Sabbir M., Syed Maruful Huq, and M. Ashikur Rahman. "Comparative Study within Scrum, Kanban, XP Focused on Their Practices." In 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1-6. IEEE, 2019.

Improving Detection Accuracy for Imbalanced Network Intrusion Classification using Cluster-based Under-sampling with Random Forests

Md. Ochiuddin Miah, Sakib Shahriar Khan, Swakkhar Shatabda, and Dewan Md. Farid*
Department of Computer Science & Engineering, United International University
United City, Madani Avenue, Badda, Dhaka 1212, Bangladesh
Email: mmiah131145@bscse.uiu.ac.bd, skbshariar@gmail.com, swakkhar@cse.uiu.ac.bd, dewanfarid@cse.uiu.ac.bd

***Abstract*—Network intrusion classification in the imbalanced big data environment becomes a significant and important issue in information and communications technology (ICT) in this digital era. Presently, intrusion detection systems (IDSs) are commonly using tool to detect and prevent internal and external network attacks/ intrusions. IDSs are majorly bifurcated into host-based and network-based systems, and use pattern-matching techniques to detect intrusions that known as misuse-based intrusion detection system. Machine learning (ML) and data mining (DM) algorithms are widely using for classifying intrusions in IDS over the last few decades. One of the major challenges for building IDS employing machine learning and data mining algorithms is to improve the intrusion classification accuracy and also reducing the false-positive rate. In this paper, we have introduced a new method for improving detection rate to classify minority-class network attacks/ intrusions using cluster-based under-sampling with Random Forest classifier. The proposed method is a multi-layer classification approach, which can process the highly imbalanced big data to correctly identify the minority/ rare class-intrusions. Initially, the proposed method classify a data point/ incoming data is attack/ intrusion or not (like normal behaviour), if it's an attack then the proposed method try to classify attack type and later sub-attack type. We have used cluster-based under-sampling technique to deal with class-imbalanced problem and popular ensemble classifier Random Forest for addressing overfitting problem. We have used KDD99 intrusion detection benchmark dataset for experimental analysis and tested the performance of proposed method with existing machine learning algorithms like: Artificial Neural Network (ANN), naïve Bayes (NB) classifier, Random Forest, and Bagging techniques.**

***Keywords*—Data Sampling; Ensemble Learning; Intrusion Classification; Imbalanced Big Data; Random Forest;**

I. Introduction

In recent times with the encroachment of information and communication technologies, computer networks are facing threats like worms, viruses, adware, spyware, rootkits, trojan horses etc. [1]. These sorts of threats/ intrusions have turn into a major apprehension and need to be perceived before destructing the computer networks. These intrusions are affecting the productivity of computer networks, organizational electronic assets and other resources. Intrusion detection system is a network security technology to identify malicious activities/ intrusions from host or network by inspecting networks and systems [2]. Network-based IDS (NIDS) and host-based IDS (HIDS) are frequently used to identify internal or external malicious activities/ intrusions. NIDS scrutinizes and monitors incoming network traffic to detect network intrusions [3]. HIDS is devoted to a singular host or computer device to detect intrusions by scrutinizing inbound and outbound packets [4]. In detection approach, IDSs are categorized as misuse-based or anomaly-based IDS [5]. Misuse-based IDS is also acknowledged as pattern-based or signature-based IDS [6]. It is incapable of detect anonymous intrusions and detection rate is comparatively low because attackers endlessly try to attack with dissimilar signatures [7]. Otherwise, anomaly-based IDS is extensively used to detect known or unknown network intrusions [3]. Network intrusions are identified by ascertaining hidden intrusion patterns from large volume of network intrusion detection dataset [8]. Many machine learning (ML) and data mining (DM) algorithms such as ANN, SVM, decision tree (DT), genetic algorithm, naïve Bayesian, fuzzy logic etc. [9] have been employed to classify intrusions in IDS over the last few decades. Today's obtainable network intrusion detection datasets are pointedly imbalanced, multi-faceted, dynamic and composed of different inconsistent data that causes problem to handle entire data. Many machine learning algorithms have been miscarried to identify all sorts of attack with high detection accuracy because some of the attacks occurrence in the dataset are too low compared to other attacks. Today's accessible intrusion detection systems have low detection accuracy and high false positive rate for all sorts of known or unknown network attacks/ intrusions [10].

In this paper, we have introduced a new method for improving detection rate to classify minority-class network attacks/ intrusions using cluster-based under-sampling with Random Forest classifier. Initially, the proposed method classify a data point/ incoming data is attack/ intrusion or not (like normal behaviour), if it's an attack than the proposed method try to classify attack type and later sub-attack type. We have used cluster-based under-sampling technique to deal with class-imbalanced problem and popular ensemble classifier Random Forest for addressing overfitting problem. The foremost purpose of this paper is to enhance the detection accuracy of low-frequency attacks in imbalanced network intrusion classification. We have built several classification models employing ANN [9], NB classifier [11], Random Forest [12], Bagging [13] and compare the performance of these existing learning methods with proposed method. The proposed method achieved 98% detection rate on average on the KDD99 benchmark dataset in compare with existing classifiers.

The rest of the paper is organized as follows. Section II presents the related works. Section III presents an effective multi-layer hybrid method. Section IV provides experimental results based on KDD99 benchmark dataset. Finally, Section V presents conclusions and future works.

II. Related Work

Singh et al. [1] presented an Online Sequential Extreme Learning technique to reduce the slow learning drawback of neural network for intrusion detection. They reduce the volume of the training instances and time complexity by applying beta and alpha profiling respectively. They also discarded irrelevant features by using consistent feature selection techniques and ensemble of filtered. KDD 2009 dataset is used to measure the performance of proposed technique. The proposed IDS achieved an accuracy of 98.66% and 97.67% with false positive rate of 1.74% for binary class and multi class respectively. The proposed IDS outperformed other published techniques based on the accuracy and detection time. Farid et al. [14] introduced a new learning method using NB classifier and an ensemble approach for adaptive intrusion detection. The proposed method can classify a known or unknown instance by considering the votes of a series of classifiers. This method can update the weights of training instances based on the outcome of classification error rates, which reduces the false positive (FP) rates and improves the detection rates (DR). They applied existing data mining algorithms and the proposed method by employing on the KDD99 dataset. The proposed method can abate the false positive rates and achieves high detection rates on different sorts of network intrusions. Farid et al. [15] introduced a new learning algorithm for anomaly-based IDS to detect known or unknown intrusions employing decision tree. It splits and adjusts the weights of each instance based on probabilities until all the sub-dataset belongs to the same class. The weights of every instance is same in conventional decision tree algorithms but in this method weights of every instance changes based on posterior probability. The proposed learning procedure achieved 98% detection rate on KDD99 dataset in comparison with other existing algorithms.

Li et al. [9] proposed a two-step hybrid intrusion detection method construct on k-NN technique and binary classification employing NSL-KDD dataset. Firstly, the proposed method uses the C4.5 algorithm to train binary classifier to identify uncertain classes. Secondly, the proposed method uses the k-NN algorithm to classify uncertain classes. By the amalgamation of these steps, the proposed technique outperforms baselines algorithms like Random Forest, naïve Bayes, backward propagation neural network and k-NN. The F1-scores of this proposed technique is greatly greater than these of baselines algorithms for U2R and R2L low-frequency attacks. Zarpelão et al. [16] presented an analysis on intrusion detection system research for IOT. They were experimented that the IoT research for IDS is still impending and their purposes are to recognize open issues, leading trends and the upcoming research possibilities. They were categorized the IDSs based on the IDS placement stratagem, validation strategy, detection scheme and security intimidation. They were discussed signature-based, specification-based, anomaly-based and hybrid detection mechanisms to develop IDSs for IoT. They also discussed about centralized IDS placement, distributed IDS placement and hybrid IDS placement techniques.

III. Learning Algorithms

A. Learning Scheme

We have used Random Forest as learning scheme in our proposed method. Random Forest is an ensemble learning method and used for classification and regression in supervised learning[17]. It has the capability of classifying big data with higher accuracy [18]. It constructs multiple decision trees depending on arbitrarily chosen group of attribute at training time and outputting the class that is picked from the votes of ensemble of trees as furthermost popular class [11]. There is a relationship between the accuracy and the number of trees, higher number of trees supposed to produce higher accuracy rate. One of the big problems of machine learning is overfitting, it occurs when model overfits to the training data that when new test instance comes to the model from outside of training data, it can not classify it correctly, it gives higher performance only if the training data is used as test data. Random Forest algorithm has solved this problem as it takes vote from the ensemble of trees, so it will not overfit the model. As it considers votes of every decision tree it has constructed, it is supposed to achieve much higher accuracy than a single decision tree [11].

B. Proposed Method

In the first layer of our proposed method, we built a model *M* using dataset *D* and Random Forest technique. When a new instance x_{new} comes in, model *M* predicts that weather it is an attack or not. Following that if x_{new} has been identified as normal we will leave it here but if it is predicted as an attack, we have divided the dataset *D* into two different datasets D_{attack} and D_{normal} respectively. We have worked only with the dataset D_{attack} and it includes 4 different category of attacks, $D_{attack} = D_{DOS} \cup D_{U2R} \cup D_{R2L} \cup D_{Probe}$. Among the 22 attacks in the KDD99 dataset, there is a huge imbalance data in the context of quantity. We had to do resampling to make it balance. In order to do that, in the second layer we have applied under sampling and over sampling techniques to the dataset D_{attack}. Cluster centroid technique was applied to do the under sampling, it partitions the instances *X* with a number of clusters $C_1, C_2, ..., Cn$ and takes only cluster centroids $Centroid(C_i)$ so that informative instances are not lost. After the proceeding, we have replaced the labels of the dataset D_{attack} with category of attacks. Then we have built another model M_{attack} using dataset D_{attack} and Random Forest. The category of attacks of x_{new} dataset has been predicted by M_{attack}. Now the domain of attacks has been shorted, so we have separated the data D_{new} from D_{attack} in the third layer. D_{new} contains instances *X* with only attacks of the predicted category of X_{new}. We have built our final model M_{new} with dataset D_{new} and Random Forest. It will predict the actual class or attack of x_{new}. The architecture of the proposed hybrid method is illustrated in Fig. 1 and algorithm is summarised in Algorithm 1. In Fig. 1, (a) the proposed method classify an incoming data is attack/ intrusion or not (like normal behaviour), if it's an attack then the proposed method try to classify main attack types shown in Fig. 1 (b) and later classify the final attack/ intrusion shown in Fig. 1 (c).

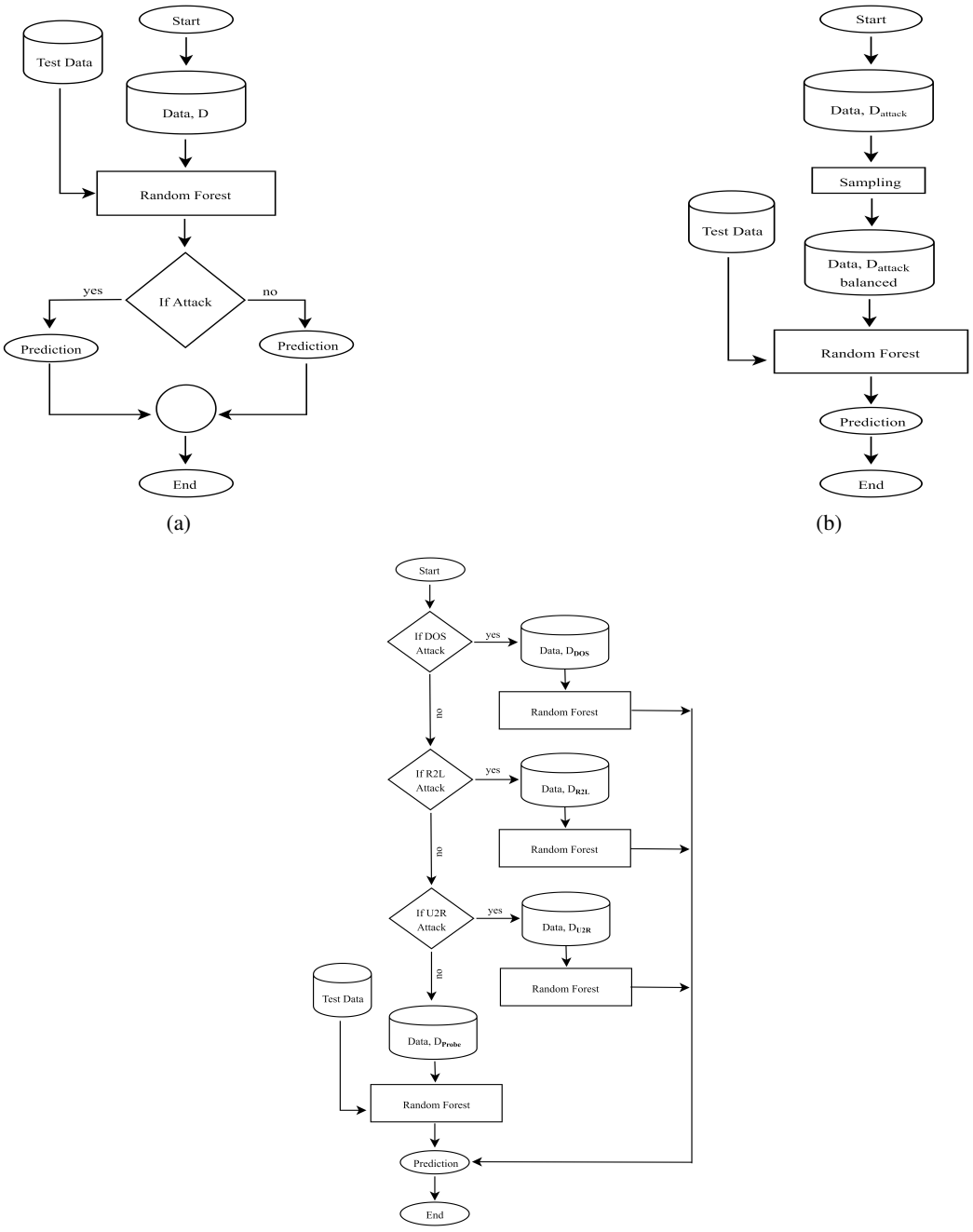


Fig. 1: Proposed method (a) normal or attack detection; (b) main attack types detection; (c) final attack/ intrusion detection.

IV. Experimental Analysis

The performance of classification algorithm is measured with several evaluation metrics in the literature. In this paper, we have used detection rates (DR) [17], false positives (FP) [17] as the standard evaluation metrics. Detection rate and false positive are widely used to estimate the performance of IDS that are shown in Eqs. 1 and 2.

$$DR = \frac{Total\_dected\_attacks}{Total\_attacks} \times 100 \qquad (1)$$

$$FP = \frac{Total\_misclassifed\_process}{Total\_normal\_process} \times 100 \qquad (2)$$

A. Intrusion Detection Dataset

The intrusion detection dataset availability is very limited because it contains information about host, topology and other classified information. The KDD99 standard intrusion detection dataset was used to constitute network intrusion detector to differentiate between normal connections and intrusions in

Algorithm 1 Proposed Hybrid Method

Input: Data D, and a learning scheme.
Output: Prediction, C
Method:

Replace Class labels of D with Normal and Attack;
Derive model, M using D and Random Forest;
Classify, x_{new} **using** M:
$c = M(x_{new})$; // binary class prediction by M
if $c == normal$ **then**
  return c;
else
  $D_{attack} = \emptyset$;
  for each $x_i \in D$ **do**
    if $x_i \in attack$ **then**
      $D_{attack} = D_{attack} \cup x_i$;
    end if
  end for
  Derive model, M_{attack} **using** D_{attack} **and Random Forest**;
  Classify, x_{new} **using** M_{attack}:
  $c = M_{attack}(x_{new})$; // prediction by M_{attack}
  C = $\{DOS, U2R, R2L, Probe\}$
  for each $k_i \in C$ **do**
    if $k_i == c$ **then**
      for each $x_i \in D_{attack}$ **do**
        if $x_i \in k_i$ **then**
          $D_{new} = D_{new} \cup x_i$;
        end if
      end for
    end if
  end for
  Derive model, M_{new} **using** D_{new} **and Random Forest**;
  $c = M_{new}(x_{new})$; // prediction by M_{new}
  return c
end if

the 3rd International Knowledge Discovery and Data Mining Tools Competition [18], [17]. In this dataset, each data point denotes feature values of a class and every class is labeled either normal or attack and all instances are labeled with one of the five types which are mentioned bellow:

Normal: Normal connections are usually behaviors of the daily normal user and generated by visiting web pages, uploading files, downloading files etc [17].

Denial of Service: DoS attack makes the memory resources or power of a device too full or too busy to serve appropriate networking demands. Denial of service is accomplished by flooding the targeted resource with superfluous requests in an attempt to overload systems [11].

Remote to User: In R2L attack, isolated user achieves access of a native account by transferring packets to a device over the internet, it includes xlock, xnsnoop, guest, sendmail dictionary, phf etc [17].

User to Root: In U2R attack hacker takes access of the system with a normal user account and tries to find vulnerabilities for gaining root-user access. Common example of user to root attacks are perl, buffer-overflows, fd-format, load-module, ffb-config, and xterm [18].

Probing: In Probe attack the hacker scans a networking device or a machine to find known vulnerabilities or gather information [13].

KDD99 dataset have 22 dissimilar sorts of attacks that drop into four central categories and total of 41 input attributes and network connections are divided into three groups and they have discrete or continuous values [18]. The basic attributes of each TCP connections are in the first group, it includes the prototype, duration, bytes quantity of IP addresses of source or IP addresses of destination, some flags in TCP connections, and service. The features of network connections suggested by domain knowledge are in the second group and the third group has the statistical attributes that which are calculated by using a two-second time period. We have used training data to train the classifiers and evaluate the classifier models supplying test data from UCI machine learning repository. In testing data, there are some unknown attack examples those are not available in the training data. The number of training and testing examples are shown in Table I.

TABLE I: Number of instances in KDD99 dataset.

Attack Types	Training	Test
Normal	972781	60593
R2L	1126	7015
U2R	52	39
DoS	3883370	223298
Probing	41102	2377
Total	4898431	293322

B. Result

We have implemented the proposed method in Python 3.7 and used scientific Python development environment (Spyder) 3.3.1 (https://www.spyder-ide.org). We have experienced the performance of some popular machine learning algorithms like Random Forest, Bagging, naïve Bayes and ANN on KDD99 dataset. We have supplied testing data for evaluating our proposed method using detect rate and false positive rate. DR is very low for minority class instances which are mostly in R2L and U2R attack types. Some of the strong classifiers like Random Forest, ANN, Bagging, naïve Bayes failed to classify the minority class instances. For our proposed method, we have momentous improvement and DR for minority class instances in our proposed method is above 81%. The proposed method raises detection rates for R2L and U2R attacks type and achieved 87% and 81% detection rate where some popular algorithms able to achieve less than 30%. The improvement of false positive rate also noticeable for our proposed hybrid method. Comparison of results are tabulated in Table II and detection rates of Proposed Method, Random Forest, Bagging, ANN and naïve Bayes are shown in Figure 2.

V. Conclusions & Future Works

Recently, information security has evolved significantly and turn into a vital concern in information technology. IDSs have been employed to protect information from network or host based attacks by utilizing computational intelligence. However, today's accessible IDSs are pattern based also known as misuse

TABLE II: Performance comparison of different strong classifiers with proposed method.

Method	Normal	Probe	DoS	U2R	R2L
Proposed Method (DR %)	**0.99**	**0.99**	**0.99**	**0.81**	**0.87**
Proposed Method (FR %)	**0.03**	0.02	0.02	**0.01**	**0.01**
Random Forest (DR %)	0.99	0.99	0.99	0.12	0.20
Random Forest (FR %)	0.10	0.02	**0.01**	0.01	0.01
Bagging (DR %)	0.98	0.95	0.99	0.10	0.18
Bagging (FR %)	0.12	**0.01**	0.02	0.02	0.03
NB Classifier (DR %)	0.90	0.92	0.95	0.09	0.15
NB Classifier (FR %)	0.11	0.03	0.01	0.02	0.01
ANN (DR %)	0.99	0.99	0.99	0.16	0.23
ANN (FR %)	0.98	0.95	0.99	0.84	0.77

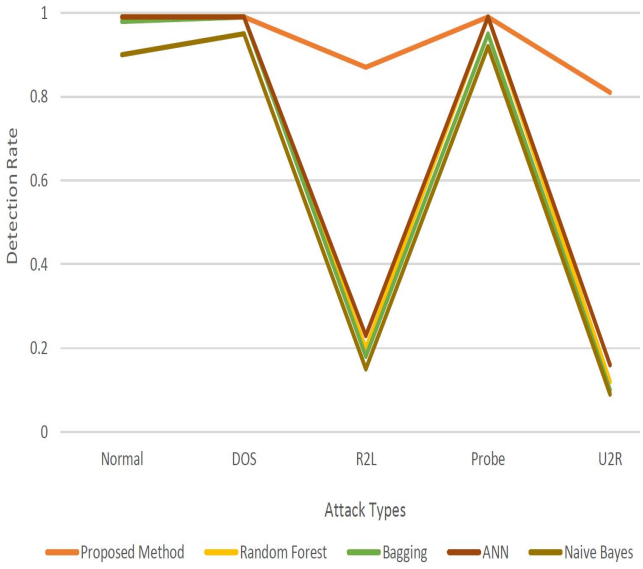


Fig. 2: Detection rates of proposed method, Random Forest, Bagging, ANN, and NB classifier.

based that are not able to detect unknown intrusions. In this paper, we have introduced a new method for improving detection rate to classify minority-class network attacks/ intrusions using cluster-based under-sampling with Random Forest classifier. The proposed method is a multi-layer classification approach, which can process the highly imbalanced big data to correctly identify known or unknown network intrusions. Initially, the proposed method classify a data point/ incoming data is attack/ intrusion or not (like normal behaviour), if it's an attack than the proposed method try to classify attack type and later sub-attack type. We have used cluster-based under-sampling technique to deal with class-imbalanced problem and popular ensemble classifier Random Forest for addressing overfitting problem. The central purpose of this paper is to enhance the detection accuracy of low-frequency attacks in imbalanced network intrusion detection classification. We compared the performance of proposed method with standard data mining algorithms. The experimental results on KDD99 benchmark dataset manifest that the proposed hybrid method raises detection rates and abates false positive rates. The future work will focus on boost the detection rates (DR) of low-frequency attacks in imbalanced dataset and apply this hybrid method into real world NIDS.

REFERENCES

[1] R. Singh, H. Kumar, and R. K. Singla, "An intrusion detection system using network traffic profiling and online sequential extreme learning machine," *Expert Systems with Applications*, vol. 42, no. 22, pp. 8609–8624, 2015.

[2] W. Wang, J. Liu, G. Pitsilis, and X. Zhang, "Abstracting massive data for lightweight intrusion detection in computer networks," *Information Sciences*, vol. 433, pp. 417–430, 2018.

[3] D. M. Farid, L. Zhang, C. M. Rahman, M. Hossain, and R. Strachan, "Hybrid decision tree and nave bayes classifiers for multi-class classification tasks," *Expert Systems with Applications*, vol. 41, pp. 1937–1946, March 2014.

[4] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges," *Computers & Security*, vol. 28, no. 1-2, pp. 18–28, 2009.

[5] M. S. Pervez and D. M. Farid, "Feature selection and intrusion classification in nsl kdd cup 99 dataset employing svms," *8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), Dhaka, Bangladesh*, pp. 1–6, December 2014.

[6] D. M. Farid, N. H. Hoa, J. Darmont, N. Harbi, and M. Z. Rahman, "Scaling up detection rates and reducing false positives in intrusion detection using nbtree," *International Conference on Data Mining and Knowledge Engineering (ICDMKE), Rome, Italy*, pp. 186–190, April 2010.

[7] D. M. Farid and M. Z. Rahman, "Learning intrusion detection based on adaptive bayesian algorithm," *11th International Conference on Computer and Information Technology (ICCIT), Khulna, Bangladesh*, pp. 652–656, December 2008.

[8] D. M. Farid, M. Z. Rahman, and C. M. Rahman, "Chapter title: Mining complex network data for adaptive intrusion detection," in *Advances in Data Mining Knowledge Discovery and Applications*, September 2012, ch. 15, pp. 327–348.

[9] S. B. Y. H. Longjie Li, Yang Yu and X. Chen, "An effective two-step intrusion detection approach based on binary classification and k-nn," *IEEE Access*, vol. 6, pp. 12 060–12 073, 2018.

[10] D. M. Farid and M. Z. Rahman, "Anomaly detection model for network intrusion detection using conditional probabilities," *6th International Conference on Information Technology in Asia (CITA), July, 2009, Kuching, Sarawak, Malaysia*, pp. 104–110.

[11] D. M. Farid, L. Zhang, A. Hossain, C. M. Rahman, R. Strachan, G. Sexton, and K. Dahal, "An adaptive ensemble classifier for mining concept drifting data streams," *Expert Systems with Applications*, vol. 40, no. 15, pp. 5895–5906, November 2013.

[12] D. M. Farid, M. A. Al-Mamun, B. Manderick, and A. Nowe, "An adaptive rule-based classifier for mining big biological data," *Expert Systems with Applications*, vol. 64, pp. 305–316, December 2016.

[13] D. M. Farid, A. Nowe, and B. Manderick, "Ensemble of trees for classifying high-dimensional imbalanced genomic data," in *Proceedings of SAI Intelligent Systems Conference*. Springer, September 2016, pp. 172–187.

[14] D. M. Farid, C. M. Rahman, and M. Z. Rahman, "Adaptive intrusion detection based on boosting and naïve bayesian classifier," *International Journal of Computer Applications*, vol. 24, no. 3, pp. 12–19, 2011.

[15] D. M. Farid, C. M. Rahman, N. Harbi, E. Bahri, and M. Z. Rahman, "Attacks classification in adaptive intrusion detection using decision tree," *International Conference on Computer Science (ICCS), Rio De Janeiro, Brazil*, pp. 86–90, March 2010.

[16] B. B. Zarpelão, R. S. Miani, C. T. Kawakani, and S. C. de Alvarenga, "A survey of intrusion detection in internet of things," *Journal of Network and Computer Applications*, vol. 84, pp. 25–37, 2017.

[17] D. Farid, J. Darmont, N. Harbi, H. H. Nguyen, and M. Z. Rahman, "Adaptive network intrusion detection learning: attribute selection and classification," *International Conference on Computer Systems Engineering (ICCSE), Bangkok, Thailand*, pp. 82–86, December 2009.

[18] D. M. Farid, N. Harbi, and M. Z. Rahman, "Combining naive bayes and decision tree for adaptive intrusion detection," *International Journal of Network Security & Its Applications (IJNSA)*, vol. 2, no. 2, pp. 12–25, April 2010.

Thermal Characteristics of Hybrid Photovoltaic Solar Thermal System during Summer.

Md. Faisal Kader, Ph.D.
Department of Mechanical Engineering
Bangladesh Military Academy,
Chattogram, Bangladesh.
mfaisaliut@yahoo.com

Md. Awal Khan
Department of Mechanical Engineering
Bangladesh Military Academy,
Chattogram, Bangladesh.
awalsetu@yahoo.co.uk

Md. Habibul Huq
Department of Civil Engineering
Bangladesh Military Academy,
Chattogram, Bangladesh.
mdhabibulhuq@gmail.com

Abstract:

A hybrid photovoltaic (PVT) solar thermal system combines a simultaneous conversion of the absorbed solar radiation into electricity and heat. In this study the design of an experimental PVT dual system, both air and water circulation with modifications in the air channel is presented. First modification is to place a thin flat metallic sheet (TMS) inside the air channel and second one is to use painted Black ribbed surfaces at the bottom of the air channel. To observe the variations of heat transmittance with change of the shape of the ribs four experimental setups with a Trapezoidal, Saw teeth forward, a Saw teeth **backward ribbed surfaces and a flat plate have been used. Natural convection is applied instead of forced convection to increase the system's net electrical output & thereby the overall system efficiency. All setups were of the same capacity, projected area, and water heat extraction method and average depth. Performance study was carried out during the months of February to June, 2012 at IUT campus, Gazipur, Bangladesh. Significant performance has been observed with the above stated modifications. In an intense sunny day of March, the maximum temperature of water was found to be 45ºC for Trapezoidal, 44º C for Saw teeth forward, 43ºC for Saw teeth backward and 41ºC for flat plate setup. Maximum temperature of air inside the air channel was found to be 39ºC for Trapezoidal, 38ºC for Saw teeth forward, 37ºC for Saw teeth backward ribbed surfaces and 36ºC for flat plate in an intense sunny day of march 2012 with an ambient temperature of 34ºC. The average efficiency from all calculated values is found to be 64% for Trapezoidal, 62% for Saw teeth forward, 61% for Saw teeth backward and 58% for flat plate setup.**

Keywords— Hybrid thermal system, solar thermal, solar photovoltaic

I. INTRODUCTION

The need for alternative energy sources is getting urgent, hence the development of renewable energy is moving fast. Solar energy is an important clean, cheap and abundantly available renewable energy. The earth receives an incredible supply of solar energy. Different Photovoltaic cell converts only 5-15% of the incoming solar radiation into electricity, with the greater percentage converted into heat. The solar radiation converted into heat increases the temperature of the PV modules, resulting in the drop of their electrical efficiency. This undesirable effect can be partially avoided by applying a suitable heat extraction mode with a fluid circulation, keeping the electrical efficiency at a satisfactory level. Furthermore, this extracted heat can be utilized for heating water. For this purpose the concept of Hybrid photovoltaic/ thermal (PV/T or PVT) solar systems developed. The PV modules that were (are) combined with thermal units, where circulating water of lower temperature than that of PV modules was heated to constitute the Hybrid PVT solar systems. It provides electrical and thermal energy, increasing therefore the total energy output from PV modules using the same projected area. The objectives of this study are to carry out an experimental investigation to compare the performances of four PVT collectors having three different ribbed surfaces (Trapezoidal, Saw teeth (tooth) forward, Saw teeth backward rib) and a flat plate opposite to each absorber plate. All collectors were of same capacity, projected area, and average depth.

Many theoretical and experimental works have been carried out for the improvement of hybrid PVT solar systems since its appearance in 1980's. Among the first who provided the main concepts of these systems with results, by the use of water or air as heat removal fluid. Bhargava *et al.* [1] present results regarding the effect of air mass flow rate, air channel depth, length and fraction of absorber plate area covered by solar cells (packing factor, PF) on single pass, Sopian *et al.* [2] on single pass and double pass hybrid PVT system performance. Garg and Adhikari [3] present a variety of results regarding the effect of design and operation parameters on the performance of air type PVT systems.

In the above works the calculated thermal efficiencies of liquid type PVT systems are in the range of 45% to 65%, the higher values derived for systems that include thermal losses suppression by using air gap with glazing. Regarding air type PVT systems, the thermal efficiency depends strongly on air flow rate, air duct depth and collector length. For higher values of air flow rate, small air duct depth and long PVT systems thermal efficiencies up to about 55% are given by the theoretical models. The packing factor is an important parameter in most of the above papers.

Because of their easier construction and operation, hybrid PVT systems with air heat extraction are more extensively studied, mainly as an alternative and cost effective solution

to building integrated PV (BIPV) systems. Following the above referred studies, test results from PVT systems with improved air heat extraction are given by and from roof integrated air-cooled PV modules by present design and performance studies regarding air type building integrated hybrid PVT systems.

The building integrated photovoltaic is going to be a sector of a wider PV module application in future. The works on building integrated air-cooled photovoltaic include the studies on the multi-operational ventilated PVs with solar air collectors, the ventilated building PV facades and the design procedure for cooling air ducts to minimize efficiency loss [4, 5].

Despite these improvements, commercial application of PVT/air collectors is still marginal, but it is expected to be wider in the future where many building facades and inclined roofs will be covered with photovoltaic. PVT/water systems are more expensive than PVT/air systems due to additional cost of the thermal unit with the pipes for the water circulation. On the other hand water from mains does not often exceeds 25^0C and ambient air is usually higher during summer in low latitude countries and water heat extraction is of more practical value at these locations as it can be used during all seasons. The liquid type hybrid PVT systems are less studied than air type systems.

Regarding recent works, modeling results [6], the study on domestic PVT systems [7] can be referred. PVT/ water collectors can replace thermal collectors for water heating in the domestic and industrial sectors, but they are not yet cost effective and this is the main reason for their niche market penetration.

An experimental study of facade-integrated photovoltaic/water-heating system has been reported done by Chow *et al.* [8]. This work describes an experimental study of a centralized photovoltaic and hot water collector wall system that can serve as a water pre-heating system. Collectors are mounted at vertical facades. Different operating modes were performed with measurements in different seasons. Natural water circulation was found more preferable than forced circulation in this hybrid solar collector system. The thermal efficiency was found 38.9% at zero reduced temperature, and the corresponding electricity conversion efficiency was 8.56%, during the late summer of Hong Kong.

In China, the energy performance of PV/water-heating collector systems with natural circulation of water has been examined by the authors Chow *et al.* [9, 10] based on the weather conditions of Hong Kong and Hefei. These studies showed that the use of a flat-box type thermal absorber is very effective. In Hong Kong for a stand-alone PVT collector system with mono-crystalline solar cells, the daily thermal efficiency was found 48.3% in winter and 45.4% in summer at zero reduced temperature (i.e. when the initial water temperature in the storage tank is as cold as the mean ambient temperature on the day of measurement). This was for water storage to collector-area ratio (M/Ac) of 96.6 kg/m^2 and at a desirable tilt angle.

Design and performance improvements of hybrid PVT systems with water or air as heat removal fluid have been carried out at the IUT including modifications that contribute to the decrease of PV module temperature and to improve the total energy output (electrical and thermal) of the PVT systems [11]. Two systems (PVT/UNGLAZED and PVT/GLAZED) were tested outdoors, consisted of pc-Si PV modules and heat exchanger of copper sheet with copper pipes. Also, PVT solar water heaters of ICS [12] and of thermosiphonic type have been studied.

Tripanagnostopoulos [13] at the University of Patras, Greece has done an extended research on PVT systems aiming at the study of several modifications for system performance improvement. A new type of PVT collector with dual heat extraction operation, either with water or with air circulation is presented. Experiments with dual type PVT models of alternative arrangement of the water and the air heat exchanging elements were performed. The most effective design was further studied, applying to it low cost modifications for the air heat extraction improvement. The modified dual PVT collectors were combined with booster diffuse reflectors, achieving a significant increase in system thermal and electrical energy output.

The present work studied for performance of Hybrid Photovoltaic Thermal (PVT) dual solar collector using different ribbed surfaces placed opposite to the air channel wall. Tests were carried out in the months of February to June, 2012 at IUT campus, Gazipur, Bangladesh.

II. Experimental Setup

Four experimental setups were fabricated with similar design and dimensions except the shape of the ribs. All works were done in the Mechanical Workshop of Islamic University of Technology (IUT), Gazipur, Bangladesh. The whole setup was constructed in a wooden box. PV panels were set at the top of the box. Thin Metallic Sheet (TMS) was placed at the middle part of the box on top of which pipes were set for water circulation. Air channel of 0.1m height was kept under the water heat exchanger. Different ribbed surfaces (*Trapezoidal, Saw* **teeth** *forward, and Saw* **teeth** *backward) and a Flat plate* of same height and dimension were placed on opposite wall of the air channel. The whole inner portion of the box was insulated. TMS and

Ribbed plates were painted black for better heat absorption. Fig . 1 shows the schematic diagram of each experimental setup with different ribbed surfaces. Details of hybrid collector and test specimens used in this study are shown in Fig. 2 and in Fig .3.

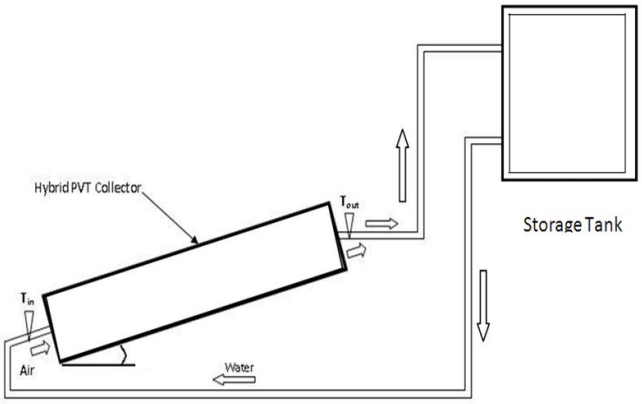


Fig. 1: Schematic diagram of the experimental setup.

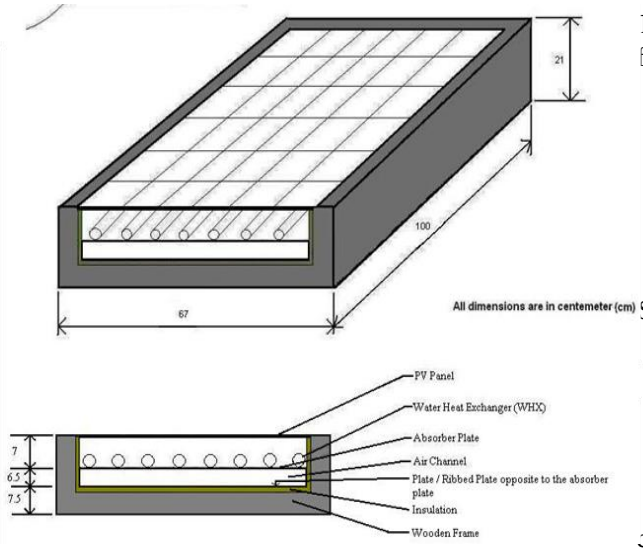
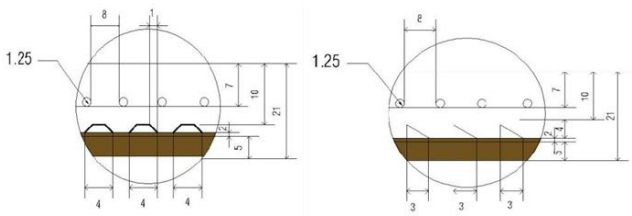


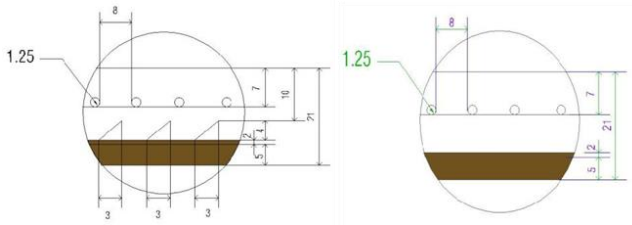
Fig . 2: Details of Hybrid Collector. (All dimensions are in cm)

A. *Test procedures:*

All the collectors were installed in front of the IUT workshop where there is no obstacle to sunshine and faced towards south with an inclination angle of 23.5^0 which is the best angle to collect as much as available radiation. Photograph of which are shown Fig . 1. All storage tanks were filled early in the morning with fresh water. Temperatures were measured for PV panel (T_{PV}) and water heat exchanger (T_{WHX}), by using seven 36 S.W.G. Chromel-Alumel thermocouples. Selector switches were used to switch between the thermocouples. Ambient temperatures (Tamb), solar radiation (G) were recorded hourly each day from at 8 AM up to 6 PM. All readings were recorded in specific data sheets.



(a) Trapezoidal rib (b) Saw teeth forward rib



(c) Saw teeth backward rib (d) Flat plate

Fig 3: Test specimens. (All dimensions are in mm)

B. *Equations used*

To calculate the thermal efficiency (η_{th}) of the collector the following equations are used:

Heat energy absorbed at any time for the projected area,

$$Q_{ab} = m \times C_p \times \Delta T_w / A_a \quad (Wm^{-2})$$

Here, ΔT_w is the temperature difference between inlet and outlet of water.

Solar radiation received in that time for the projected area, G (Wm^{-2})

So thermal efficiency can be found from, $\eta_{th} = Q_{ab} / G$

III. Results and Discussions

Performance study of four hybrid PVT systems were carried out in IUT Campus, Gazipur, Bangladesh during the months of February to June 2012. PV panel temperature and water heat exchanger surface temperature were recorded and plotted against time. These are presented from Fig. 4 to Fig. 13. Thermal efficiency is calculated for all the setups and average plotted in and presented in Fig. 14. Thermal output is emphasized more than the electrical output in this study. As no electrical load is used as a part of the setup, output from PV panel is not hampered.

A. *PV panel temperature*

Temperature distributions of PV panels of different setups with time are shown in Fig.4 to Fig. 8. As the PV panel receives the solar energy directly on top of the setup, panel temperature rises very quickly with time. The figures shows similar patterns of temperature survey with slight difference among four setups. PV panel temperature rises from 8:00 am to 12:00 noon and then decreases after 2:00 pm rapidly. In a sunny day of 15 February 2012, the maximum temperature of PV panel is found 61ºC for Flat plate, 57ºC for

Trapezoidal, 57ºC for Saw teeth forward and 59ºC for Saw teeth backward plate. In a sunny day of 30th February 2012 the maximum temperature of PV panel is found 60ºC for Flat plate, 56ºC for Trapezoidal, 57ºC for Saw teeth forward and 58ºC for Saw teeth backward plate. In an intense sunny day of 15 March 2012, the maximum temperature of PV panel is found 62ºC for Flat plate, 56ºC for Trapezoidal, 57ºC for Saw teeth forward and 59ºC for Saw teeth backward plate. In a sunny day of 30March2012, the maximum temperature of PV panel is found 62ºC for Flat plate, 56ºC for Trapezoidal, 57ºC for Saw teeth forward and 60ºC for Saw teeth backward plate. In a sunny day of 15 April 2012, the maximum temperature of PV panel is found 62ºC for Flat plate, 55ºC for Trapezoidal, 56ºC for Saw teeth forward and 59ºC for Saw teeth backward plate. In a sunny day of 30April2012, the maximum temperature of PV panel is found 61ºC for Flat plate, 54ºC for Trapezoidal, 56ºC for Saw teeth forward and 59ºC for Saw teeth backward plate. In a partly cloudy day of 15 May 2012, the maximum temperature of PV panel is found 40ºC for Flat plate, 44ºC for Trapezoidal, 43ºC for Saw teeth forward and 42ºC for Saw teeth backward plate at the time was 02:00 PM. In a Rainy day of 30 May 2012, the maximum temperature of PV panel is found 40ºC for Flat plate, 41ºC for Trapezoidal, 40ºC for Saw teeth forward and 39ºC for Saw teeth backward plate. In an intense sunny day of 15 June 2012, the maximum temperature of PV panel is found 64ºC for Flat plate, 63ºC for Trapezoidal, 62ºC for Saw teeth forward and 60ºC for Saw teeth backward plate. In a sunny day of 30June 2012, the maximum temperature of PV panel is found to be 62ºC for Flat plate, 61ºC for Trapezoidal, 60ºC for Saw teeth forward and 59ºC for Saw teeth backward plate.

Setup with Saw teeth forward and Saw teeth backward ribs setup gave lower temperature of the panel which shows better cooling of PV. Trapezoidal ribbed plate gave moderate cooling better than that of flat plate setup. In flat plate setup PV temperature is found always high, which shows inadequate heat transfer for PV cooling.

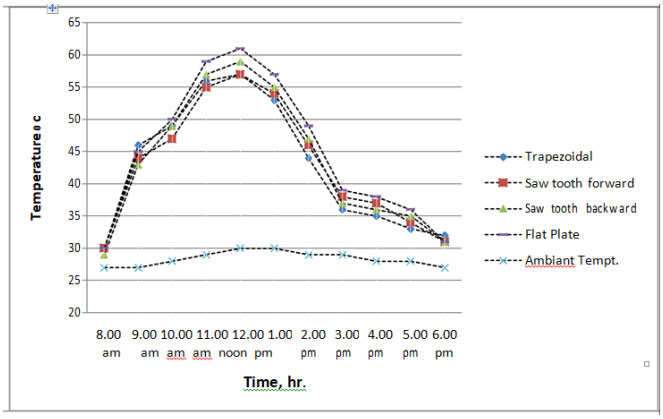


Fig. 4: Temperature distribution of PV panel temperature on 15/02/2012

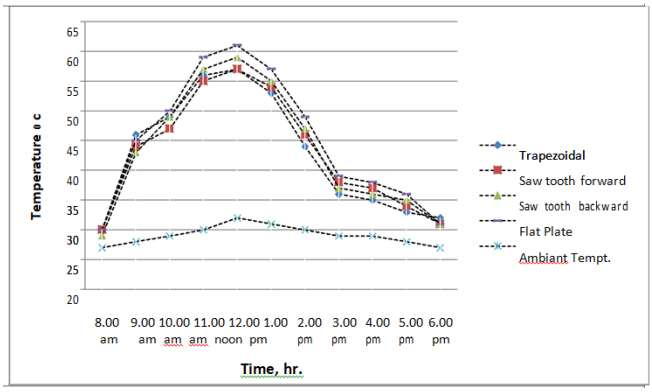


Fig. 5: Temperature distribution of PV panel temperature on 15/03/2012

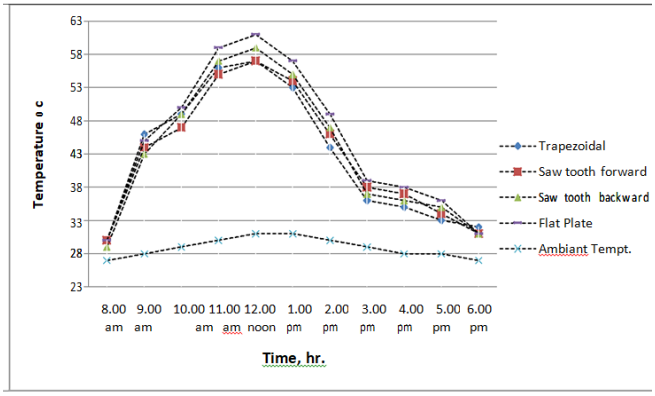


Fig. 6: Temperature distribution of PV panel temperature on 15/04/2012

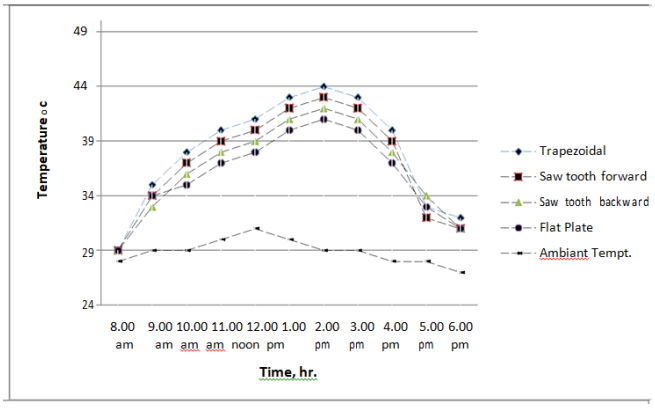


Fig. 7: Temperature distribution of PV panel temperature on 15/05/2012

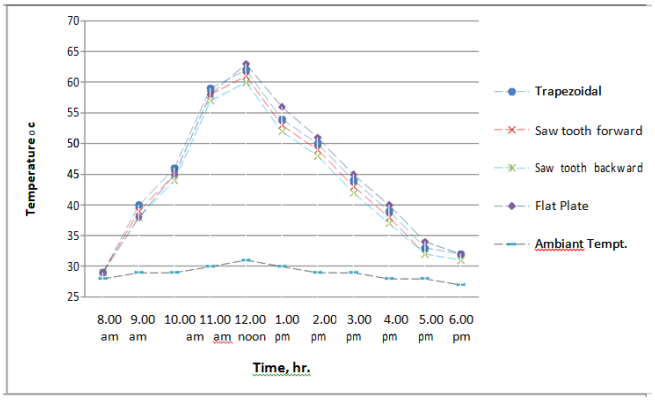


Fig.8: Temperature distribution of PV panel temperature on 15/06/2012

B. Water Heat Exchanger temperature

Water Heat Exchanger (WHE) surface temperature distributions of different setups with time are shown in Fig. 9 to Fig. 13. Water heat exchanger with TMS is placed below the PV panel at the middle section of the setup. Heat energy is absorbed here from the PV rear surface and water is heated flowing through it. It is evident from these figures that the temperature of this heat exchanger also rises with time from 8:00 am to 12:00 noon and then decreases after 2:00 pm rapidly. In a sunny day of 15February 2012, the maximum temperature of WHE is found 50ºC for Flat plate, 48ºC for Trapezoidal, 49ºC for Saw teeth forward and 51ºC for Saw teeth backward plate. In a sunny day of 30 February 2012 the maximum temperature of WHE is found 49ºC for Flat plate, 46ºC for Trapezoidal, 50ºC for Saw teeth forward and 48ºC for Saw teeth backward plate. In an intense sunny day of 15March 2012, the maximum temperature of WHE is found 48ºC for Flat plate, 46ºC for Trapezoidal, 49ºC for Saw teeth forward and 47ºC for Saw teeth backward plate. In a sunny day of 30March 2012, the maximum temperature of WHE is found 48ºC for Flat plate, 46ºC for Trapezoidal, 49ºC for Saw teeth forward and 48ºC for Saw teeth backward plate. In a sunny day of 15April 2012, the maximum temperature of WHE is found 49ºC for Flat plate, 52ºC for Trapezoidal, 52ºC for Saw teeth forward and 49ºC for Saw teeth backward plate. In a sunny day of 30 April 2012, the maximum temperature of WHE is found 49ºC for Flat plate, 47ºC for Trapezoidal, 51ºC for Saw teeth forward and 49ºC for Saw teeth backward plate. In a partly cloudy day of 15 May 2012, the maximum temperature of WHE is found 36ºC for Flat plate, 38ºC for Trapezoidal, 37ºC for Saw teeth forward and 36ºC for Saw teeth backward plate at the time was 02:00 PM. In a Rainy day of 30 May 2012, the maximum temperature of WHE is found 34ºC for Flat plate, 37ºC for Trapezoidal, 36ºC for Saw teeth forward and 35ºC for Saw teeth backward plate. In an intense sunny day of 15 June 2012, the maximum temperature of WHE is found 45ºC for Flat plate, 48ºC for Trapezoidal, 47ºC for Saw teeth forward and 46ºC for Saw teeth backward plate. In a sunny day of 30 June 2012, the maximum temperature of WHE is found 45ºC for Flat plate, 47ºC for Trapezoidal, 48ºC for Saw teeth forward and 46ºC for Saw teeth backward plate. Average good temperature is found in case of setup with Saw teeth forward rib. Average temperature difference between PV panel and WHE is around 10-12ºC.

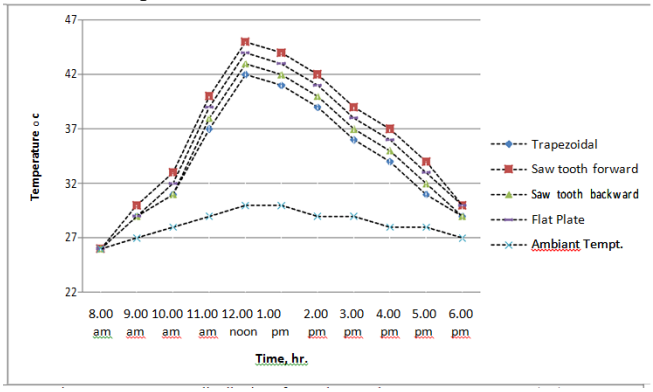


Fig. 9: Temperature distribution of water heat exchanger temperature on 15/02/2012.

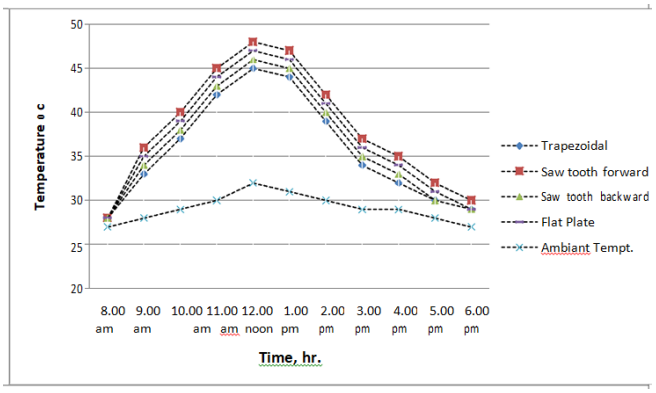


Fig. 10: Temperature distribution of water heat exchanger temperature on 15/03/2012.

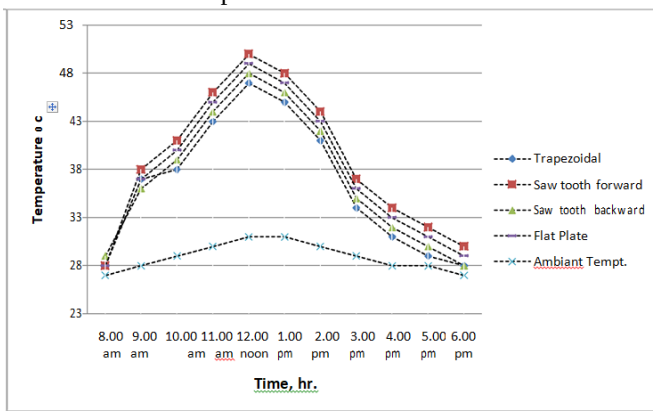


Fig. 11: Temperature distribution of water heat exchanger temperature on 15/04/2012.

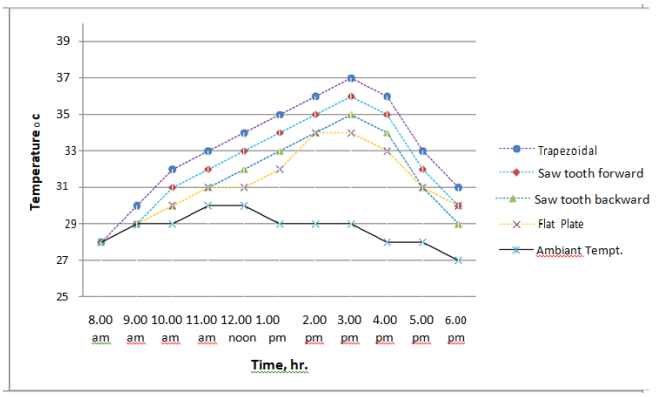


Fig. 12: Temperature distribution of water heat exchanger temperature on 15/05/2012.

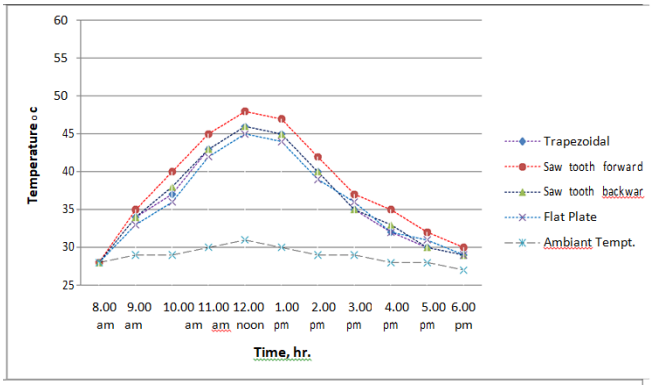


Fig. 13: Temperature distribution of water heat exchanger temperature on 15/05/2012.

C. *Thermal efficiency comparison*

Thermal efficiency regarding water is compared as a function of $\Delta T/G$ ($Kw^{-1}m^2$).ΔT is the temperature difference between input fluid and the ambient temperature. Efficiency of all four setups is compared along with the work of Tripangnostopoulos (2007) and Rezwan (2011) in Fig. 14 for comparison. Tripangnostopoulos studied a hybrid PVT/dual system having corrugated rib on opposite air channel with WHE placed just under the PV panel in 2007 at the University of Patras, Greece. .Karim and Akhanda (2011) studied hybrid photovoltaic thermal solar system using three different rib surfaces (Triangular, Semicircular, and Square) and flat plate in 2011 at IUT.

The present study was carried out using hybrid photovoltaic thermal solar system using three different ribbed surface (Trapezoidal, Saw teeth forward, Saw teeth backward) and a flat plate in 2012 at IUT, Bangladesh. The efficiency of Trapezoidal is 36% to 73%, Saw teeth forward is 37% to 70%, Saw teeth backward is 36% to 67% and Flat plate is 33% to 59%. Tripangnostopoulos (2007) PVT system is found to vary from 45% to 62% and Rezwan (2011) reported results to vary from 35% to 70%. The average efficiency from all calculated values in the present study is found to be 67% for Trapezoidal, 62% for Saw teeth forward, 61% for Saw teeth backward and 54% foe flat plate setup.

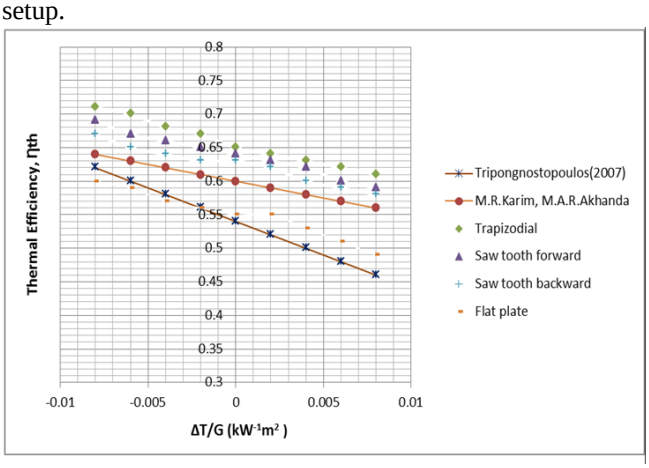


Fig. 14: Comparison of thermal efficiencies of all four systems.

The range of efficiency varies with the operating temperature of the system. The thermal efficiency of PVT/dual system for water heat extraction is extended in negative $\Delta T/G$ axis, as some experiments were performed for ambient temperature being higher than the water temperature at system input. Data taken for comparisons of setups are similar with $\Delta T/G$ values, to make the comparison more correct.

IV. Conclusions

From this experimental study the following conclusions are drawn;

- With the increase of PV panel temperature the water temperature also increase and reach a maximum value at a maximum PV temperature around noon. The temperature then decrease slowly with the ambient temperature.
- The energy absorbed in the system from 8AM to 12 Noon is faster than that absorbed in the afternoon for all systems, tested in this study.
- Average water temperature rise is found to be better in Trapezoidal and Saw teeth forward ribbed surface then those of surfaces with Saw teeth backward and Flat plate. Setup with Saw teeth forward ribbed surface shows better performance than that of Flat plate but its average performance is lower than that of setup with Trapezoidal and Saw teeth forward ribs.
- Thermal efficiency is found to be the best in the setup with Trapezoidal ribs among all setups, applying similar experimental conditions (T/G). Efficiency of Saw teeth forward, saw teeth backward and flat plate is also found satisfactory within the range of experimental conditions.
- PVT/dual system can be used to heat water depending on the weather conditions and building needs. The water heat extraction part could operate mainly during periods of higher ambient temperature and the air heat extraction part to operate mainly when the ambient temperature is low. In mild weather conditions (like Bangladesh) it is possible to operate both heat extraction modes, if it is considered useful for the application.

References

[1] A. K. Bhargava, H. P. Garg and R. K. Agarwal, "Study of hybrid solar system- Solar air heater combined with solar cells," *Energy conversion and Management*, vol. 31, pp. 471-479, 1991.

[2] K. Sopian, H. T. Liu, S. Kakac and T.N. Veziroglu,"Performance analysis of photovoltaic thermal air heaters," Energy Conversion and Management vol 37, no. 11,pp 1657-1670, 1996.

[3] H. P. Garg and R. S. Adhikari, "Conventional hybrid photovoltaic /thermal (PV/T) air heating collectors: steady state simulation," Renewable Energy, vol. 11, pp 363-385, 1997.

[4] B. J. Brinkworth, R. H. Marshall and Z. Ibrahim, "A validated model of naturally ventilated PV cladding," Solar Energy vol. 69, pp. 67-81, 2000.

[5] B. J. Brinkworth and M. Sandberg, "Design procedure for cooling Ducts of minimize efficiency loss due to temperature rise in PV arrays," Solar Energy, vol. 80, pp. 89-103, 2006.

[6] J. Ji , T. T. Chow and W. He, "Dynamic performance of hybrid photovoltaic/thermalcollector wall in Hong Kong," Building and Environment, vol. 38, part 11, pp.1327-1334 2003.

[7] J. S. Coventry, , Lovegrove. K., 2003. “Performance of a concentrating photovoltaic/thermal solar collector,” Solar Energy, vol. 78, pp. 211-222, 2003.

[8] T. T. Chow, W. He and J. Ji, “Hybrid photovoltaic-thermosyphon water heating system forresidential application,” Solar Energy, vol. 80, part 3, pp. 298-306, 2006.

[9] T. T. Chow, W. He, J. Ji and A. L. S. Chan, “Performance evaluation of photovoltaic-thermosyphon system for subtropical climate application,” Solar Energy, vol. 81, part 1, pp. 123-130, 2006.

[10] T. T. Chow, W. He and J. Ji, “Photovoltaic-thermal collector system for domestic application,” ASME Journal of Solar Energy Engineering vol. 129, pp. 205-209, May 2007.

[11] M. R. Karim and M. A. R. Akhanda,”Study of a hybrid photovoltaic thermal (PVT) solar systems using different ribbed surfaces opposite to absorber plate,” Journal of Engineering and Technology (JET), vol. 09, no.01, June 2011.

[12] Y. Tripanagnostopoulos, T. H. Nousia and M. Souliotis,” Hybrid PV ICS systems,” Proceedings of International Conference, WREC V, Florence, Italy, pp. 1788-1791. 20-25 September, 1998.

[13] Y. Tripanagnostopoulos,”Aspects and improvements of hybrid photovoltaic/thermal Solarenergy systems,”Solar Energy, vol. 81, pp 1117–1131, 2007.

Seizure and Non-Seizure EEG Signals Detection Using 1-D Convolutional Neural Network Architecture of Deep Learning Algorithm

Tanima Tasmin Chowdhury[1,2], Afrin Hossain[1], Shaikh Anowarul Fattah[1], Celia Shahnaz[1]
Department of Electrical and Electronic Engineering
[1]Bangladesh University of Engineering and Technology, [2]University of Asia Pacific, Dhaka, Bangladesh
Email: {tanimatasmin, afrinfive, fattah92, celia.shahnaz}@gmail.com

***Abstract*—In this paper, seizure activities of EEG signals have been detected exploiting 1-D convolutional neural network architecture of deep learning algorithm. In past some years, many algorithms have been developed for seizure detection. Traditional method of seizure detection includes manual feature selection and classifier separately. In recent years, convolutional neural network architecture of deep learning gets very much popularity in place of traditional method.Though, very few works of convolutional neural network are found in literature for one dimensional signal classification. However, one dimensional EEG and ECG signal classifications are necessary to detect and investigate diseases like seizure activity, sleep apnea or arrhythmia and so on. The proposed method of one dimensional EEG signal classification for epileptic seizure detection involves the smallest and simplest 1-D convolutional neural network architecture for automated feature selection. The projected architecture is implemented to detect several kinds of seizure and non-seizure EEG signals and found achieving greater accuracy and sensitivity in comparison to that acquired by state-of-the-art methods involving manual feature selection and the identical EEG dataset.**

***Keywords*—Electroencephalogram (EEG), epilepsy, seizure, 1-D convolutional neural network (CNN)**

I. Introduction

There are many types of seizures. Seizure types depend on where and how they begin in the brain. Mostly they sustain from 30 seconds to two minutes. It is an emergency situation if any person is attacked by seizure and the duration is more than five minutes. Seizures are more common disease. Seizures can happen for many reasons. Among the reasons, some are person experiencing head injury, brain infection, stroke, fever and many other health problems. But still, when a person is attacked by seizure, cause cannot be defined accurately [1].

Many processes or algorithms have been used to detect seizure. As we see in traditional methods, features have to be extracted from the EEG signal. Normal Inverse Gaussian (NIG) statistical modeling parameters or higher order statistics (HOS) have been used as features for EEG signal. These features have been analyzed in Empirical Mode Decomposition (EMD) domain or in Discrete Wavelet Transform (DWT) domain. However, no specific IMF or time-frequency band has been mentioned in these works and thus increased the size of feature set and computational burden. Then classifier like SVM or ANN hve been used to detect the seizure from the normal EEG signal. These are actually lengthy process as well as time consuming method [2], [3], [4]. So, we have been motivated in this work to use Convolutinal Neural Network (CNN) scheme for detection and classification of seizure and non-seizure 1D EEG signal which will replace the above mentioned lengthy and time consuming methods with automatic feature selections.

In recent years, the deep learning with CNN architecture for automatic feature selection is gathering popularity in signal processing. This method involves fast computation as the manual feature and classification step are not done here. EEG signal has the feature of low frequency signal for long time period and the feature of high frequency for small time period. There is a kind of hierarchy among these features of EEG signal. Deep learning is also known as hierarchical learning which encodes hierarchy of features. So for EEG signal classification, we are proposing deep learning based method.

The paper is arranged as follows: we have presented in section II method description; in section III results & analysis which includes dataset description, goodness of features, performance assessment & comparison of proposed and existing methods; and in section IV conclusions.

II. Method Description

The sequence of EEG signal ranges up to 60 Hz [5]. 60 Hz above of an EEG signal are noise. The noise of EEG signal has been removed by passing the signal over a 6th order Butterworth filter. The cutoff frequency of the low pass filter has been selected 60 Hz.

In this paper, we have used convolution neural network method. This method is replacing the traditional & manually selecting feature extraction and classification steps.

CNN model automatically reads the pattern of EEG signal from the data and performs classification. On the other hand in traditional approach, at first features are extracted, major features are separated and then a classifier is used for classification. Full traditional approach is a manual process. But through CNN, the scheme is fully automated.

A CNN has three layers. One input layer, one output layer and multiple hidden layers. The CNN architecture consists of convolutional layers, maxpooling layers and fully connected layers. In these layers, unknown number of mathematical

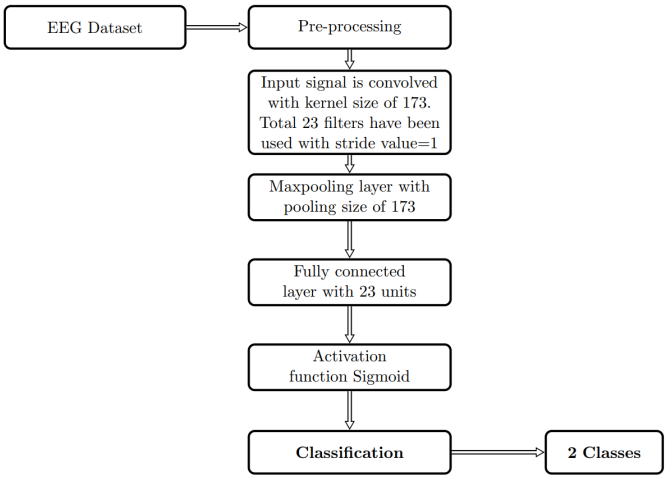


Fig. 1: The Proposed Model Architecture

operations are being performed.

CNN model is based on convolutional process. The considered data signals are seizure and the non-seizure EEG signals. Each time two different data-class has been considered for analysis of this work. The input signal is convolved with the selected kernel size 173. Total 23 numbers of filters has been considered in each analysis of 2 classes. The kernel size of 173 and 23 number of filters have been selected because the sampling rate of EEG signal is 173.61. The length of input signal is 4097 as the sampling rate is 173.61 and the each set data segments are of 23.6 seconds. After each convolution operation, the output data contains the feature map of the input signal. Here, stride is selected at 1.

Down sampling operation is performed in max pooling layer. Performing max pooling layer on the output of the convolutional layer reduces the dimension of the output neuron. Through this layer, maximum value has been selected from the output feature maps.Max pooling size is set in this work equal to the kernel size that is 173. So, this down sampling operation consequently reduces the computational intensity. After that, the output from max pooling is flattened out and the output is applied as the input for the fully connected layer (FC). FC layers are 1D vector. FC layer see the output of the previous layer.At first, fully connected layers are set to 23 which is equal to the number of filters. In the next step, FC layers are set to 2. It actually works as the classifier. This layer determines the features correlating to a particular class.

Two activation functions have been used in this paper. One is sigmoid and another is softmax. These are non linear functions. Activation functions are used as a decision maker as activation function activates or fires a neuron from the output of fully connected layer. Activation function sigmoid is used which thresholds at zero. Thus, it keeps the volume unchanged. This layer increases the nonlinear properties of the model. The methodology used in this paper is presented in Fig 1.

III. Results and Analysis

A. Dataset Description

For analyzing, we have used the EEG signal data from Bonn university seizure database [6] and proposed CNN for epileptic seizure detection. Dataset description are as follows:

Recorded data has been collected by placing 10-20 electrodes which is an internally recognized standard system. In the database, there are five data sets. The notations for five data sets are declared as A to E. The EEG signals of each set has been recorded by 128 amplifier system. The data available in each set are 100 single channels EEG segments with duration of 23.6 seconds. Set A consists of those EEG data segments which have been recorded at relaxed with awaken state and eyes open. Set B consists of those EEG data segments which have been recorded at relaxed with awaken state and eyes closed. Both A and B sets data are of five healthy volunteers. The EEG signal segment in set C, D & E have been taken from 5 seizure patients. Sets C, D and E have been taken from EEG archive of pre-surgical diagnosis. EEG signal segments in D set have been recorded from the epileptogenic zone during seizure-free intervals. On the other hand, EEG signal segments in C set have been recorded from hippocampal formation of opposite hemisphere of the brain during seizure-free intervals. E set is found solely with the recording of those EEG signal segments of the patients undergoing epilepsy .

The data acquiring process of EEG signal involves sampling rate of 173.61 Hz and 12 bit analogue-to-digital converter. From the visual examination of continuous multichannel EEG recordings, the band pass filter of 0.5340 Hz (12 dB/oct) has been used to remove the desired segments.

In this investigation, we have used all the five datasets from A to E of Bonn university seizure database. The CNN has been done using keras neural network library. We have run the code to differentiate between 2 classes of EEG signals: A-E, B-E, C-E, D-E, ABCD-E & A-C. Figure II-VI show the raw EEG signals of the above mentioned pairs which are required to be classified. It is well-understood from the figures that without extracting appropriate features, it is very difficult to differentiate the classes from the raw EEG signals.

Training set and the testing set are not definitely defined. For binary classification, randomly fifty percent of the total segments are selected for training purpose and the rest fifty percent are selected for testing purpose. Ensuring the performance of the adopted methodology and the system for each binary classification, the methodology has been run for 10-times randomly for cross checking. As a result, this step ensures the unbiased performance of the adopted methodology as the each time randomly different sets of training and testing sets have been automatically selected. Standard deviation has been calculated for each binary classification which shows the deviation of the ten sets of results.

B. Goodness of Features

The features extracted by CNN architecture have been exploited to understand the features quality by Geometrical Separability Index (GSI) and Bhattachariya Distance (BD)

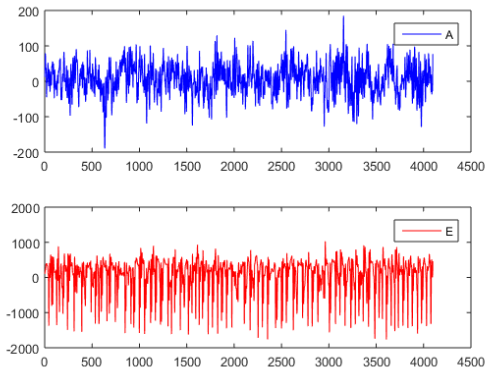


Fig. 2: Time Domain Plot of A and E Classes of EEG Signals

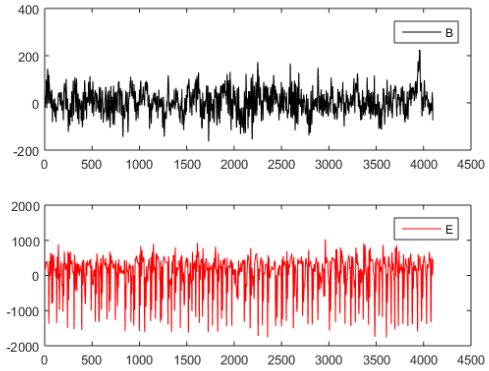


Fig. 3: Time Domain Plot of B and E Classes of EEG Signals

value [4]. GSI and BD values are two measures of inter-class distance and intra-class compactness; respectively. To which extent two classes can be distinct together in the neighboring sense is indicated by GSI value. The 0 GSI value directs two class as inseparable while the GSI value greater than 0.5 estimates two classes as detachable.

Intra-class compactness of clusters is shown by BD value. The smaller BD value postulates more compactness of the features in similar class. The GSI and BD values for mentioned all five classes in the dataset are shown in Table I and Table II.

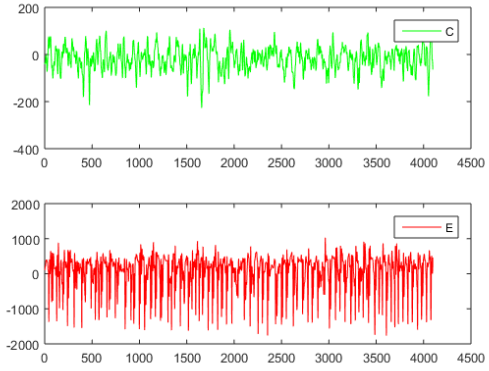


Fig. 4: Time Domain Plot of C and E Classes of EEG Signals

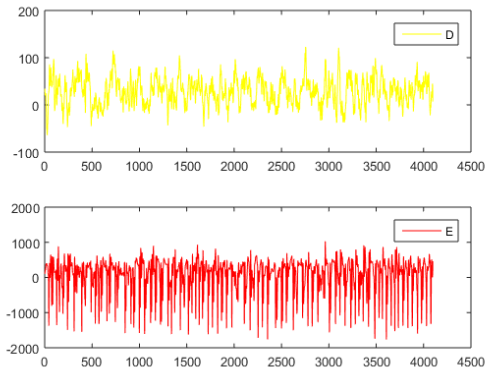


Fig. 5: Time Domain Plot of D and E Classes of EEG Signals

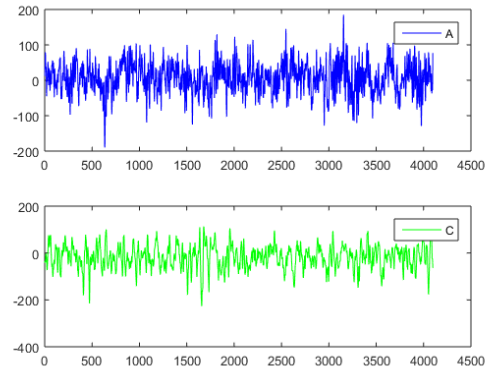


Fig. 6: Time Domain Plot of A and C Classes of EEG Signals

TABLE I: GSI VALUES OF THE PROPOSED SCHEME

Class	A	B	C	D	E
A	0.04	0.90	0.81	0.90	0.93
B	0.90	0.33	0.96	0.95	0.50
C	0.81	0.96	0.4	0.98	0.98
D	0.90	0.95	0.98	0.36	0.95
E	0.93	0.50	0.98	0.95	0.42

TABLE II: BD VALUES OF THE PROPOSED SCHEME

A	B	C	D	E
0.0193	0.0389	0.0487	0.0951	0.0583

C. Performance Assessment

To evaluate the performance of the proposed model, accuracy with standard deviation and sensitivity has been calculated. Accuracy with standard deviation and sensitivity results for binary classificatinons are stated in Table III, Table IV, Table V , Table VI, Table VII & Table VIIII. The following equations have been used for the analysis of the output result

TABLE III: PERFORMANCE RESULT FOR CLASS A-E

Methods	Sensitivity% for A%	Sensitivity% for E	Accuracy% ± Avg STD
Scheme of ref [3] using HOS parameters	76.6	99.3	87.9
Scheme of ref [2] using NIG parameters	97.5	94	95.75
Proposed CNN structure for classes A-E	98.80	100	99.4±0.52

of proposed model [7]:

$$\text{Sensitivity}_i = \frac{TP_i}{TP_i + FN_i} \times 100\% \quad (1)$$

$$\text{Accuracy}_i = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \times 100\% \quad (2)$$

All these are calculated in percentage. In the above equation TP stands for true positive, TN stands for true negative, FP stands for false positive and FN stands for false negative.

D. Comparison of Proposed and Existing Methodology

The methodology used in this paper is very simple CNN architecture . Accuracy result for A-E through the proposed methodology is 99.4% which is an improved result from the EMD based methodology in scheme of reference [3]. Accuracy result for D-E is 99.4% which is higher than the accuracy found through EMD based methodology in scheme of reference [3]. The sensitivity & the standard deviations are presented in Table III to Table VIII.

In some paper, it has been found that dual-tree complex wavelet transform has been used with features obtained from five sub-bands gamma, beta, alpha, theta and delta. And after that separately SVM classifier has to be used for classification [2].

The absorbed methodology in this paper used only single layer of max pooling layer and two layers of fully connected layer. So, definitely the time for running the code gets reduced as well as the complexity of CNN architecture has also been reduced. If we see other works on CNN methodology, it is found that complex CNN architecture are adopted like three or four convolutional layer, three or four FC layer etc [8]. These architectures are time consuming as the computational time and complexity of these architectures are higher than the proposed methodology.

In this paper, the binary classification has been performed for A-C & B-E class. The scheme of reference [2] and the scheme of reference [3] have not been used for the binary classification of A-C & B-E class. But in this paper, all the binary classification has been performed from class A to class E with a better accuracy result.

It is experimented that steps in proposed CNN architecture can differentiate two classes of seizure with minimal time and also replacing the EMD based analysis and wavelet transformation based analysis with better experimental result.

IV. CONCLUSION

We have explored the proposed method for 1-D signal. The computational complexity has been very much reduced through the proposed method of CNN architecture and the accuracy is found satisfactory. So it can be a preferable methodology for seizure classification.

TABLE IV: PERFORMANCE RESULT FOR CLASS B-E

Methods	Sensitivity% for B	Sensitivity% for E	Accuracy% ± Avg STD
Proposed CNN structure for classes B-E	99.60	99.80	99.7 ±0.48

TABLE V: PERFORMANCE RESULT FOR CLASS C-E

Methods	Sensitivity% for C	Sensitivity% for E	Accuracy% ± Avg STD
Scheme of ref [2] using NIG parameters	97.88	95.44	96.66
Proposed CNN structure for classes C-E	98.60	99.80	98.5 ±0.42

TABLE VI: PERFORMANCE RESULT FOR CLASS D-E

Methods	Sensitivity% for D	Sensitivity% for E	Accuracy% ± Avg STD
Scheme of ref [3] using HOS parameters	66.1	88.8	77.4
Scheme of ref [2] using NIG parameters	95.38	96.63	96
Proposed CNN structure for classes D-E	99.20	99.60	99.4 ±0.52

TABLE VII: PERFORMANCE RESULT FOR CLASS A-C

Methods	Sensitivity% for A	Sensitivity% for D	Accuracy% ± Avg STD
Proposed CNN structure for classes A-C	97.60	99.37	98.5 ±0.53

TABLE VIII: PERFORMANCE RESULT FOR CLASS ABCD-E

Methods	Sensitivity% for ABCD	Sensitivity% for E	Accuracy% ± Avg STD
Scheme of ref [3] using HOS parameters	58.8	98.1	90.2
Scheme of ref [2] using NIG parameters	97.13	94.41	94.95
Proposed CNN structure for classes ABCD-E	99.80	98.34	99 ± 0.33

REFERENCES

[1] Truong, Nhan Duy, Anh Duy Nguyen, Levin Kuhlmann, Mohammad Reza Bonyadi, Jiawei Yang, Samuel Ippolito, and Omid Kavehei, "Integer Convolutional Neural Network for Seizure Detection," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems,* 2018.

[2] Das, Anindya Bijoy, Mohammed Imamul Hassan Bhuiyan, and SM Shafiul Alam, "Classification of EEG signals using normal inverse Gaussian parameters in the dual-tree complex wavelet transform domain for seizure detection," *Signal, Image and Video Processing,* 10, no. 2 (2016): 259-266.

[3] Alam, SM Shafiul, and Mohammed Imamul Hassan Bhuiyan, "Detection of seizure and epilepsy using higher order statistics in the EMD domain," *IEEE journal of biomedical and health informatics,* 17, no. 2 (2013): 312-318.

[4] Chowdhury, Tanima Tasmin, Shaikh Anowarul Fattah, and Celia Shahnaz, "Classification of Seizure and Non-Seizure Activity in Seizure Patients Using Time-Frequency Domain Processing of Gamma Band EEG Signals," In *2017 4th International Conference on Advances in Electrical Engineering (ICAEE)*, pp. 537-540. IEEE, 2017.

[5] Vavadi, Hamed, Ahmad Ayatollahi, and Ahmad Mirzae, "A wavelet-approximate entropy method for epileptic activity detection from EEG and its sub-bands," *Journal of Biomedical Science and Engineering,* 3, no. 12 (2010): 1182

[6] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elge,"Indications of nonlinear deterministics and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, *Physical Review E,* vol. 64, no. 6, p. 061907, November 2001.

[7] Qazi Khurram I., H. K. Lam, Bo Xiao, Gaoxiang Ouyang and Xunhe Yin, "Classification of epilepsy using computational intelligence techniques," *CAAI Transactions on Intelligence Technology 1,* no. 2 (2016): 137-149.

[8] Ilah, Ihsan, Muhammad Hussain, and HatimAboalsamh, "An automated system for epilepsy detection using EEG brain signals based on deep learning approach," *Expert Systems with Applications,* 107 (2018): 61-71.

An IoT based Smart Energy Meter for the Idle Use of Electricity: An Industry Perspective

Abdul Kadar Muhammad Masum
Dept. of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh
akmmasum@yahoo.com

Saquib Manzur
Dept. of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh
saquib.manzur@gmail.com

Md. Kalim Amzad Chy
Dept. of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh
kalim.amzad.chy@gmail.com

Shahidul Islam Khan
Dept. of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh
nayeemkh@yahoo.com

Md. Golam RabiulAlam
Dept. of Computer Science and Engineering, BRAC University, Bangladesh
rabiul.alam@bracu.ac.bd

S.M. Taslim Reza
Dept. of Information and Communication Technology, Central Queensland University, Australia
smtaslimreza@gmail.com

***Abstract*—In the perspective of the Bangladesh garment industry, a vast amount of power is wastage due to forgetting in turning off the machines at break time or other reason. Also, the existing prepaid meter doesn't include any system to understand the wastage of the power and control the machines remotely. This research demonstrates a smart energy meter to understand the idle use of the loads and control the loads remotely. This efficient and cost-effective system offers an accurate estimation of the consumed power and cost, reducing the labor cost. A microcontroller is involved in calculating the voltage, current, energy and cost. After deploying in the garment industry, the first few daysthe systemmeasures the consumed amount of power by different loads and store it. From these values, an average value is estimated as a threshold value for that load. In case of exceeding the threshold value of this load, an auto-notification is sent to the authority via a GSM module. Being notified about the excessive usage via that load, the authority confirms about the machine is in working state or idle state through announcing in a microphone. According to the response, the authority takes the decision and sent a specific command through SMS to the system. The GSM module converts the SMS into a signal that is supplied to the microcontroller to turn off the load. In consequence of this feature, the system will assist the authority of the garment industry to be economical in power consumption, and the amount of power wastage will be reduced drastically.**

Keywords—smart energy meter, GSM module, Arduino Uno,

I. INTRODUCTION

Electricity is the prime source of economic development for most of the country. According to the WorldData total consumption of electric energy is 53.65 billion kWh per year where per capitaaverage is 326 kWh in Bangladesh [1].Bangladesh Power Development Board (PDB) stated that 90% of the population had access to electricity[2]. However, per capita energy consumption in Bangladesh is considered very low. Power deficiency is the most significant issue for any developing country in the world. Without proper administration of power use, it is almost difficult to reduce power deficiency, keep up proper billing and increase client satisfaction. A massive amount of electrical energy, a real scenario presented in figure 1, requires for moving forward the garment industry of Bangladesh. Also, lots of energy is wasted just because of the whimsical of the garment worker by keeping the machine on state while it doesn't assign any task.

In this situation, to remotely control the heavyweight machine while it doesn't assign any task, an automated smart energy meter with better service is in need. The proposed GSM based energy meter needs a few days to observe the amount of the power consumption of the industry. Through this observation data, the system generates some threshold value for each load used in the industry. Later, in case of exceeding this threshold value the authority confirm that anyone is working on that area or not. In case of idle use of the load, the authority turns of the machine.A simple #A.sys on* short message and #A.sys off* message is transmitted to control the meter on and off. So, it will assist the authority to subside the wastage of the power and save a lot of money through this system.

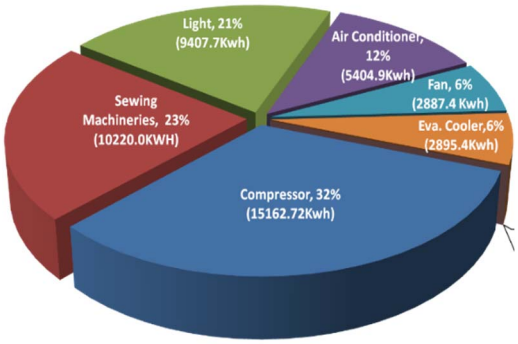


Fig. 1. Power Consumption Scenario inYOUNGONE(CEPZ) Ltd.

The primary objective is to reduce the energy consumption of the garment industries through a smart monitoring system. By monitoring the day to day consumption, the system can easily determine the sudden change in electricity consumption. It will help to avoid any kind of energy wastage in the factory premises.GSM based smart energy meter is already present in the market, but not so cheap and accurate.Also, to show a way of a cheaper alternative smart energy monitoring system is another objective of this research.

The rest of the paper is organized as follows. Section II deliberatesthe related work. In section III methodology including circuit operation, hardware, software, and the

working procedure are explained. Section IV demonstrates an experimental resultwhile section V concludes the paper.

II. Related Work

AMR (Automatic Meter Reading)transfers the unit of used power of the industry or household at a certain period to the central system of a distribution company via wireless communication to produce the monthly bill remotely [3].

The magnitude of the consumed power of various electronic devices is measured in a microcontroller-based smart wattmeter[4]. Digital transformation and signal processing techniques have been utilized to design the meter. As the GSM or GPRS module isn't integrated with this meter, it can't interact with the consumer spontaneously that increases the employment cost. Masum et al. [5] introduced a smart meter for residential customers in Bangladesh that can predict the next day usage by using machine learning regression analysis. In another work, the next day usage had been predicted through Support Vector Regression [6]. Numerous features like meter reading, tempering, power cut information, total used load and disconnect on demand or frequently through SMS alert is combined in an Automatic Meter Reading system for remotely monitoring and controlling purpose [7].The large expense is the major downside of this meter. A summary of previous work has been demonstrated in table 1.

TABLE I. Summary of the Previous Work

Authors & Year	Used Sensor	Methodology	Findings
Riddhi Ghosh [8]	Voltage Sensor & Current Sensor	Smart Energy monitoring through ARM processor	Power quality events can be recorded but no controlling
S.V.Anushree, T.Shanthi[9]	Voltage Sensor & Current Sensor	IOT,DNO, Microcontroller ATMEGA	Electricity theft Prevention
Shivani B. Sonune, R.P. Borole[10]	Current Transformer, Potential Transformer,	Wireless energy metering through radio frequency	Zigbee networking is a cheaper alternative to internet administration for remote control
S Chopra et al. [11]	Voltage sensor & Current Transformer, Peak Detector Circuit	Peak load setting peak load calculation	Meter reading can be collected from remote areas.
Motiur, Fakhrul, Ashiq, Tanver[12]	Current Sensor (ACS712-20A), Optocoupler	IOT, ARDUINO, GSM, and GPRS	Desco meter can be monitored

In the context of Bangladesh, a significant amount of power is waste daily due to the whimsical act of worker, and heavy machines continue to run while it doesn't assign any task at all. Also, there doesn't exist any system to control the load remotely in case of their idle usage. To the best of our insight, this work is the first attempt to subside the wastage of the power of the industry. Through this system, we succeed to minimize a significant amount of energy, showing a case study considering YOUNGONE (CEPZ) Ltd with their legal permission.

III. Methodology

Figure 1 demonstrates a picture of the proposed system how it is designed and worked. After supplying the power, the system is initialized. At the beginning stage, the framework takes a few moments to initialize and estimate the current and voltage. Power is calculated in inductive load as the product of current, voltage and cosine of the angle of the phase difference between current and voltage while in resistive load it is simply a product of current and voltage. There are three types of power which could be measured named real power, apparent power, and reactive power. Among these three types, we have chosen real power consumed in the sewing industry. Real power is measured using equation 1:

$$P = V_{rms} \times I_{rms} \times Cos(Ø)\ Watt \quad (1)$$

Here, P= Power, V_{rms}= Voltage, I_{rms} = Current, Cos(Ø)= Power Factor and Ø= The angle between $V_{rms}$and I_{rms}.

The real power measurement is done accurately anddepending on the load we get either leading or lagging power factor. Determining the phase difference between the line voltage and current is an essential factor to calculate the real power. Ideally, both have zero phase difference, but in practice, they have very minimum phase difference and sometimes depends on the connected loadseither resistive or inductive or capacitive. So, by taking samples within a little period, the total power of a cycle is measured then power is calculated and after all to gain total power average of the complete cycle is undertaken.

Taken Steps in Power Calculation:

1. Taking samples of 60 ms
2. Averaging V_{rms}values and power factor
3. Running these two things for 1 second
4. Averaging I_{rms}values and power factor
5. Calculating power, P= (V_{rms} × I_{rms}× power factor) W
6. Getting energy for this 1 second, E= [{ (P × 1) ÷ 1000 } ÷ 3600]KWh

To measure the total used unit, we will be adding this energy at every second and then we can find out the cost just by multiplying the per unit cost with it. In case of this project simplest power calculation formula has been taken in hand. Power factor has been considered as unity or 1. Only the voltage and current has been considered for the calculation.

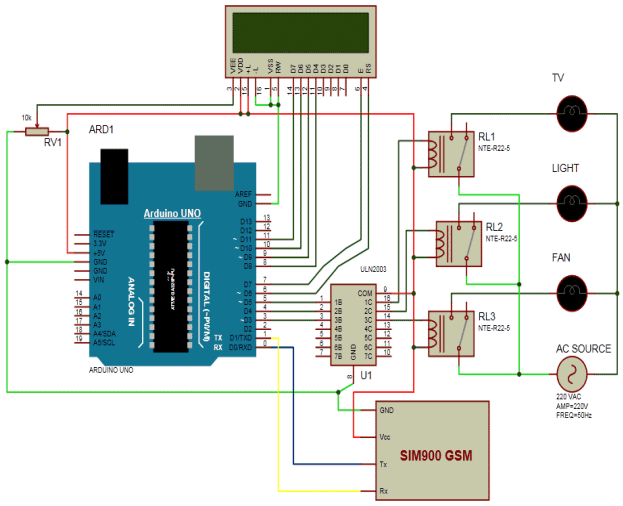


Fig. 2. The proposed system

After successfully determining the total used unit, the system is trained a few days to manipulate the threshold level of the integrated area. The average value of the few days observation is considered as a threshold value. Next time if the system finds out that the total used unit exceeds the threshold level then it notifies the authority through SMS service to take necessary steps. Afterward, the authority makes sure either anyone is working in this area or not. If any response doesn't found from that area, then the authority will send a command to the system through SMS to turn off the load. In case of any response, the authority will decide either he turns off the load or not.

The proposed system is classified into the following parts:

A. *Circuit operation*

The whole system is developed by following the circuit design shown in figure 2. The RX and TX pin of the GSM module relates to the RC6 and RC7 pin of the PIC16F876A respectively while pin 21 to 28 of the PIC16F876A except pin 24 relates to16X2 LCD display. Other loads are attached according to the circuit diagram.

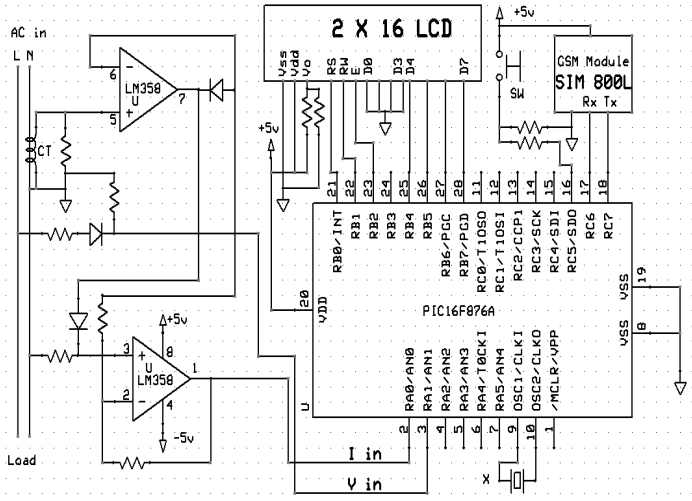


Fig. 3. Circuit design of the proposed framework

B. *Hardware:* Following pieces of equipment are utilized to develop the hardware part.

1)	MCU	PIC16F876A
2)	LCD	2X16 LCD 2 nos
3)	GSM Module	SIM800L/SIM900A
4)	Crystal oscillator	16MHz
5)	Current transformer	50A
6)	Op-Amp	LM358
7)	Diode	IN4007
8)	IC base	28 pin
9)	Resistor	
10)	Capacitor	
11)	Touch switch	
12)	Vero board	
13)	Arduino	UNO/Mega
14)	Relay	One/Two Channel
15)	Bread Board	
16)	Connecting Wires	
17)	Power Supply	

C. *Softwares:*

Arduino IDE and language is used to write the code and burnt into the system that is easy to understand and reliable.

D. *Algorithm:*

At first, the system gets power from the AC power source which can be used from the load or from any other source and Preparation of the system is an event that is taking place after providing the power. In this process, it's read the PIC memory. Then it measures the voltage and current in real time. Voltage can be taken easily for the calculation in the peripheral interface controller. But for taking the current measurement, a peak detector circuit has been used to get the current data easily. A diode has been used to make the alternating current or AC to DC value. After converting the voltage value to DC, the information goes to the PIC for easy calculation. Providing the voltage and current information we get the power and watt-hour using the equation. While calculating the watt-hour, the PIC updates the present current reading and updates the current memory and count day. After taking the average of the few days readings, the system sets a threshold value. From the next day, if the system observes any time the exceeding of the threshold level, it informs the authority through delivering SMS by SIM800L GSM module. If the power consumption does not exceed the threshold level, the system continues to monitor the area. The whole system relates to the relay. In reply to the notification of the exceeding threshold level, the authority confirms either anyone using the load of that area or not. Based upon the confirmation the controller of the load

sends a specific code through SMS. The GSM module converts the code into a signal which proceeds by Arduino. The Arduino sends a signal to the relay to turn off the load and LCD monitor show the Status.

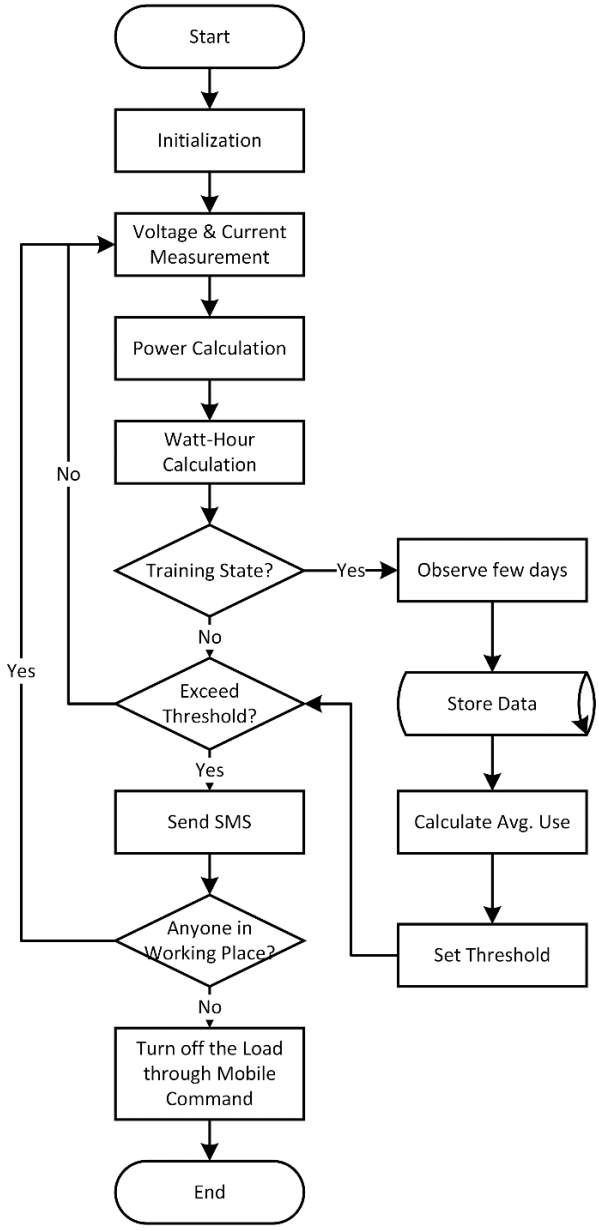


Fig. 4. Smart meter working process

IV. Experimental Results and Discussion

A prototype of the proposed system appears in figure 4 and5 where figure 4 demonstrates the monitoring unit of the system and figure 5 illustrates the controlling unit of the system. In the process of gathering data from actual sewing floor, the developed smart energy monitoring device has been integrated successfully. The data is deduced from office, air condition, light and fan connection. Light and fan connections of the sewing floors are considered for the experiment purpose. This data has been taken for a week. While taking the data, we have checked the amperage rating with an actual amperage meter for any inconsistency. After checking it, we found that it is almost the same.

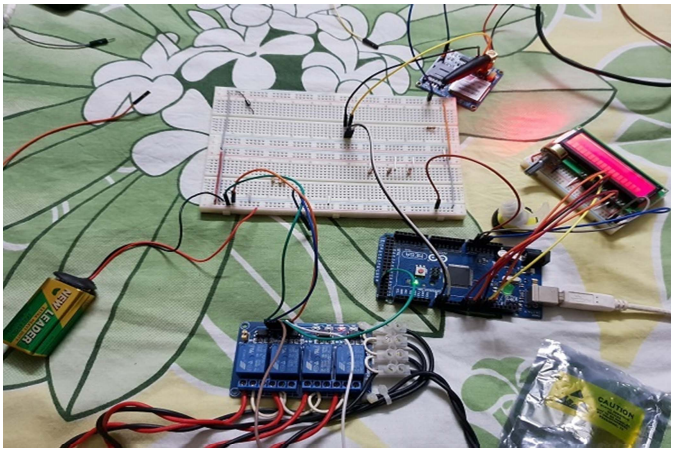


Fig. 5. Monitoring Unit of the system

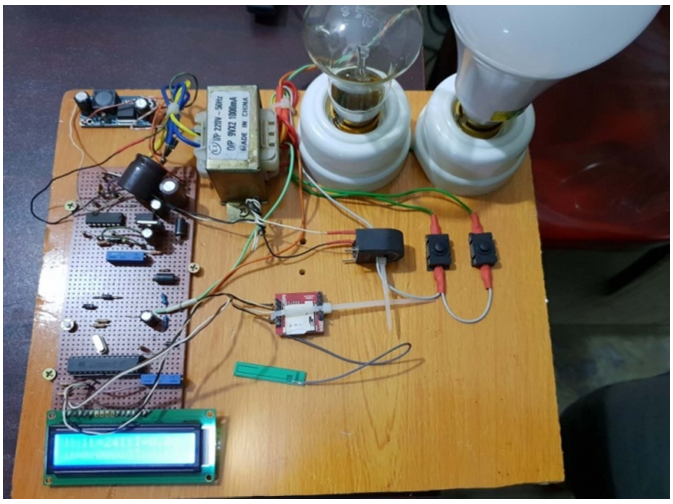


Fig. 6. Controlling Unit of the system

The figure 6 shows a typical sewing floor layout of a garment industry where there have almost 16 lines. From this line, the system works in the red marked lines. Among various loads of a garment industry the lights, fans and air condition of sewing floors are considered.

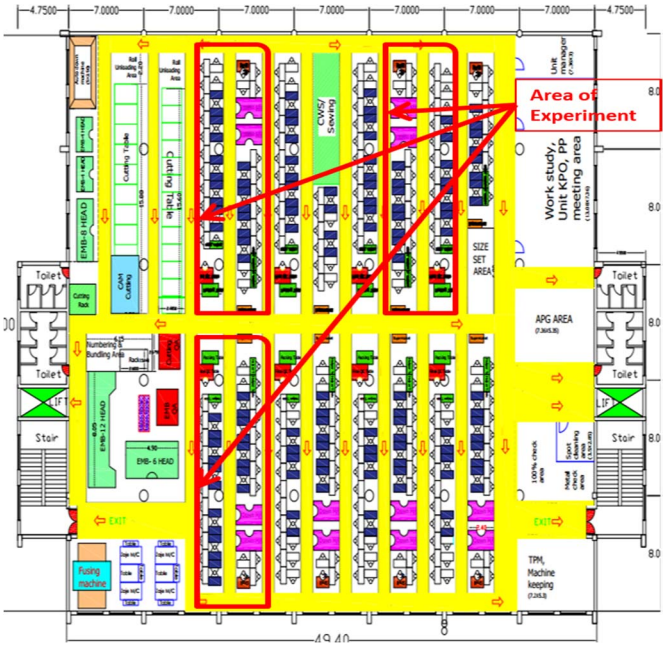


Fig. 7. Typical Sewing Line Layoutof YOUNGONE(CEPZ) Ltd.

Estimated actual data of per day electricity consumptionfor different loads appear in table 2. The sewing floor runs up to 9 hours per day in a regular working hour while 11 hours in case of overtime. During usual working hour all the machines are in operation, but in overtime some machines are inactive. Here considering 11 hours overtime load for sewing is 20.02 kW and by using IOT based smart energy meter, we calculated the same overtime hour load 18.3 kW which gives us 9% saving of energy. In the same way, we can get 10% savings from conventional light and 12% savings from the LED. Also in the case of fans(Korean and Bangla) and AC, we found consequently 7%, 10% and 11% of energy savings.

TABLE II. PER DAY CONSUMPTION OF DIFFERENT LOAD

Area	Light Consumption In total (Per Day)								
	Total Sheds	Watt	Total kW	Usage Calculated for 9 Hours	Cost in tk (per unit 8.99 taka as per BEPZA)	Usage Calculated for 11 Hours	Cost in tk (per unit 8.99 taka as per BEPZA)	After IOT Energy Meter	Percentage of Saving
Sewing	65	28	1.82	16.38	147.2562	20.02	179.9798	18.3	9%
Conventional	32	72	2.304	20.736	186.41664	25.344	227.84256	22.7	10%
LED	65	20	1.3	11.7	105.183	14.3	128.557	12.6	12%
Area	Fan Consumption In Total (Per Day)								
	Total Quantity	Watt	Total kW	Usage Calculated for 9 Hours	Cost in tk (per unit 8.99 taka as per BEPZA)	Usage Calculated for 11 Hours	Cost in tk (per unit 8.99 taka as per BEPZA)	After IOT Energy Meter	Percentage of Saving
Korean	12	40	0.48	4.32	38.8368	5.28	47.4672	4.887	7%
Bangladeshi	12	75	0.9	8.1	72.819	9.9	89.001	8.923	10%
Area	Consumption In Total (Per Day)								
	Total Quantity	Watt	Total kW	Usage Calculated for 9 Hours	Cost in tk (per unit 8.99 taka as per BEPZA)	Usage Calculated for 11 Hours	Cost in tk (per unit 8.99 taka as per BEPZA)	After IOT Energy Meter	Percentage of Saving
AC Ofiice	1 (2ton)	2100	2.1	11.34	101.9466	23.1	207.669	20.445	11%

A comparison in terms of cost between an existing energy meter and our IoT based smart meter for regular light power consumption is shown in table 3. The same information is displayed for LED light, Korean Fan /Bangladeshi Fan intable 4 and 5.A weekly consumption scenario is given in these tables

TABLE III. LIGHT CONSUMPTION IN SEWING FLOOR WEEKLY

Item / Day		1st October 2018	2nd October 2018	3rd October 2018	4th October 2018	6th October 2018	7th October 2018	8th October 2018	Total
IOT	Light	18.056	18.157	18.014	18.856	17.67	18.689	18.782	128.224
Regular		20.178	20.457	20.746	20.111	19.054	20.258	20.347	141.151

TABLE IV. LED LIGHT CONSUMPTION IN SEWING FLOOR WEEKLY

Item / Day		15th September 2018	16th September 2018	17th September 2018	18th September 2018	19th September 2018	20th September 2018	22nd September 2018	Total
IOT	LED Light	12.45	12.137	12.87	12.42	11.804	12.06	12.49	86.231
Regular		14.45	14.85	14.12	14.65	14.68	14.99	14.55	102.29

TABLE V. FAN CONSUMPTION IN SEWING FLOOR WEEKLY

Item / Day		6th September 2018	8th September 2018	9th September 2018	10th September 2018	11th September 2018	12th September 2018	13th September 2018	Total
IOT	Korean Fan	4.882	4.47	4.25	4.31	4.21	4.95	4.87	31.942
Regular		5.28	5.45	5.17	5.46	5.49	6.2	5.78	38.83
Item / Day		29th August 2018	30th August 2018	1st September 2018	2nd September 2018	3rd September 2018	4th September 2018	5th September 2018	Total
IOT	Bangladeshi Fan	8.45	8.41	8.21	8.16	8.78	8.78	8.46	59.25
Regular		9.92	9.78	9.8	9.78	9.65	10.23	9.94	69.1

In the training phase, the average value of a few days' usages is set as a threshold limit. In the second phase, the system will inform the user by delivering an SMS about the current reading in case of crossing the limit. Figure 7 shows a saving scenario by using our proposed smart meter where the red line demonstrates the consumption through normal energy meter, and the blue line demonstrates the consumption through the GSM-based smart energy meter. From the figure, it's recognized that a significant reduction in energy usage is achieved after the use of the smart energy Meter.

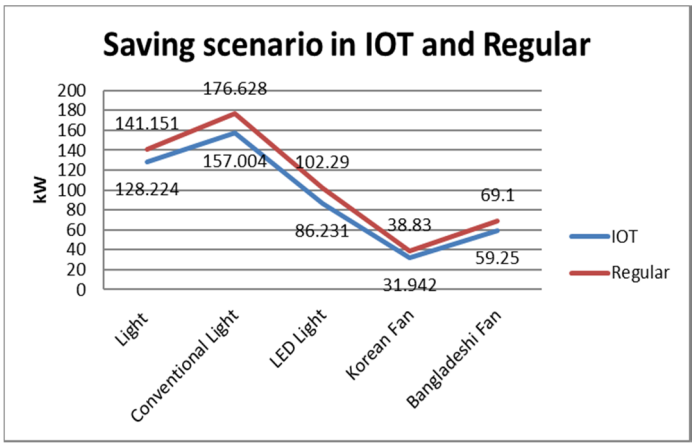


Fig. 8. The saving scenario in the consumption places weekly basis

In table 6 all the electrical loads' savings has been shown in a chart for ease of understanding.

TABLE VI. PER DAY CONSUMPTION OF DIFFERENT LOAD

Consumption From	Total kW	Cost In tk (per unit 8.99 taka as per BEPZA)	After using IoT based Smart Energy Monitoring Total kW	Percentage of savings %
Light (65 sheds)	141.151	1268.94	128.224	9.05%
Conventional Light	176.628	1587.88	157.004	11.11%
LED Light	102.39	919.5871	86.231	15.78%
Korean Fan	38.83	349.0817	31.942	17.73%
National Fan	69.1	621.209	59.25	14.254%

V. CONCLUSION

Notifying and controlling the idle usage of the machine of the garment industry could save a vast amount of power wastage. This study proposes a smart energy meter where the authority is notified about the excessive usage and idle usage of the loads through SMS service and can control the load remotely. Smart controlling of the electrical load can be achieved by sending a simple message to the designated number of the GSM module. Anyone knowing the designated SIM number can operate the load from sitting far away through a specified command. The proposed system is more cost-efficient and smarter than existing energy meter. Tampering the connection of existing meter did not make any good result to read data in the first training step of the proposed meter. It needs to make a separate circuit to measure energy consumption. Smart energy meter can be modified for the detection of illegal use of electricity. In the prototype of the proposed system, at present light, LED light, fan of the sewing floor is considered, and data manually collected from the experimental YOUNGONE(CEPZ) Ltd. area. Controlling the wastage of the heavy weighted machine and developing a data acquisition model to collect the energy consumption data from the loads existing on a single section of a single industry and transmit the data to a centralized server for further processing and analysis is a future scope of our proposed work. Also, other means of communication could be used instead of GSM networks. The authority of the garment industry of Bangladesh should integrate our efficient and effective energy meter to defeat the power wastage problem.

VI. REFERENCES

[1] WorldData. (2019) Energy consumption in Bangladesh. Available: https://www.worlddata.info/asia/bangladesh/energy-consumption.php. [Accessed: Feb. 5, 2019].

[2] Bdnewsnet. (2019). Booming Energy Sector of Bangladesh: 90 Percent Have Access To The Electricity. Available: http://bdnewsnet.com/bangladesh/economy/booming-energy-sector-of-bangladesh-90-percent-have-access-to-the-elctrecity/. [Accessed: Feb. 5, 2019].

[3] S. Shahidi, A. Gaffar, and K. M. Salim, "Design and implementation of digital energy meter with data sending capability using GSM network," in Advances in Electrical Engineering (ICAEE), 2013 International Conference on, 2013, pp. 203-206.

[4] N. Makanjuola, O. Shoewu, L. Akinyemi, and A. Ajasa, "Design and Development of a Microcontroller Based Digital Wattmeter (MIDIWAT)," The Pacific Journal of Science and Technology, vol. 16, pp. 147-158, 2015.

[5] A. K. M. Masum, M. H. Sayeed, M. K. A. Chy, M. T. Hasan, S. M. T. Reza. Design and Implementation of Smart Meter with Load Forecasting Feature for Residential Customers. International Conference on Electrical, Computer and Communication Engineering (ECCE), February 07-09, 2019.

[6] A. K. M. Masum, M. K. A. Chy, M. T. Hasan, M. H. Sayeed, S. M. T. Reza. Smart Meter with Load Prediction Feature for Residential Customers in Bangladesh. International Conference on Energy and Power Engineering (ICEPE), March 14-16, 2019.

[7] M. R. G. Sarangle, U. P. Khot, and J. Modi, "GSM Based Power Meter Reading and Control System," Journal of Engineering Research and Applications (IJERA) ISSN, vol. 2248, p. 9622, 2012.

[8] R. Ghosh, "A Microcontroller Based Single Phase Smart Energy Meter Capable of Classifying Power Quality Events," Citeseer, 2012.

[9] T. Shanthi, S. Anushree, S. Prabha, and D. Rajalakshmi, "DAC to monitor solar-powered home appliances and usage control using bluetooth enabled mobile application and IoT," in 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017, pp. 1-4.

[10] P. D. R. P. B. Shivani B. Sonune, "Wireless Energy Meter Monitoring System with Automatic Tariff Calculation," International Journal of Innovative Research in Computer and Communication Engineering, vol. 4, 2016.

[11] S. Chopra, V. Prasanth, B. El Mansouri, and P. Bauer, "A contactless power transfer—Supercapacitor based system for EV application," in IECon 2012-38th Annual Conference on IEEE Industrial Electronics Society, 2012, pp. 2860-2865.

[12] M. Nayem, M. Rahman, F. I. Mahmud, A.-U. Rahman, and T. Jahan, "Real time monitoring and controlled power distribution system with SMART meter to support smart grid," BRAC Univeristy, 2017.

Data Mining Techniques for Predicting User Interest in Facebook Pages: A Comparison

Nusrat Jahan Farin*, Morium Akter†, Puthi Roy† and Mohammad Shorif Uddin†
*Department of Computer Science and Engineering
Stamford University Bangladesh, Dhaka
†Department of Computer Science and Engineering
Jahangirnagar University, Savar, Dhaka
Email: farin.cse@stamforduniversity.edu.bd, ecs_morium@yahoo.com, puthiroy@gmail.com and shorifuddin@gmail.com

***Abstract*—Now-a-days the use of social media (e.g; facebook, twitter etc.) is increased rapidly. It is rational to exploit interest of users based on their facebook-page activities. However, it cannot be easily predicted manually as there are number of attributes including social interactions. The aim of this paper is to do analysis of 7 data mining algorithms and their variants for determining the best appropriate prediction of people interest in facebook pages. For this, we will analyze and compare various algorithms such as KStar, LWL, IBK, J48, Naive Bayes, Bayes Net, Decision Table. Their performances have been evaluated by using a dataset with 10172 instances of 6 attributes each. Among all the utilized algorithms Bayes Net demonstrates the best execution in regard to the accuracy. However, in respect to the computational time KStar demonstrates the best execution instead of Bayes Net.**

Keywords - data mining, facebook pages, prediction, performance, weka

I. Introduction

Social media turns out to be increasingly more well known in recent days. Individuals utilize online life for various purposes. In web based life, individuals can visit diverse sorts of online business destinations. By these locales advertisers can speak with their clients. Client can put request, think about the items, think about the refreshed market cost of various items and do a lot more things from anyplace at whenever. Web based life is a standout amongst the most critical apparatus to recover information of client enthusiasm for web based business destinations.

Recognizing individuals interest is imperative points for online trade. There exists distinctive sorts of mining algorithms for recognizing the clients and their interests. Clients are verified by various sorts of procedures, for example, PIN, password, user name, email, identification cards and so forth.

Social Networking sites are the massive source for big data that gives the opportunity to study about the human behaviors and interactions in different perspectives. Numerous analysts do their exploration to distinguish the client through the online networking communications. Facebook is a stage that empowers clients to speak with different clients through Internet associations. As indicated by a report of Statistica [1] practically 1.79 billion dynamic clients are utilizing facebook. Consistently 16 percent clients are expanding. Every day 1.18 billion individuals do login to facebook. Research demonstrates that billions of clients every day impart to one another through facebook.

In facebook, clients can communicate by means of post or status, comment, reply, chat and so forth. Individuals additionally speak with one another by means of video call. There is an alternative to respond (e.g, happy, wow, angry etc.) to a post. Clients can make remark and reply to a post. The posts involve with writings as well as suit pictures, or different connections like URLs. To recognize the notoriety of various online business pages of facebook, the posts of the distinctive pages are partitioned into two classifications: textual and interactive. The texual part contains the content, hashtag or URLs of the post and the interactive part contains the like, remark and answer of the specific post. Textual status can be posted in users timeline and the interactive posts can be shared by other users [3, 4]. Users can follow any posts of any page or any user. Clients can follow any posts of any page or any client. To post the status client can utilize hashtag to characterize a specific points [5]. In this paper, we exhibit an instrument to discover the prominence of the facebook pages through data mining algorithms.

The rest of the paper is organized as follows. The following section demonstrates the background study. Section 3 describes the methodology of this paper. In Section 4, experimental results are discussed. The paper is concluded along with the direction of future works in Section 5.

II. Background Study

Identification of people interest is very important, as it finds diverse applications. Many researchers work on finding the human interest in diverse social media [6–8]. Sleeper [9] tried to find out the people interest from facebook posts and likes. There exists some researches to identity the users mentality for business purpose [10]. Through social media usually many things are happening such as message spreading, voting, relationship, personalized advertising [11].

Foster [12] have been studied about to find out the stress, closeness from online interactions. A novel relationship between personal advertising and posts from facebook is presented by Kruike et al. [13]. Scissor [14] extracted the attitudes and behaviors from facebook like's option. Another study [15]

demonstrates the learning interest and satisfaction in social media. Human activities in social media should have some patterns for being used in identifying the user sentiments. This has prompted us to study the users interactions in facebook pages to explore human interests. Many data mining algorithms are available to identify or classify data for finding the patterns. Hence, we took 7 algorithms to classify information of different facebook pages to categorize them. This research gives us the opportunity to figure out the better algorithm for extracting the people's interest from online interactions.

In this study, 7 unique kinds of data mining strategies are examined. The strategies are KStar, Decision Table, J48, Naive Bayes, Bayes Net, IBK, LWL. Scientists work with these data mining systems for various issues [16, 17]. A case based languid learning classifier, KStar [18], stores the preparation information and after that performs handling to discover the test information. KStar utilizes entropy based separation which varies it from the other occurrence based learners [19]. It is like the K-Nearest-Neighbor classifier [20].

IBK is a k-closest neighbor classifier. In this calculation an item is characterized by a greater part vote of its closest neighbors [21]. Boyer-Moore lion's majority vote calculation is utilized to discover most of a grouping of components linear time in constant space. Vote is a prototypical case of a streaming algorithm [22].

Decision Table [23] is appropriate for test knowledge exploration. The tree structure relies upon the information which performs forecast. In this model, result esteems are spoken to with leaves and characteristics of datasets are known as branches. Decision Tree does not settle on any choice in data mining instead of it helps for settling on choice. There exists varieties in Decision Tree, for example, CART, CHAID, J48, ID3, C4.5 and so forth. In our research, J48 calculation is utilized. J48 enables client to anticipate the target variable of another dataset record [24, 25].

Naive Bayesian classifier utilizes the Bayes restrictive likelihood rule for performing order, expecting traits are measurably autonomous. The word Naive methods qualities of the informational index are considered as autonomous [26].

Bayes Nets or Bayesian systems [27] are graphical portrayal of probabilistic connections. It makes relations among a lot of self-assertive factors. A Bayesian system is deciphered with Directed Acyclic Graph (DAG) that encodes a joint likelihood dissemination [28]. It utilizes the likelihood capacity to anticipate distinctive data.

III. Proposed Methodology

In this paper, user interest in online social media is identified via user interactions on facebook page. Our proposed methodology involves the following steps.

1) Acquire posts from selected facebook pages.
2) Extract interactive data such as likes, comments, replies, posts name, and URLs (page names), texts from posts.
3) Analyze acquired data to explore human behavioral interest.
4) Apply matching techniques on training and test data set for identifying the people interests.

We have conducted experiment with the collected dataset for different data-mining techniques. The experiment is performed on the basis of test mode of k-fold cross-validation with supplied test sets (where k=10). For the performance analysis, some performance factors are focused like predictive accuracy and model building time, true positive (TP) rate, false positive (FP) rate, F-measure, etc. The flow chart of this presented study is given in Fig. 1.

A. Data Acquisition

We manually collected 10172 posts from 20 facebook pages over a period of more than three months. Initially, 50 active pages, where both Bengali and English languages are used, from facebook are selected on a random basis. Among them, the most prolific 20 pages are identified who produces more than 20 posts per week. All the data are collected manually and no tools are used to collect the data. Some sample data are presented in Table I. We have divided our data (posts) into five sessions. Each record (post) comprises of total 6 attributes: likes, comments, replies, post title, page likes and page name along with texts. Collected data set are available in our dropbox link [2]. The attributes are collected on the basis of the user input. Mainly 6 attributes are more effective for showing peoples interest and interaction in facebook pages. By analysing those attributes the popular or more interactive post can be identified easily. The collected attributes help to identify the user interest by classifying inputed data.

We extracted the top 50 interests (e.g; Interior decorating, Painting, Shopping, Cooking, Leather, Devices, Gaming, Cricket, Football, Travelling etc.) based on occurrence in the facebook pages from 10172 random posts. We shorten the list of interests by excluding any interest that occurred within multiple post in different facebook pages. By using different types algorithm we can identify people interest. All the analysis of predicting user interest are shown in result section.

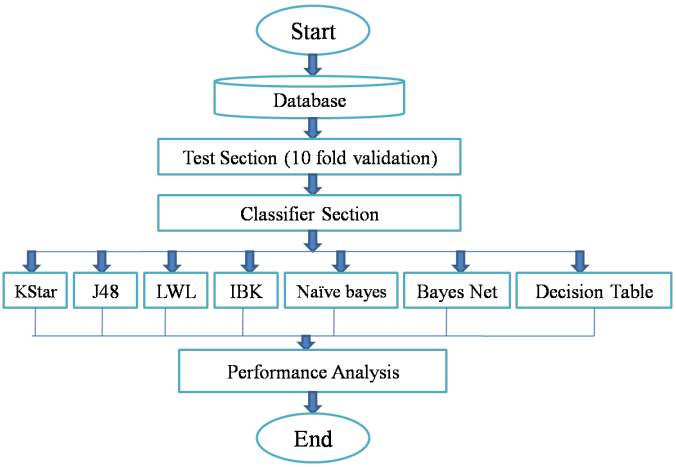


Fig. 1. Data mining classification process

B. Preparing dataset

We converted the dataset from CSV (Comma Separated Value) format to ARFF (Attribute Relation File Format) which can be easily uploaded in Weka software [29].

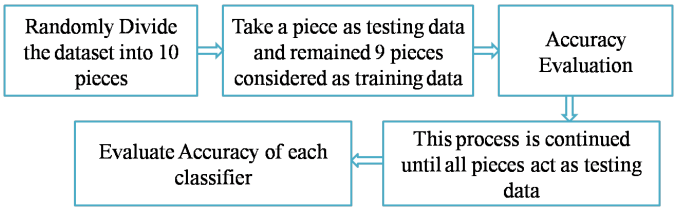


Fig. 2. 10 cross-validation process for our experiment

C. Experiment using Weka software

We investigated each mining algorithm for the prepared dataset. The investigation process is shown in Fig. 1. Cross-validation divide a dataset into 10 pieces (folds). Moreover, hold out each piece in turn as testing and train on the remaining 9 pieces together. This gives 10 evaluation results, which are averaged. By doing 10-fold cross-validation and then computed the evaluation results, Weka invokes the learning algorithm 11th time on the entire dataset to obtain the model [30]. Fig.2 shows the graphical view of cross-validation process that has been used in the current experiment. In this study, different accuracies are found for different classifiers. All information about the classifictaion results of classifiers are collected from Weka software. Performance of all classifiers have been compared and shown in Table II.

IV. Results and Discussions

In this study, We ran 7 separate experiments to decide how well we can foresee individuals performance for various facebook pages utilizing diverse data mining procedure. We ran the majority of our experiments in weka utilizing 7 types classifiers (e.g; KStar, LWL, IBK, J48, Naive Bayes, Bayes Net and Decision Table) more than 5 times running of 10-overlay cross-approval.

Confusion matrix along with computation time gives the entire information about correctly and incorrectly classified instances for each classifier. Based on the data, we consider the following statistical measures such as TP rate, FP rate, Precision, Recall, F-Measure, Accuracy for comparative performance analysis of different data mining algorithms. These statistical measures are calculated through equation $1-6$. Equation 1 helps to find the accuarcy of the classifiers that used in this experiments.

$$Accuracy = \frac{\sum TruePositive + \sum TrueNegative}{\sum TotalInstances} \quad (1)$$

$$TruePositiveRate(TPR) = \frac{TP}{TP + FN} \quad (2)$$

$$FalsePositiveRate(FPR) = \frac{FP}{FP + TN} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$FMeasure = \frac{precision.recall}{precision + recall} \quad (6)$$

Where, TP, FP, TN, FN denotes True Positive, False Positive, True Negative, False Negative respectively.

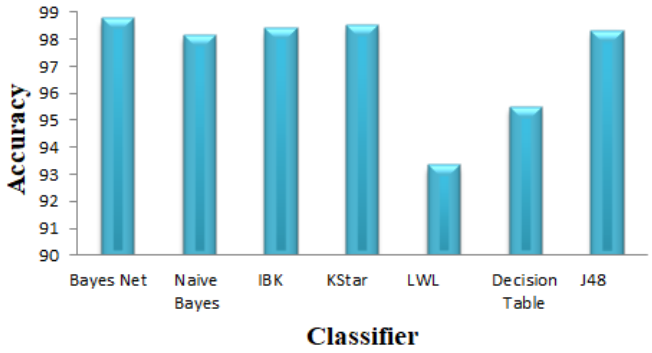


Fig. 3. Accuracy of classifiers

As accuracy is an essential execution parameter, so we exhibited the accuracy of all classifiers graphically in Fig. 3. KStar, IBK, J48, Naive Bayes, Bayes Net perform superior to different classifiers. Table II demonstrates the point by point accuarcy, TP rate (TPR), FP rate (FPR), Precision, F-Measure. TPR, FPR is calculated through equation 2 and equation 3. Moreover, Presicion, F-Measure and Recall are derived from equation 4, equation 5 and equation 6. From Table II we locate that the greater part of the utilized classifiers false positive rate is low and genuine positive rate is very significant. By consolidating the TPR, FPR, precision, recall, F-measure and accuracy, we find that bayes and tree learners give the preferred outcome over the others do. Bayes Net effectively grouped 10019 cases; be that as it may, it can't characterize 103 occurrences accurately. Naive Bayes classified correctly 9966 instances and incorrectly 206. Whereas, other used classifiers as IBK, Kstar, LWL J48 and Decision Tree correctly classified 9989, 10003, 9460, 9663, 9980 instances respectively. KStar gives the best performance other than 7 classifiers concerning computational time and accuarcy.

V. Conclusion and Future Work

The presented paper focuses to find the correct data mining calculation to predict the general population interest for various facebook pages as indicated by info properties. It cannot be effectively anticipated physically as it is troublesome assignment which requests higher learning and aptitude. In this paper, we have made a thorough relative investigation of 7 diverse order calculations, for example, KStar, LWL, IBK, J48, Naive Bayes, Bayes Net, Decision Table. Their results have been assessed through Weka programming by utilizing

TABLE I
SAMPLE DATASETS

Page Name	Title of Post	No. of Likes	No. of Comments	No. of Replies	Page Likes
Joblisting.com	Management Trainee Officer	91	0	0	89562
BBC News	This massive sinkhole	610000	686	0	33143694
Berger Paints Bangladesh Limited	Visit us at Furniture	184	75	0	222235
Studying in Germany	Learning German	3.5k	239	0	455323
SYMPHONY	Symphony P7	4400	372	106	1471671
Architecture Fashion Design	Wow	6000	205	0	1123987
PRAN UP	Chittagong Vikings	227000	3030	876	1422125
Selena Gomez	Walks through Calgary	766000	4336	0	61632168
Tahsan	Clothes don't make a man	132000	1545	0	6185464
Grameenphone	Grameen Phone gives 2 GB	40000	1723	876	8828070
Robi Axiata Limited	Back in action!	9600	293	76	8558453
Theatre Live	Live at 2.00 PM	2217	59	9	37950
Mashrafe Bin Mortaza	Today Bangladesh captain	144000	1160	0	7905189

TABLE II
PERFORMANCE MEASURE OF CLASSIFIERS

Classifier	TP Rate	FP Rate	Precision	Recall	F-Measure	Accuracy	Computational Time(s)
Bayes Net	0.986	0.001	0.987	0.986	0.986	98.5727	0.15
Naive Bayes	0.986	0.001	0.987	0.986	0.986	97.9742	0.01
IBK	0.982	0.001	0.984	0.982	0.982	98.21	0.02
KStar	0.983	0.001	0.985	0.983	0.983	98.34	0
LWL	0.931	0.005	0.936	0.931	0.916	93.14	0
Decision Table	0.953	0.003	0.928	0.953	0.937	95.3039	0.05
J48	0.981	0.001	0.986	0.981	0.980	98.1123	0.12

gathered dataset containing 10172 records. In view of the execution factors KStar, Naive Bayes, Bayes Net, J48, IBK work practically comparable and give the preferable execution over other grouping techniques. KStar demonstrates the best execution over the 7 classifiers concerning computational time and accuarcy.

In future, we intend to research individuals' enthusiasm to look for progressively precise expectation and make a regressional investigation to proposed the connection between computational time and accuracy.

REFERENCES

[1] The top 20 valuable facebook statistics updated november 2016, Available online, https://zephoria.com/top-15-valuable-facebook-statistics/, Accessed on November 15, 2016.

[2] Collected dataset, Available in, https://www.dropbox.com/s/zsd42tiiok4htsb/data.csv?dl=0

[3] Wenhong Chen and Kye-Hyoung Lee, "Sharing, liking, commenting, and distressed? the pathway between facebook interaction and psychological distress", *Cyberpsychology, Behavior, and Social Networking*, 16(10):728–734, 2013.

[4] Bimal Viswanath, Alan Mislove, Meeyoung Cha, and Krishna P Gummadi, "On the evolution of user interaction in facebook", *Proceedings of the 2nd ACM workshop on Online social networks*, pages 37–42. ACM, Barcelona, Spain, August 17, 2009.

[5] Bjarki Valtysson, "Facebook as a digital public sphere: Processes of colonization and emancipation", *tripleC: Communication, Capitalism & Critique, Open Access Journal for a Global Sustainable Information Society*, 10(1):77–91, 2012.

[6] Michel Ballings, Dirk Van den Poel, and Matthias Bogaert, "Social media optimization: Identifying an optimal strategy for increasing network size on facebook", *Omega*, 59:15–25, 2016.

[7] Wenzhi Zheng, Chih-Hung Yuan, Wei-Hung Chang, and Yen-Chun Jim Wu, "Profile pictures on social media: Gender and regional differences", *Computers in Human Behavior*, 63:891–898, 2016.

[8] Derek Summerfield, "The invention of post-traumatic

stress disorder and the social usefulness of a psychiatric category", *The Bmj*, 322(7278):95–98, 2001.

[9] Manya Sleeper, Rebecca Balebako, Sauvik Das, Amber Lynn McConahy, Jason Wiese, and Lorrie Faith Cranor, "The post that wasn't: exploring self-censorship on facebook", *Proceedings of the ACM conference on Computer supported cooperative work*, Texas, USA, February 23-27, 2013, pages 793802.

[10] Han K Kang, Benjamin H Natelson, Clare M Mahan, Kyung Y Lee, and Frances M Murphy, "Post-traumatic stress disorder and chronic fatigue syndrome-like illness among gulf war veterans: a population-based survey of 30,000 veterans", *American journal of epidemiology*, 157(2):141–148, 2003.

[11] Linchi Kwok and Bei Yu, "Spreading social media messages on facebook an analysis of restaurant business-to-consumer communications", *Cornell Hospitality Quarterly*, 54(1):84–94, 2013.

[12] Jamie E Foster and Allison R Thorson, ""i'm too stressed to 'like' your post!": Exploring the connections among facebook relational maintenance, stress, and closeness", *Iowa Journal of Communication*, Vol. 48, Issue 1/2, p76-95, 20p, 2016.

[13] Sanne Kruikemeier, Minem Sezgin, and Sophie C Boerman, "Political microtargeting: Relationship between personalized advertising on facebook and voters' responses", *Cyberpsychology, Behavior, and Social Networking*, 19(6):367–372, 2016.

[14] Lauren Scissors, Moira Burke, and Steven Wengrovitz, "What's in a like?: Attitudes and behaviors around receiving likes on facebook", *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, ACM, San Francisco, CA, USA, 2016, pages 1501–1510.

[15] Jon-Chao Hong, Ming-Yueh Hwang, Elson Szeto, Chi-Ruei Tsai, Yen-Chun Kuo, and Wei-Yeh Hsu, "Internet cognitive failure relevant to self-efficacy, learning interest, and satisfaction with social media learning", *Computers in Human Behavior*, 55:214–222, 2016.

[16] Jiawei Han, Jian Pei, and Micheline Kamber, *"Data mining: concepts and techniques"*, Elsevier, 2011.

[17] Marjia Sultana, Afrin Haider, and Mohammad Shorif Uddin, "Analysis of data mining techniques for heart disease prediction", *Electrical Engineering and Information Communication Technology (ICEEICT), 3rd International Conference on*, IEEE, Military Institute of Science and Technology, Dhaka, Bangladesh, 2016, pages 1–5.

[18] Eibe Frank, Mark Hall, Geoffrey Holmes, Richard Kirkby, Bernhard Pfahringer, Ian H Witten, and Len Trigg, "Weka-a machine learning workbench for data mining", *Data mining and knowledge discovery handbook*, pages 1269–1277, Springer, 2009.

[19] John G Cleary, Leonard E Trigg, et al, "K*: An instance-based learner using an entropic distance measure", *Proceedings of the 12th International Conference on Machine learning*, volume 5, pages 108–114, Miami, FL, USA, 1995.

[20] Padraig Cunningham and Sarah Jane Delany, "k-nearest neighbour classifiers", *Multiple Classifier Systems*, 34:1–17, 2007.

[21] David Frederick Cutler and Ian BK Richardson, *"Tree roots and buildings"*, Longman Group Ltd., 1989.

[22] W Nick Street and YongSeog Kim, "A streaming ensemble algorithm (sea) for large-scale classification", *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 377–382, ACM, San Francisco, CA, USA, 2001.

[23] Jafar Tanha, Maarten van Someren, and Hamideh Afsarmanesh, "Semi-supervised self-training for decision tree classifiers", *International Journal of Machine Learning and Cybernetics*, 8(1):355–370, 2017.

[24] Richard J Roiger, "Data mining: a tutorial-based primer", pages 1112–1117, Available online, https://www.questia.com/library/journal/1P3-930077291/data-mining-a-tutorial-based-primer.

[25] Eibe Frank, *"RepTREE. java"*, Revision, 1999, Available online, http://weka.sourceforge.net/doc.dev/weka/classifiers/trees/REPTree.html.

[26] Kevin P Murphy, *"Naive bayes classifiers"*, 2006, Available online, https://en.wikipedia.org/wiki/Naive_Bayes_classifier, Accessed on March 15, 2017.

[27] Robert Fung and Kuo-Chu Chang, "Weighing and integrating evidence for stochastic simulation in bayesian networks", *arXiv preprint arXiv:1304.1504*, 2013.

[28] Finn V Jensen, *"An introduction to Bayesian networks"*, volume 210, UCL press London, 1996.

[29] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H Witten, "The weka data mining software: an update", *ACM SIGKDD explorations newsletter*, 11(1):10–18, 2009.

[30] Kohavi, Ron "A study of cross-validation and bootstrap for accuracy estimation and model selection" *InIjcai*, 1995 Aug 20 (Vol. 14, No. 2, pp. 1137-1145).